

Some Euskaro-Caucasian Homologies

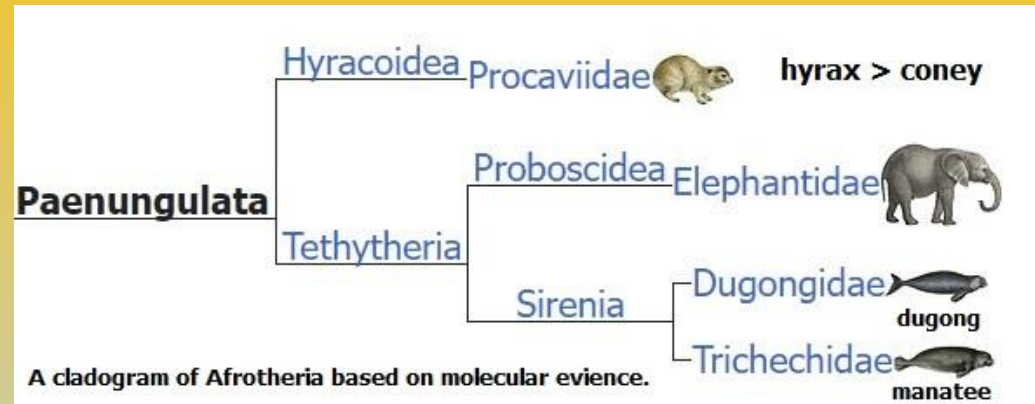
Lexical forms and patterns
common to Basque and North Caucasian

John D. Bengtson, Association for the Study of Language in Prehistory
(ASLIP) <https://www.mother-tongue-journal.org/>

For the 18th Sergei Starostin Memorial Conference: March 24, 2023

Introduction: Homology

In biology a homology is any characteristic of biological organisms that is derived from a common ancestor.



In anthropology, of which historical linguistics is a sub-field, a homology is an analogy between human beliefs, practices or artifacts owing to genetic or historical connections. "Human practices" include language, so a homology (or homologue) in compared languages consists of a phoneme, morpheme, or word shared by two languages or dialects, via cognation (common genetic origin) or naturalization (borrowing).

Since there is no evidence of direct borrowing between modern Basque and modern North Caucasian languages, the homologies discussed here are all of the genetic type, and are therefore some of the evidence for a common genetic origin of Basque and the North Caucasian languages.

Introduction:

So, do the homologies discussed here serve to “prove” the Euskaro-Caucasian hypothesis? As I have written elsewhere, I do not think the concept of “proof” of a linguistic hypothesis is the best approach. I prefer to think in terms of “best explanation (of the facts),” as is common in “hard” sciences like physics.

For example, when was the Indo-European (IE) hypothesis “proven”? Parts of the IE family had been recognized already for centuries when terms like “Indo-Germanic” and “Indo-European” began to be used by 1820. “Sound laws” (correspondences) had been proposed by Rask and Grimm, but nobody at that time thought of them as “proving” IE. When Brugmann and Osthoff asserted that “sound laws have no exceptions” (1878), nobody thought that the sound laws “proved” IE, which had been widely accepted by “philologists” for at least six decades.

Therefore, I prefer to think about the homologies discussed below as *properties* expected of a linguistic family of a common genetic origin, rather than as evidence leading to a “threshold of proof” of a language family. “Threshold of proof” is a subjective concept. Anyone who would like to deny the validity of a language family can always “move the goalposts” higher or farther. The “best explanation” approach, on the other hand, simply challenges the critic to propose a better explanation of the linguistic facts.

Introduction: The Euskaro-Caucasian hypothesis

The Euskaro-Caucasian (EuC) hypothesis – postulating that the Basque (Euskara) language is genetically related to some or all of the indigenous languages of the Caucasus region – has developed gradually over the past two centuries. It was supported by some eminent linguists, like Uhlenbeck, Dumézil and Lafon.

In its present form the EuC hypothesis has been adopted – and adapted – as a part of the Sino-Caucasian hypothesis posited by Sergei A. Starostin in the 1990s.



