

Another Member of Vasco-Caucasian?

There is another mysterious “isolate” language spoken high in the mountains of northeastern Pakistan by some 60,000 people. This is the language formerly known by names such as Kanjut, Khajuna or Werchikwar, and now usually as Burushaski. It is spoken in the Hunza, Nager and Yasin valleys, a mountainous environment best known in the West as a legendary haven of health and long life. The possible connections of this language with Basque and the Caucasian languages have been explored throughout the twentieth century, by, for example, the scholars Robert Bleichsteiner, Karl Bouda, and Hermann Berger. The following table lists some of the most promising proposed cognates. Footnotes indicate meanings that vary from the canonic meaning in the left column, dialectal provenance of words cited, and other notes.

This table only lists some of the cognates common to all three branches. Other possible lists would show cognates common only to Basque and Burushaski, or to Caucasian and Burushaski. Again, as in the Basque-Caucasian table, most of the words are basic, only the last two (‘thread’ and ‘house’) having any cultural content. The words for ‘milk’, ‘goat’ ~ ‘sheep’, and ‘horse’ ~ ‘donkey’ may bear witness to the beginnings of animal husbandry before the three groups diverged.

And what of the tests applied earlier to the Basque-Caucasian case, phonology and grammar? Rather than discussing the phonology in detail, as we did in the Basque-Caucasian case, I will just mention a few of the correspondences. Some “trivial” correspondences ($b = b = b$; $n = n = n$) are readily apparent in several of the comparisons. More interesting are comparisons such as these:

(‘stomach’)	Bur. <i>-phaṭ</i>	= Bsq. <i>eperdi</i>	= Cauc. <i>*pHVrṭwV</i>
(‘slime’)	Bur. <i>giṭ</i>	= Bsq. <i>lerde</i>	= Cauc. <i>*λwirdi</i>
(‘narrow’)	Bur. <i>čhaḍ-úm</i>	= Bsq. <i>estu</i>	= Cauc. <i>*čHVrdV</i>

In these words Burushaski has **retroflex** plosives – voiceless /t/ in word-final position, voiced /ḍ/ between vowels. The contrast between **dental** plosives /t/, /d/ and **retroflex** plosives /t/, /ḍ/ is typical of languages in the South Asian (Greater Indian) linguistic area. It is found in Indo-European (Indo-Aryan) languages such as Hindi and Bengali (and ancient Sanskrit), in some Iranian languages (such as Pashto), in Dravidian languages such as Tamil and Malayalam, and even in some Munda (Austroasiatic) and Sino-Tibetan languages of the region. Therefore it is thought that Burushaski, like Indo-Aryan, acquired the contrast through contact with languages that were in India before them. In the three words shown above it appears that the Burushaski retroflex plosives developed out of the older clusters of /r/ + /t/ or /r/ + /d/, which are preserved as such in Basque and Proto-Caucasian. This type of change also took place in Indo-Aryan languages, for example, Hindi *āṭh* ‘eight’ < Old Indic *aṣṭā* < PIE **oktō*; Sanskrit *nīḍá* ‘resting place’ <

Table

(gloss)	Burushaski ¹	Basque	Caucasian ² (modern)	Proto-Caucasian
‘face’	<i>buk</i> ³	<i>beko</i> ⁴	<i>buk</i> ⁵	* <i>bč̣kwo</i>
‘stomach’	<i>-phať</i> ⁶	<i>eperdi</i> ⁷	<i>p̄art</i> ⁸	* <i>pHVr̄t̄wV</i>
‘abdomen’	<i>-úl</i>	<i>urdail</i> ⁹	<i>b-etl’ tl’ u</i> ¹⁰	*= <i>iraĹV</i>
‘back’	<i>-sqa</i> ¹¹	<i>bizkar</i>	<i>azkwa</i> ¹²	* <i>zok^wa</i> ¹³
‘fur’	<i>bišké</i>	<i>bizar</i> ¹⁴	<i>bizal-ba</i> ¹⁵	* <i>bil̄ǵV</i>
‘pus, tears’	<i>nagéi</i> ¹⁶	<i>negar</i> ¹⁷	<i>not’ q’ a</i> ¹⁸	* <i>nč̣wq̄ũ</i>
‘slime’	<i>giť</i>	<i>lerde, lirdi</i> ¹⁹	<i>x^werd</i> ²⁰	* <i>ł^wirdi</i>
‘milk’	<i>šiŋ</i>	<i>ezne</i>	<i>šin</i> ²¹	* <i>š̄am̄V</i>
‘goat’	<i>tshigir</i> ²²	<i>zikiro</i> ²³	<i>ts’ ts’ ik’ er</i> ²⁴	* <i>zīkV</i>
‘horse’	<i>čhardá</i> ²⁵	<i>asto</i> ²⁶	<i>čada</i> ²⁷	* <i>čada</i> ²⁸
‘rodent’	<i>čargé</i> ²⁹	<i>sagu</i> ³⁰	<i>sok</i> ³¹	* <i>cārgwā</i>

¹ Dialects are designated as Hunza (H), Nager (N), and Yasin (Y).