Verbal forms: The Archi verbal suffix **-t:u** was defined by S.A. Starostin as “adjective substantivator / participial suffix.” This and other NC cognates (like Ubykh **-də** [participle]), strengthen the case that the Basque participial suffix **-tu** is of ancient (Euskaro-Caucasian) origin, secondarily reinforced by the Latin participial suffix **-tum**, “the idea being that so many Latin participles were borrowed with this suffix that it came to be regarded as the normal marker of the perfective participle in Basque” (Trask).

Archi	Basque	PNC / PEC
guli-t:u- `hidden, secret	*gal-tu `lost, misplaced, disappeared	Tsakhur a=gwal- `to get lost', etc. < PEC *=igwV† `to lose, get lost; steal'
xala-t:u- `old (person)'	*a=gure-tu `aged, having become an old man'	Andi =oxor `old (of a person)', etc. < PEC *=öñqVIV `old (person)'
hōlo-t:u- `liquid'	*hur-tu `melted, watered, poured	Lezgi hül `sea', etc. < PEC *hwiV ~ *lihwV `river, reservoir'
hiba-t:u- `good'	*hobe-tu `made better' (*hobe `better', *habo-ro `more')	PNC *HǎbV ~ *bǎHV `big, many'
q'ᶜwaq'ᶜwar-t:u- `narrow'	*garhi-tu `made (oneself) thin, slimmed down'	Avar q':warí-da- `narrow, cramped', etc. < PEC *qwarHV `narrow, thin'

Phonology: The non-trivial correspondences of Basque laterals /l/ and clusters *-rd- [r̥ð] ~ *-rt- [r̥t] with North Caucasian lateral affricates *λ, *λ̄, *ł = [tʃ], [tʃʰ], [dʒ].

In initial and final positions Basque /l/ corresponds to all PNC / PEC lateral affricates:

- Basque *laru 'skin, hide, leather' = PNC *łöli 'skin, color' (Avar *tʃ:er* 'color' < *'skin')
- Basque *limu-ri 'slippery, moist, smooth', etc. = PEC *λHwemV 'liquid' (Avar *tʃ:amí-ya-* 'liquid')
- Basque *lincu-ra 'marsh, bog', PEC *λwilçwV 'bog, marsh' (Karata *ɬic':u* 'bog, marsh')
- Basque *lo 'sleep' = PNC *=HVwλĀn 'to sleep' (Akhwakh *tʃ:u-nu-* 'to sleep')
- Basque *hil 'to die; dead; death' = PNC *=iwλĒ 'to die, kill' (Avar *=a/ɣ-* 'to die')
- Basque *oi=hal 'cloth' = PEC *χw[a]łłV 'a kind of clothing' (Bezhta *χitʰo* 'trousers, breeches')

In medial positions Basque *-rd- (pronounced [r̥ð], with trilled [r̥] + voiced fricative [ð] in most dialects) corresponds to all PNC / PEC lateral affricates:

- Basque *erdi 'middle, half' = PNC *=ěλĒ 'middle, half' (Karata *b=attʃ:i* 'in the middle, between')
- Basque *i=serdi 'sweat; sap' = PEC *čāłwV 'blood; life' (Batsbi *c'eygi* 'blood')
- Basque *barda 'belly, bowels' = PEC *bħěřλV '(large) intestine' (Bezhta *batʰa* 'large intestine')
- Basque *mardo 'bunch, group' = PEC *măr[λ]o 'handful, armful' (Hinukh *metʰu* 'handful')
- Basque *urde 'swine, pig', *ord-oć 'boar' = PNC *wHārλwə 'pig, boar' (Hunzib *butʰu* 'boar, pig')

Less frequently the Basque intervocalic reflex is the voiceless cluster *-rt-:

- Basque *urte 'year' = Avar *utʃ'i-sa* 'last year' < PEC *ʔVλwV 'last year'.

This alternation seems to have been determined by the ancient position of accent.

Phonology: The Basque /rc/ ~ /št/ stem alternation. As noted by Luis Michelena, in a very few cases, Basque words exhibit a stem alternation between medial/final /rc/ and /št/. It so happens that PNC provides a key, since all of the words with the /rc/ ~ /št/ stem alternation on this short list have NC cognates with clusters of *r + sibilant affricates (*-rc-, *-rč-, *-rč-):

(Proto-) Basque	Basque (attested)	North Caucasian	PNC / PEC
*perc ~ *berc ~ bešt 'cauldron, pail'	(G, AN) <i>pertz</i> 'bucket, pail'; (BN-ms) <i>best</i> id.	Tsakhur <i>parč</i> 'water jug'; Lak <i>parš</i> 'brass water jug'; Avar <i>parč:í</i> 'earthenware water jug'	PEC *parčē 'water jug'
*borc ~ bošt 'five', *bošte-ko 'hand'	<i>bost</i> ~ <i>bortz</i> 'five'; (B, G) <i>bosteko</i> 'hand', (L-Sara) <i>bortzeko</i> id.	Inkhokwari <i>buso</i> 'fist'; Avar <i>púrc:i</i> 'ham'	PEC *b[ə][r]cV 'paw; part of arm or leg'
*herce ~ *hešte 'intestine'	(BN, L) <i>hertze</i> 'intestines, guts'; (B, G, AN) <i>este</i> id.	Avar <i>Ɔorčó</i> 'intestine', Andi <i>ručo-l</i> id; Abkhaz <i>á-ca</i> 'abomasum, rennet'	PNC *HǎrčV (~ *rHǎčV ~ *čǎrHV) 'stomach, abomasum'
*horci ~ *hošti 'sky, storm'	(AN, BN) <i>orz</i> 'sky', (Z) <i>ortze</i> 'storm'; (B) <i>osti</i> 'storm, thunder'; (c) <i>ost-egun</i> 'Thursday'	Karata (Anchikh, Archo dialects) <i>orč-</i> 'to become cloudy', Karata tʃ'- <i>orč-</i> , Avar <i>róč:n-</i> id.	PEC *HǎrčwVn 'to become cloudy, gloomy'
*mucu-ř (< *murcu-ř) ~ *muštu-ř 'snout; edge, extremity'	(B) <i>mustur</i> 'snout, muzzle; corner, edge, extremity' (pl.) <i>musturrak</i> 'lips'; (BN, R) <i>mutzur-keta</i> 'rooting made by snout of a pig'	Andi <i>maču</i> 'snot'; Chechen <i>marš</i> 'snot'	PEC *mHǎrčwV 'pus; mucus, snot'

Phonology: Homologies in Basque and PNC / PEC final vowels. In a significant number of cases, present-day Basque final vowels precisely match final vowels in reconstructed Proto-North Caucasian (or East Caucasian).

There is no possibility that Nikolayev & Starostin were influenced by the Basque forms, since PNC and PEC were developed strictly from the comparison of NC languages only.

Basque	PNC / PEC	North Caucasian
*a=śka `trough, manger, crib'	PNC *ćăqwaǎ ~ ćăqwaǎ 'wooden vessel, scoop, spoon'	Abkhaz a-ćaq'wá 'wooden scoop with a handle'
*pista `rheum, gleet'	PNC *pĩncwAǎ 'resin, juice'	
*e=rece `corner, edge, side', etc.	PEC *ʔrəyçweǎ 'hind, bottom'	
*gośe `hunger, hungry'	PNC *gašē 'hunger'	
*śare ~ *śal- `net, grill'	PEC *čhaťē 'enclosure, fence'	
*heři `people, nation', etc.	PNC *ʔwǎhri 'troops, army'	[Archi ʔari 'army, troops', an old loanword from Lak]
*minhi `tongue' (< *minši)	PNC *mǎlcǐ 'tongue'	Andi mic':i tongue'
*mardo `bunch, group, cluster'	PEC *mǎr[λ]o 'handful, armful'	
*tanko (~ *tanka) `drop (of liquid)'	PNC *ʔHänkǒ 'drop, spray'	
*[h]uśu `pus, serum'	PNC *h[a]mʒū 'saliva, mucus; herpes'	Lak unction 'snot, nasal mucus'; Avar ħác':u 'saliva'
*mośu `nose, face'	PEC *mħərçū 'protruding part, point, edge'	Avar móc':u 'teat, nipple; tip'

Phonology: Recurrent transpositions can be seen in a series of Basque and NC verbs, in which Basque has the structure **-UC-** and PNC/PEC has the structure **-CwV-** or **-CVw-**.