² In this column representative words are cited from actual Caucasian languages. The next column is made up of reconstructions, usually Proto-Caucasian, but occasionally lower-level reconstructions, as indicated in footnotes.

³ ‘throat’; cf. meanings in Basque and Caucasian: ‘mouth’ in Nakh languages, ‘muzzle’ in Tsakhur, ‘lip’ in Kryz and Budukh. And cf. the IE cognates Eng. *mouth* = Latin *mentum* ‘chin’, for the type of semantic shift.

⁴ ‘forehead’

⁵ Rutul ‘part of face’ (lips + nose)

⁶ (Y) *phať* ‘stomach (of fowl)’, (H,N) *-phať* ‘viscera (of fowl)’. [t] is a retroflex plosive.

⁷ ‘rump, buttocks’

⁸ Archi ‘large intestine’

⁹ ‘stomach, abomasum, rennet’

¹⁰ ‘stomach, abomasum, rennet’

¹¹ ‘on one’s back’ – must be used with a possessive prefix, such as *á-sqa* ‘on my back’.

¹² Abkhaz

¹³ Proto-Abkhaz-Tapant (a subgroup of West Caucasian)

¹⁴ ‘beard’

¹⁵ Bezhta ‘mustache’

¹⁶ (H) *nagéi*, (N) *magéi* ‘boil, ulcer’

¹⁷ ‘tears, weeping’

¹⁸ Chechen ‘pus’

¹⁹ ‘drivel, saliva’

²⁰ Avar ‘pus’

²¹ Chechen ‘udder’

²² (N) ‘(she-)goat’

²³ ‘castrated ram’

²⁴ Karata ‘kid’ (young goat)

²⁵ (H,N) *čhardá*, (Y) *čardé* ‘stallion’

²⁶ ‘ass, donkey’

²⁷ Abkhaz ‘ass, donkey’

²⁸ Proto-Abkhaz-Tapant (a subgroup of West Caucasian)

²⁹ (Y) ‘flying squirrel’

³⁰ ‘mouse’

(gloss)	Burushaski ³²	Basque	Caucasian ³³ (modern)	Proto-Caucasian
‘flea’	<i>khin, khen</i> ³⁴	<i>akain</i> ³⁵	<i>gēnig</i> ³⁶	*ǵānʔV
‘fruit’	<i>gaiŋ</i> ³⁷	<i>ar(h)an</i> ³⁸	<i>gēni</i> ³⁹	*yōnʔV
‘forest’	<i>hun</i> ⁴⁰	<i>oihan</i>	<i>ħun</i> ⁴¹	*fānV
‘stone’	<i>xóro</i> ⁴²	<i>harri</i>	<i>χirχem</i> ⁴³	*χHērχV
‘slope’	<i>hurgó</i> ⁴⁴	<i>hegi</i> ⁴⁵	<i>urqi</i> ⁴⁶	*ħwōrqē
‘fire’	<i>ši</i> ⁴⁷	<i>su</i>	<i>ts’u</i> ⁴⁸	*čājš
‘bitter’	<i>gaqáy-um</i> ⁴⁹	<i>kharats</i> ⁵⁰	<i>q’ala</i> ⁵¹	*ǵčhIV
‘narrow’	<i>čhađ-úm</i> ⁵²	<i>estu</i>	<i>č’art’a</i> ⁵³	*čHVrdV
‘to know’	<i>-ki-</i> ⁵⁴	<i>jakin</i> ⁵⁵	<i>=iq’</i> ⁵⁶	*=iqE
‘thread’	<i>gay, hari</i>	<i>hal-</i>	<i>χal</i> ⁵⁷	*χāfV
‘house’	<i>-yeš</i> ⁵⁸	<i>etxe</i> [eče]	<i>ts’a</i> ⁵⁹	*č[ī]ju

³¹ Tsakhur ‘weasel’

³² Dialects are designated as Hunza (H), Nager (N), and Yasin (Y).