Basque	Basque (attested)	PNC / PEC	NC (attested)
*e=aući 'to go down, descend'	(AN) <i>jautsi, jatsi, jatxi, jetxi</i> 'go down'	*=ičwĒ 'to come; return'	Lak =u=č'a-n- 'to go down'
*e=uci 'to leave, let'	(c) <i>utzi</i> 'to leave, abandon; let, permit'	*=ičwV 'to run'	Andi =uč':in- 'to run (away)'
*e=ući 'to consider, think'	(c) <i>etsi</i> 'to despair, be resigned, lose hope', <i>uste</i> 'opinion, belief, idea, plan; trust, confidence, faith, hope'	*=ičwĒ 'to count; understand'	Avar =ič':- 'to understand'
*e=ući 'to take, hold, keep'	(c) <i>eutsi</i> 'to take hold, seize, grasp'	*=áčwV 'to take, carry'	Akusha =uč'- 'to gather, collect'
*e=uka-n ~ *e=uki 'to have'	(BN) <i>ukhan, ukan</i> 'to have', (B) <i>euki</i> 'to have, possess, contain; keep (secret); power, property'	*=iqwV 'to hold, catch'	Hinukh =uχ- 'to take, hold, catch'
*hausi 'to break, shatter'	(BN, L) <i>hautsi</i> 'to break', shatter, tear, rip', (B, G) <i>ausi</i> id.	*=ičV(w) 'to break, hit, wound'	Tsezi =ešu+- 'to break, be broken'
*Husi 'to divide, split'	(L) <i>huzi, huzi-tu</i> 'to distribute, divide out'	*=VrswĒ 'to cut (in parts)'	Tindi =us:i'- 'to cut (lengthwise)'
*e=aurti 'to throw, hurl, launch'	(B) <i>jaurti</i> 'to throw, hurl, launch'	*=äłĒw 'to lie; to put; lead'	Chechen =ill' 'to put'

Phonology: There are some precise matches of Basque laryngeal clusters like **-rh-*, **-lh-* with similar PNC/PEC reconstructed forms, with the Northern Basque laryngeals in the same position as in PNC / PEC.

(Proto-) Basque	Northern Basque (attested)	PNC / PEC	NC (attested)
*arhe ~ *arhe 'harrow, rake'	(BN, L) arhe, ahe '5-spiked harrow', (Z) árrhe, áhe id.	*ɤarhV 'harrow'	Andi ɤari 'harrow'
*beŋa-ř 'forehead; side'	(BN, L) belhar, belar 'forehead'	*bɤāŋhō 'edge, end; forehead'	Kryz bel 'forehead', Tindi bala 'edge, end, corner'
*buŋa-ř 'chest; female breast'	(BN, L) bulhar, (Z) búlhar 'chest; female breast'	*gɤǎŋhē 'udder, breast'	Avar ɤwári 'udder'; Hunzib ɤeru 'breast, udder'
*eŋhi 'cattle, herd; throng'	(L) elhi 'cattle, livestock, herd, flock', (Z) elhi 'throng, gang, group, swarm'	*χwōlhV ~ *χHwōlV 'cattle, cows'	Chechen hēlī 'cows', Ingush hāliy id.
*erhi ~ *el- 'finger'	(BN, L) erhi, (Z) é(r)hi, (AN, R, Sal) eri 'finger' [(B) el-aztun '(finger-)ring']	*χěŋHe 'sleeve'	Rutul χili-dī 'mitten', Tabasaran χil 'sleeve'
*garhi 'thin, slim'	(BN, L) garhi 'thin, slender, slim, slim-waisted', (Z) gárhi, gáhi id.	*ǵwarHV 'narrow, thin'	Avar q':warí-da- 'narrow, cramped'; Agul q':ure- 'thin, emaciated'
*gurhi 'fat, butter'	(Z) gurhi, guhi 'fat, lard, butter', (Bzt) guri id.	*χərHV ~ *χHərV 'butter, cheese'	Lezgi ɤeri 'butter', Rutul χʕara-d id.