³³ In this column representative words are cited from actual Caucasian languages. The next column is made up of reconstructions, usually Proto-Caucasian, but occasionally lower-level reconstructions, as indicated in footnotes.

³⁴ (H,N) *khin*, (Y) *khen*

³⁵ ‘(large) tick’

³⁶ Chechen ‘louse’

³⁷ ‘grapes’

³⁸ ‘plum, sloe’; I analyze the word as *ar-han*, the first element being a fossilized class prefix, as seen in Tibetan *r-gun* ‘grape, vine’

³⁹ Avar ‘pear’

⁴⁰ (Y) ‘wood, firewood’, (H,N) ‘wood, timber, beam, hewn trunk’

⁴¹ Chechen ‘forest’

⁴² (Y) ‘small stones’

⁴³ Lezgi ‘small stones, gravel’

⁴⁴ ‘steep slope, uphill, up the mountain’

⁴⁵ ‘top, summit, ridge; border, corner, edge’

⁴⁶ Dargwa (Akushi) ‘mountains’

⁴⁷ (H,N) ‘fireplace, hearth’

⁴⁸ Lak

⁴⁹ (H,N) *gaqáy-um*, (Y) *qaqám* ‘bitter’

⁵⁰ Zuberoan ‘bitter, sour’

⁵¹ Archi

⁵² (N) *čhađ-úm*; (H) *čhan-úm*, (Y) *čan-úm* (influenced by the verb *du-čhan*?)

⁵³ Dargwa (Akushi)

⁵⁴ ‘to learn’

⁵⁵ ‘to know (a fact)’ (< **e-aki-n*)

⁵⁶ Bezhta ‘to know’

⁵⁷ Chechen ‘a piece of thread’

⁵⁸ ‘dauernder Wohnsitz; Seßhaftigkeit, Beständigkeit’ (H. Berger: ‘lasting dwelling-place; sedentariness, permanence’).

⁵⁹ Chechen ‘house, room’

Indo-Iranian **niždas* < PIE **nizdó-s* ‘sitting place, nest’ (cf. Lat. *nīdus*, Lith. *lizdas*, Arm. *nist*, Eng. *nest*, etc.); in the Vedic texts this word has instead a retroflex lateral /l/: *nīlá*.

Another interesting non-trivial correspondence is seen in the following comparisons:

(‘bitter’)	Bur. <i>gaqáy</i> -um	= Bsq. <i>k(h)arats</i>	= Cauc. <i>*q̣čfiIV</i>
(‘thread’)	Bur. <i>gay</i>	= Bsq. <i>hari</i>	= Cauc. <i>*χāfV</i>

Burushaski /ý/ is a very unusual and peculiar sound⁶⁰ heard – as far as I know – only in two of the three dialects of Burushaski (Hunza and Nager), and in Dumaki, an Indo-Aryan language spoken in the same area (Gilgit). The correspondence of Bur. /ý/ = Bsq. /r/ = Cauc. **l* ~ **ʃ* may also be verified by:

(‘sheep’)	Bur. <i>huyóo</i>	= Bsq. <i>ahari</i> ‘ram’	= Cauc. <i>*ʃlχU</i> ‘sheep, lamb’
(‘leaf’)	Bur. <i>khiy</i>	= Cauc. <i>*kəwʃV</i> ‘sheaf’	
(‘stones’)	Bur. <i>qhiyé</i>	= Cauc. <i>*q̣wiθ</i> ‘rock, cliff, stone’	

Apart from the above, other evidence shows that Bur. /ý/ comes from an /l/-like sound (**lateral**). For example, the place name *Námay* is otherwise known as Nomal (a village between Gilgit and Hunza).

These comparisons exemplify another recurrent pattern:

(‘rodent’)	Bur. <i>čargé</i>	= Bsq. <i>sagu</i>	= Cauc. <i>*cārgwī</i>
(‘slope’)	Bur. <i>hurgó</i>	= Bsq. <i>hegi</i>	= Cauc. <i>*fiwərgē</i>

The rule can be stated as “Bur. /rg/ = Bsq. /K/ = Proto-Cauc. **rK*,” where **K* represents “velar or postvelar plosive.” At least two other Basque-Caucasian comparisons confirm this pattern:

(‘border’)	Bsq. <i>muga</i>	= Cauc. <i>*mōrqwǃ</i> ‘stripe, line’
(‘temple’)	Bsq. <i>loki</i>	= Cauc. <i>*ʎarqwě</i> ‘forehead, cap’

⁶⁰ The Norwegian linguist Morgenstierne described /ý/ (which he transcribed as /i/) as “a fricative *r*, pronounced with the tongue in the retroflex ... position.” The American linguist Greg Anderson describes /ý/ as “a curious sound whose phonetic realizations vary from a retroflex, spirantized glide, to a retroflex velarized spirant.” Besides /ý/ and /i/, it has also been transcribed as /ly/ and /gh/. Hermann Berger finds /ý/ similar to the Tamil sound commonly transcribed as /ɻ/.

Here Basque is least conservative, losing /r/ before velar plosives, while in the case of /rd/, above, Burushaski was innovative in merging /rt/, /rd/ as /t/, /d/. (Note that some Caucasian languages – independently but convergently – have undergone the same change as Basque, for example, Chechen *moga* ‘line, row’, Khwarshi *tl'oq'o* ‘forehead’.)

These comparisons show yet another correspondence:

(‘fire’)	Bur. <i>ši</i>	=	Cauc. *čǎǰǰ
(‘house’)	Bur. <i>-yeéš</i>	=	Cauc. *č[ī]ju

Overlooking the Basque reflexes for the moment,⁶¹ the correspondence Bur. /š/ = Cauc. *č/ts’/ is confirmed by other cases:

(‘blind’)	Bur. <i>šon</i>	=	Cauc. *čǎwnV ‘dark’
(‘driftwood’)	Bur. (Y) <i>šulú</i>	=	Cauc. *čwǰǰV ‘stick, branch’
(‘wool’) ⁶²	Bur. <i>še</i> (pl. <i>šemiŋ</i>)	=	Cauc. *čfiwěme ‘eyebrow’

Finally, note these parallels:

(‘horse’)	Bur. <i>čhardá</i>	=	Bsq. <i>asto</i>	=	Cauc. *čada
(‘narrow’)	Bur. <i>čhađ-úm</i>	=	Bsq. <i>estu</i>	=	Cauc. *čHVrdV

Here Burushaski and Caucasian have clear vowels between the first and second consonants, while Basque has clustered the consonants and placed a vowel initially. The same tendency is seen in some other comparisons, all involving sibilant affricates: /č/ as in *church*; /čh/ is the aspirated version, /č’/ the glottalized, /č/ the retroflex. In Basque the original sibilant affricates have uniformly become /s/, written *s*:

⁶¹ The Basque reflexes are “conditioned” (vary depending on phonetic environment), as discussed above in the comparison of Basque and Caucasian.

⁶² The semantic connection of ‘wool’ and ‘eyebrow’ is seen clearly in the Sino-Tibetan language Kanauri: *tsam* ‘wool, fleece’ and *mik-tsam* ‘eyebrow’ (lit. ‘eye-wool’). By the way, I consider Proto-Sino-Tibetan **tshām* ‘wool, head hair’ to be cognate with Caucasian **čfiwěme* ‘eyebrow’, Burushaski **še[m]* ‘wool’, and Basque *zamar* ‘fleece’.

(‘measure’)	Bur. <i>-č<u>uq</u></i> ⁶³	= Bsq. <i>aska</i> ⁶⁴	= Cauc. <i>*čăqwă / čăqwă</i> ⁶⁵
(‘all’)	Bur. <i>ç<u>iq</u></i> ⁶⁶	= Bsq. <i>asko</i> , ⁶⁷	= Cauc. <i>*čH ^ qwV</i> ⁶⁸
		<i>aski</i>	

I believe these examples are sufficient to show that the basic lexicon of Burushaski is intimately connected with the basic lexicons of Basque and the Caucasian languages, and this is further verified by the recurrent sound correspondences between and among the three groups of languages.

Morphology: We shall also explore some grammatical correspondences between Burushaski and its Vasco-Caucasian cousins, Basque and Caucasian. Let us begin with the **pronouns**.