Submerged grammar: Basque (fossilized) prefixes = NC gender markers:

I. *be= ~ *bi= ~ *b=

Uhlenbeck was aware of these Basque prefixes in the 1920s, and suggested they could be relics of an older, fossilized gender / noun class system, as in the Bantu languages. This is the interpretation I have followed, for several reasons:
(a) the nouns with these prefixes denote parts of the body, or its attributes (*bi=si 'life') or excretions (*bi=helde-ř 'slobber').
(b) the nouns with *b= prefixes are never associated with whole humans or animals: see especially *a=, below.

Basque *b= most likely corresponds historically with East Caucasian *w=/b=, marker of III-class singular, which left traces also in Abkhazian languages.

Basque form	Basque gloss	North Caucasian (modern)	North Caucasian gloss	Proto-North Caucasian / East Caucasian
*be=hac	finger ~ thumb ~ toe	Avar kwač'	paw	*kwăčĕ
*hac	paw ~ finger ~ inch			
*bi=helde-ř				
*herde ~ *helde-ř	drool, slobber	Karata hant'á	sweat	*hămlă
*be=řari	ear	Batsbi lark'	ear	*lari-k < PNC *řĕHi(-rV-kV)
*bi=hoc	heart	Dargwa k'ac' Circassian k'wac'	spleen entrails	*kwıçĕ ~ *çıkwĕ
*bi=si	life, alive	Lak s:iň Chechen sa	breath, vapor soul	*řĭHwV
*bi=ska-ř	back	Abkhaz zək ^w a	back	

Submerged grammar: Due to the symmetry of the Basque fossilized prefixes, with a mid-vowel allomorph and a high-vowel allomorph, we have postulated that the prefixes ***be=** and ***bi=** are allomorphs of the same historical morpheme, and likewise with ***e=** and ***i=**, and ***o=** and ***u=**.

II. The ***e= ~ *i=** fossilized prefix is frequently associated with words for liquids and other mass or uncountable nouns:

Basque	North Caucasian	PNC / PEC
*e=thu-ř `snow'	Chechen lō ~ lūo `snow'	PEC *łiwV ~ *yiwλV
*e=kē ~ *kē `smoke'	Avar k':uy `smoke'	PNC *kwĩñhV
*e=śene `milk'	Akhwakh (Tsegob) s:in `milk'	PNC *šǎmʔV `milk, udder'
*i=cal `shadow'	Botlikh =eč'er `black'	PEC *HčōlV̄ `black'
*i=sař `star'	Akhwakh c':wari id.	PNC *zwhǎrī ~ *zwǎhrī `star'
*i=ću `fire' ~ *śu id.	Lak c'u `fire'	PNC *čǎyř
*i=ća(-)śo `sea'	PWC *žə	*čə `salt'
*i=toi `drop'	Avar tuy `spittle', tuy- `to spit'	PEC *twəyV `spit, spittle'
*i=tuři `fountain, spring'	Karata t'or- `to drop, drip, flow'	PEC *=V̄tʷV-r- `to pour, soak'

Submerged grammar: II. The ***e= ~ *i=** fossilized prefix is also frequently associated with anatomical terms (like with ***be= ~ *bi=**, above).

Basque	North Caucasian	PNC / PEC
*e=gal 'wing'	Lak qa 'wing', Lezgi, Agul ɤil 'hand'	PEC *qǎʔi
*e=purdi ~ *purdi 'rump, buttocks'	Archi pǎrt'i 'one of the large intestines'	PEC *pHVrɥwV
*e=ra(h)i 'entrails, pluck'	Tsezi aɣ ^w a-reku 'entrails',	PEC *rēkV ~ *kērV 'chest, belly, entrails'
*e=stari 'throat'	Agul s:ür 'gullet'	PNC * $\text{s}_{w}^{\check{e}}\text{ri}$ ~ * $\text{ri}_{sw}^{\check{e}}$ 'neck'
*i=serdi 'sweat; sap'	Chechen ts'iy 'blood'	PEC *čāłwV 'blood; life'
*i=sta- 'groin, thigh'	Proto-Circassian * $\text{swt}:\acute{\text{e}}$ 'genitals'	cf. Burushic * $\text{=}\acute{\text{s}}\text{t}\text{-i}_{\text{ŋ}}$ 'loins, waist, kidney area'
*i=štu 'spittle, saliva'	Batsbi šatō 'spittle, saliva'	PEC *šHätV id.