Both Burushaski and the reconstructed Proto-Caucasian have **suppletive** pronoun stems in the second person singular. According to Nikolayev and Starostin the original Proto-Caucasian paradigms were very complicated, and difficult to reconstruct with much certainty. For the present purpose, let us compare Hunza Burushaski (as described by Berger, 1998) with two East Caucasian languages, Khinalug and Tsakhur (and Proto-East Caucasian)⁶⁹:

	direct:	genitive:	dative:
1st person singular (‘I – me’):			
Burushaski	<i>je, já</i>	<i>áa</i>	<i>áar</i>
PEC	<i>*zō(-n)</i>	<i>*ʔiz(V)</i>	<i>*ʔez(V) (erg.)</i>
Khinalug	<i>zĭ (nom.)</i>	<i>i, e</i>	<i>as</i>
	<i>yă (erg.)</i>		
Tsakhur	<i>zu</i>	<i>yiz-in</i>	<i>za-</i>
2nd person singular (‘thou – thee’):			
Burushaski	<i>un</i>	<i>góo</i>	<i>góor</i>
PEC	<i>*uō(-n)</i>	<i>*ʔeuV/*ʔiuV</i>	<i>*ʔōgwV (erg.)</i>
	<i>~ *GwV̄</i>		
Khinalug	<i>wĭ (nom.)</i>	<i>wi</i>	<i>oχ</i>
	<i>wa (erg.)</i>		
Tsakhur	<i>wu ~ Gu</i>	<i>y-iG-</i>	<i>wa-</i>

⁶³ (H) *-čuq*, (N) *-čóq* ‘a measure of grain’; (Y) *çiq* ‘sifting tray’

⁶⁴ ‘crib, manger, trough’

⁶⁵ ‘scoop, spoon, wooden vessel’

⁶⁶ (Y) ‘all, altogether’; H and N have different words.

⁶⁷ *asko* ‘much, many’, *aski* ‘enough’

⁶⁸ ‘big’ (‘many’ in Lak)

⁶⁹ Here the symbol /G/ is used to represent the voiced uvular fricative (otherwise /ɣ/ or /ʁ/).

In spite of some rearrangements of the original paradigms, much similarity can still be seen. Both Burushaski and East Caucasian have two different **allomorphs** of the word for ‘thou’. What is more, the allomorphs are phonetically as well as functionally similar, *i.e.*, *un* = **uō(-n)*, and *góo* = **GwV̄*. The Burushaski word for ‘thou’ (direct), *un*, is identical with those of the East Caucasian languages Archi and Udi (*un* ‘thou’; cf. Lezgi and Agul *wun* ‘thou’). The Khinalug and Tsakhur forms cited above lack the final *-n*.

A note on **suppletion**: Suppletion means the coexistence of lexically unrelated variants (**allomorphs**) of the same morpheme. A common example is the paradigm of English ‘to go’:

present tense: *go* **past tense:** *went* **participle:** *gone*

Though English speakers unconsciously think of *went* as the “same” word (in the past tense) as *go*, it was originally the past tense form of another verb, *to wend*, that replaced the earlier past tense of *go* (Old English *éode*, itself a suppletive form! German [*ging*] and Swedish [*gick*] preserve the original Germanic past tense forms). Suppletion is also found in the paradigm of *good*:

	positive:	comparative:	superlative:
English:	<i>good</i>	<i>better</i>	<i>best</i>
German:	<i>gut</i>	<i>besser</i>	(der/die/das) <i>beste</i>
Swedish:	<i>god</i>	<i>bättre</i>	<i>bäst</i>

All three languages (as well as other Germanic languages) use a different stem for the comparative and superlative forms of ‘good’.⁷⁰ Joseph H. Greenberg cited these forms as examples of superior evidence of genetic relationship.⁷¹ Complete paradigms showing *shared irregularities* of this type are highly unlikely to have come about by chance, and thus far more probative than the comparison of the single words *good/gut/god* by themselves. Likewise, the pattern of *r/n* alternation discussed above was viewed by linguists as a valuable piece of evidence for genetic relationship of Hittite with the Indo-European languages.

I suggest that the suppletive pronouns found in Burushaski and East Caucasian should also be considered strong evidence of genetic relationship:

⁷⁰ Those who have studied the classical languages will recall that Latin and Greek also have irregular comparative and superlative forms of ‘good’ (and some other adjectives), though with totally different stems than are used in Germanic. Sanskrit also has a few such cases.

⁷¹ “Agreement in irregularities and evidence from survivals of grammatical markers that have become petrified are worthy of special consideration [in the genetic classification of languages].” Greenberg, “Principles of Genetic Classification” (p. 30), in *Language in the Americas* (Stanford University Press, 1987).

Burushaski *un* / *góo* ‘thou’ = PEC **uō(-n)* / **GWV̄* ‘thou’

Both languages have two different allomorphs of the word for ‘thou’. What is more, the allomorphs are phonetically as well as functionally similar, *i.e.*, *un* = **uō(-n)*, and *góo* = **GWV̄*. This, coupled with the many lexical parallels and recurrent sound correspondences listed above, can only be the result of genetic relationship.

As we saw above, there is a tendency over time to even out irregularities. Thus children say things like **gooder*, **goodest*, and eventually some of these neologisms become accepted. We saw, for example, that the Germanic languages generalized either the *-r* form or the *-n* form of the word for ‘water’. In the Vasco-Caucasian family two widely separated languages – Basque and Dargwa – have independently leveled out the suppletion of first and second person singular pronouns in the same way:

Basque	<i>ni</i>	‘I’	:	<i>hi</i>	‘thou’
Dargwa (Akushi, Urakhi)	<i>nu</i>	“		<i>ħu</i>	“

The *nV* stem is preserved only in Dargwa and Lak (*na*) among the Caucasian languages (as well as in Basque), while the *ħu/hi* ‘thou’ stem (presumed cognate with Burushaski *góo*, Tsakhur *gu*, etc.) has been generalized at the expense of the **uō(-n)* allomorph.

Another morphological pattern is found when we look into the grammatical gender (or class) system of nouns in Burushaski. Most Indo-European languages have a three-gender or two-gender system, or have merged all genders and thereby lost their grammatical significance:

Three genders: masculine (M) / feminine (F) / neuter (N): Sanskrit, Greek, Latin,

German, Icelandic, Norwegian

Two genders: M / F: French, Spanish, Italian, etc.

or common (M+F) / neuter (N): Swedish, Danish

No genders: English

Since Hittite had a two-gender system (animate / inanimate), many linguists think that the three genders of Indo-European are an innovation (a split of the animate gender), and the older Indo-Hittite system was two-gender. In Burushaski there are *four* genders (or classes):

Class	singular	plural
1.a. human-masculine	<i>i-</i>	<i>u-</i>
1.b. human-feminine	<i>mu-</i>	<i>u-</i>
2. non-human animate	<i>i-</i>	<i>u-</i>
3. non-human inanimate	<i>i-</i>	<i>i-</i>

The phonetic elements in the right column are third-person possessive prefixes corresponding to each class. Most Caucasian languages also have a multiple-gender system, and the following system is reconstructed for Proto-East Caucasian:

Class	singular	plural
1. animate (masculine)	* u	* w
2. animate (feminine)	* y	* w
3. inanimate (non-collective)	* w / * b	* r
4. inanimate (collective)	* r / * d	* r

The phonetic elements in the right column are the **markers** associated with each class. Depending on the language, they may appear attached to adjectives, or verbs, or nouns, for example:

Bagwali *w-ass* ‘brother’, *y-ass* ‘sister’; *w-aša* ‘son’, *y-aš* ‘daughter’

Tindi *w-aha* ‘son’, *y-aha* ‘daughter’, *b-etl’tl’u* ‘stomach’, *b-atl’tl’i* ‘in the middle’

Dargwa *w-ǰh* ‘face (of a man)’, *r-ǰh* ‘face (of a woman)’, *b-ǰh* ‘face (of an animal)’

In some Caucasian languages the prefixes have been retained sporadically in some words even though the original class distinctions are forgotten. Elements such as these are called **petrified** or **fossilized**.

Lezgi *ru-fun* ‘belly’; Tsakhur *wu-xun* ‘belly’; Ubykh *t-χamə* ‘skin, fur’, *t-χwa* ‘ashes’, *b-ħa* ‘eye’, *b-ħa* ‘top’, etc.

The Burushaski possessive prefixes are fully functional, and (roughly) similar to the Dargwa examples above:

Bur. *á-sqa* ‘on my back’, *gó-sqa* ‘on your back’, *í-sqa* ‘on his back’, *mú-sqa* ‘on her back’

And what of the “third leg” of our trifecta, Basque? In contrast to Burushaski and Caucasian, modern Basque has no grammatical gender or class distinctions whatsoever, but there are some interesting prefixed elements that have attracted linguists’ attention. The original prefixal nature becomes clear when we compare Basque words with their proposed cognates in Caucasian and Burushaski:

1. o- / u-

Basque *olho, olho* ‘oats’ (< **o-ʃo*): cf. Cauc: PNC **λwʃwV* ‘millet’

Basque *oihal*, ‘cloth, fabric’ (< **o-xal*): cf. Cauc: PEC **χwʃlV* ‘clothing’

Basque *ohe (o-he)* ‘bed’: cf. Cauc: Tabasaran *aχin* ‘bed’, etc.

Basque *oihan (oi-han)* ‘forest’: cf. Cauc: Chechen *ħun* ‘forest’, etc.