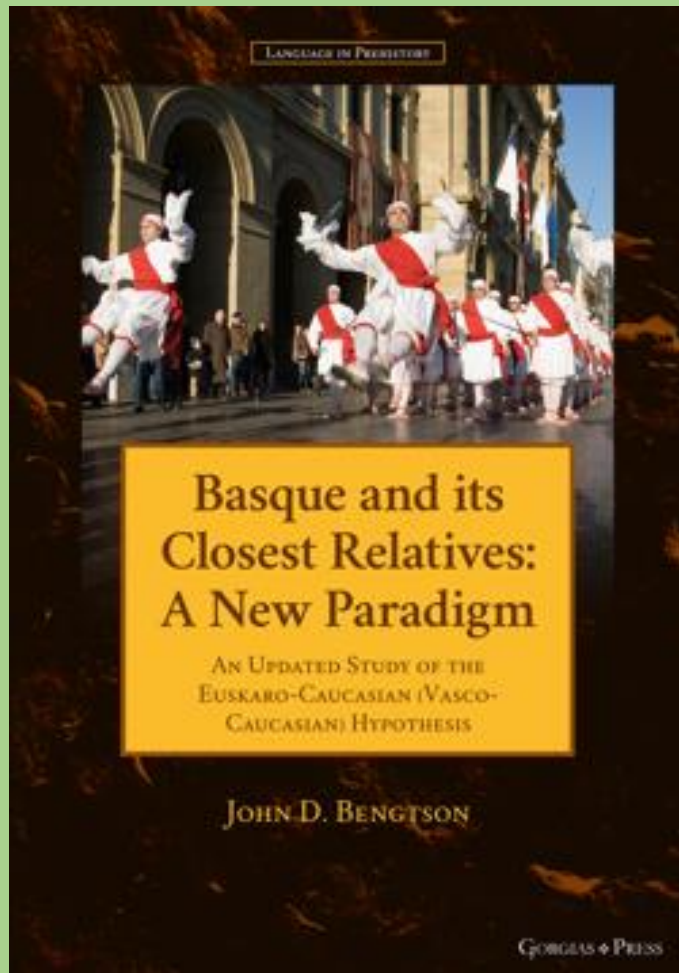
Submerged grammar: Basque (fossilized) prefixes = NC gender markers:

III. The *o= ~ *u= fossilized prefix is frequently associated with terms for artifacts and conventions (intangibles). It is likely related historically to the EC gender marker *u= [I-class singular].

Basque	North Caucasian	PNC / PEC
*o=he 'bed'	Tabasaran axi-n 'bed'	PEC *=axV-r 'to lie, fall'
*o=hol 'board, plank'	Rutul xil 'wooden trough'	PEC *xuIV ~ *xuIV 'wooden vessel'
*oi=hal 'cloth'	Bezhta xitlo 'trousers'	PEC *xw[a]lV 'a kind of clothing'
*o=rac 'needle, pin' *o=rasc 'comb'	Tsezi ric'u 'wedge'	PNC *rəwçē ~ *çəwrē 'pointed stick'
*o=ški 'shoe' < *o=soki	Tabasaran šaq' 'heel'	PEC *šwöçHwV ~ *çHwöšwV 'heel, ankle'
*u=Ice '(metal) nail'	Bezhta hūc'u 'wedge, stake'	PEC *yälçwV 'stick, branch; wedge, stake'
*u=pel 'barrel, cask'	Tsezi pelu 'pipe, reed pipe'	PEC *HpēIV 'pipe; vein'

Thank you for your kind attention!

Basque and Its Closest Relatives: A New Paradigm



Hardbound Book | available from Gorgias Press |
\$124.80 (ask your favorite library to buy it)

Paperback Book | available on Amazon.com | **\$18.50**

Free PDF | available on Academia.edu