Basque *oski (o-ski)* ‘shoe’: cf. Cauc: Tabasaran *šaq*^w ‘heel’; Bur. *šoq* ‘sole of shoe’

Basque (Bizkaian) *uzen (u-zen)* ‘name’: cf. Bur. *sén-* ‘to say, name’; *sénas* ‘named’

Basque *urdail (u-rdail)* ‘stomach’: cf. Cauc: Tindi *b-etl’tl’u*, Karata *m-etl’u* ‘stomach’; Bur. *-úl* ‘belly’

Basque *ukondo (u-kondo)* ‘elbow’: cf. Cauc: Hinukh *q’ontu* ‘knee’, etc.

2. e- / i-

Basque *elhur, elur* (< **e-ʃu-r*) ‘snow’: cf. Cauc: PEC **λʃwV* ‘snow’ > Chechen *lō*, etc.

Basque *egur (e-gur)* ‘firewood’: cf. Cauc: Udi *gor, gorgor* ‘pole’, etc.

Basque (Z) *ekhei (e-khei)* ‘material’: cf. Cauc: Lak *q’aj* ‘thing(s), ware(s)’, etc.

Basque *esne (e-sne)* ‘milk’: cf. Cauc: Chechen *šin* ‘udder’; Bur. *šij* ‘milk’

Basque *ele, elhe (e-lhe)* ‘word’: cf. Cauc: Inkhokwari *loje* ‘word, sound, voice’, etc.

Basque *izar (i-zar)* ‘star’, etc.: cf. Cauc: Tindi *tstsaru* ‘star’, etc.

Basque *ihintz (i-hintz < *i-xinc)* ‘dew’: cf. Cauc: Lak *xunts’a* ‘bog’, etc.

Basque *ilindi, ilhinti (i-lhinti)* ‘firebrand, ember’: cf. Cauc: Andi *ʃudi* ‘firewood’, etc.

Basque *idulki (i-dulki)* ‘block of wood’: cf. Cauc: Archi *dali* ‘long stick, pole’, etc.

Basque *itain, ithain (i-thain)* ‘tick’: cf. Cauc: Akhwakh *t’ani* ‘nit’, etc.

Basque (common) *izen (i-zen)* ‘name’: cf. Bur. *sén-* ‘to say, name’; *sénas* ‘named’

Basque *izerdi (i-zerdi)* ‘sweat, sap’: cf. Cauc: PEC **čāLwV* ‘blood, life’

3. be- / bi-

Basque *behatz (be-hatz)* ‘thumb, toe’ (vs. *hatz* ‘finger, paw’): cf. Cauc: Avar *kʷač* ‘paw’, etc.

Basque *belarri (be-larri)* ‘ear’: cf. Cauc: Batsbi *lark* < **lari-k* ‘ear’, etc.

Basque *belaun, belhaun, belhain (be-lhaun)* ‘knee’: cf. Cauc: Akhwakh *etlelo* ‘elbow’, etc.

Basque *behazun (be-hazun)* ‘bile’: cf. Cauc: Archi *ssam* ‘gall’, etc.; Bur. *-sán* ‘spleen’

Basque *bizkar (bi-zka-r)* ‘back’: cf. Cauc: Abkhaz *azkʷa* ‘back’; Bur. *-sqa* ‘on one’s back’

Basque *bizi (bi-zi)* ‘life, alive’: cf. Cauc: Lak *ssiħ* ‘breath, vapor’, Chechen *sa* ‘soul’, etc.

Basque (G) *bilder (bi-lder)* ‘drivel, drool’ (vs. *helder, herde* in other dialects): cf. Cauc:

PNC **fām* √*ǎ* ‘sweat’

4. a-

Basque *ahizpa* (**a-hiz-pa**) ‘sister (of a woman)’: cf. Cauc: Bezhta *is* ‘brother’, *isi* ‘sister’, etc.;

Bur. -*ço* ‘brother (of a man), sister (of a woman)’

Basque *atso* (**a-tso**) ‘old woman’: cf. Cauc: Batsbi *pst’u* ‘wife’, Lak *ttsu-* ‘female’, etc.

Basque *abere* (**a-bere**) ‘domestic animal, cattle’: cf. Cauc: Udi *bele* ‘cattle’, etc.

Basque *ahuña* (**a-huña**) ‘kid’: cf. Cauc: Andi *kun* ‘ram’, etc.

Basque *akain* (**a-kain**) ‘tick’: cf. Cauc: Chechen *gēnig* ‘louse’, etc.; Bur. *khen*, *khin* ‘flea’

Basque *ametz* (**a-metz**) ‘gall oak’: cf. Cauc: Chechen *naž* ‘oak tree’ etc.; Bur. (H) *meş*, (Y) *noş*
‘bush, shrub, sapling’

Basque *aho* (**a-ho**) ‘mouth’: cf. Cauc: PNC **χwí-* in **χwím(V)ǰV* ‘mouthful’

Basque *ahur* (**a-hur**) ‘hollow of the hand’: cf. Cauc: Dargwa *kur* ‘pit’, etc.

Basque *adar* (< ***a-rdar**) ‘horn’: cf. Cauc: Avar *tltlar* ‘horn’, etc.; Bur. -*ltúr*, *tur* ‘horn’

The original separability of these apparent prefixes is shown also by internal evidence. Some Basque words have different prefixes in different dialects, or a prefix in some dialects but not in others, or prefixed and unprefixes words can co-exist (see *hatz* vs. *be-hatz*, below):

- Basque (c) *izen* (**i-zen**) ‘name’ vs. Basque (B) *uzen* (**u-zen**) ‘name’
- Basque (c) *k(h)e* vs. Basque (AN, BN, R) *e-ke* ‘smoke’
- Basque (AN, BN, L) (*h*)*erde*, (B, BN, L, R, Z) (*h*)*elder* ‘drivel’ vs. (G) *bilder* (< ***bi-** + *helder*)
- Basque (c) *gai*, *gei* ‘material, subject, topic’ vs. (BN, Z) *e-khei* id.
- Basque *hatz* vs. *be-hatz*: (Meanings vary widely depending on dialect, e.g. in Bizkaian, *atz* ‘finger’ vs. *beatz* ‘toe’; in Zuberoan, *hatz* ‘finger’ or ‘paw’ vs. *behatz* ‘thumb’)

Even though these prefixed elements no longer carry any grammatical function in modern Basque, some linguists have speculated that they formerly did, and represent what Joseph Greenberg called “stage III articles.” Clear examples of stage III articles are seen in Haitian Creole, for example /*latab/* ‘table’, /*deze/* ‘egg’ (< French *la table*, *des oeufs*), where the former articles are no longer separable but have fused with the noun roots. The same thing has happened with the Caucasian fossilized prefixes mentioned above. For example, in Lezgi *rufun* ‘belly’ (< *ru-fun*) the old class prefix (*ru-*) has fused with the noun stem (*fun*), just as in the Haitian Creole examples. (However, the class system is still alive in Tsakhur and Rutul, which have *wu-xun* and *u-xun*, respectively, belonging to the third class.)

In sum, I think the most likely explanation of the Basque prefixes is that they are fossilized remnants of what formerly were class markers. In this light the Basque fossilized prefix *be-* / *bi-* resembles the Caucasian 3-class marker **w* / **b*:

Basque *be-larri* ‘ear’, *bi-zkar* ‘back’: cf. Cauc: Tindi *b-etl’ tl’u* ‘stomach’; Dargwa *b-ǰh̄* ‘face (of an animal)’; Tsakhur *wu-xun* ‘belly’; Ubykh *b-ǰa* ‘eye’, etc.

If so, Basque *o-* / *u-* could correspond to the Caucasian 1-class marker **u*, and Basque *e-* / *i-* could be related to the Caucasian 2-class marker **y*:

Basque *o-saba* ‘uncle’, *o-he* ‘bed’, *u-rdail* ‘stomach’, etc.: cf. Cauc: Bagwali *w-ass* ‘brother’, *w-aša* ‘son’; Tindi *w-aha* ‘son’; Dargwa *w-ǰh̄* ‘face (of a man)’

Basque *i-zeba* ‘aunt’, *i-hintz* ‘dew’, *e-l(h)ur* ‘snow’, etc.: cf. Cauc: Bagwali *y-ass* ‘sister’, *y-aš* ‘daughter’, Tindi *y-aha* ‘daughter’; Chechen *y-üñ* ‘face, end’, etc.

This hypothesis is supported by the opposition of Basque *o-saba* ‘uncle’ vs. *i-zeba* ‘aunt’, possibly an old but rare parallel to the Caucasian opposition of **u* (masculine) vs. **y* (feminine).

To summarize the grammatical evidence for classes (genders) in the Vasco-Caucasian family: Burushaski and Proto-East Caucasian agree in having a 4-class system; modern Basque has no grammatical genders or classes, but the lexical evidence of fossilized prefixes points to the earlier existence of a multiple-class system.

