Correspondences of Basque and Caucasian Final Stem Vowels: -i/-e, -u/-o

JOHN D. BENGTSON *

This essay is based on the hypothesis of a genetic relationship of the Basque language with the Caucasian language family, a hypothesis that has been current, in various forms, throughout this century1. The present writer is of the opinion that (apart from certain extinct languages, notably Aquitanian) Basque is most closely related to the (North) Caucasian family 2 (Abkhazo-Adygan and Nakho-Dagestanian), and to the Burushic family 3, and only more remotely with the other Dene-Caucasic languages 4 (Yeni-seian, Sino-Tibetan, Na-Dene).

In my pursuit of lexical studies of the Dene-Caucasic macrophylum, I have been struck by the frequent similarity of Basque final (stem) vowels with those found in presumed cognate words in the Caucasian languages.

---


More specifically, the most striking resemblances are found in the Avar-Andian group, and to a lesser extent also in the Tsezian (Dido), group, and the Dargi language. In other Caucasian languages we see varying degrees of apocope of the old final vowels, a reduction which reaches its extreme form in the West Caucasian (Abkhazo-Adygan) tongues.

In the following comparisons, the Basque form is presented first, on the left side of the page, and on the right Caucasian forms that are presumed cognate with the Basque word, and finally the proto-Caucasic form postulated by Nikolaev & Starostin 5.

Note that not all of the Caucasian cognates are cited, just those that share a similar final vowel with Basque.

[Basque -i/-e = Caucasian -i/-e]

1. Basque mih'i ‘tongue’ 6: Andi, Akhvarkh, Botlikh mi'e’ ‘tongue’;
   Tindi, Godoberi mići; Dargi (dialect.) lezmi –
   limzi, etc. (*mêlc'i’). 7.

2. Basque begi [beγi] 8:
   <*berγi “eye”
   Dargi huli–çule ‘eye’; Tsezian *hoore:
   Tsez ozuri, Bezhta hâre, Hunzib hare, etc.

3. Basque *be-lař’i ‘ear’ 10:
   Dargi lihi–lahi ‘ear’; Proto-Nakh *lari-k’
   ‘ear: Chechen ler–(dialect.) lerig–lerig
   (*lêHle ~ lêHji).

4. Basque aragi [araŋi] :
   “meat, flesh” 11
   Andi, Akhvaekh, Tindi, Karata, Botlikh
   rit ’i “meat”, Chama1a1 (dialect.) ref ’i
   (“rät’i”.


6. I exclude the proposed derivation of mih'i “tongue” from *bini (L. MICHELENA, Fonética Historica Vaca, San Sebastián, 1961, p. 412). mih'i is parallel with behi “cow”, just as the same words are parallel in Caucasian (e.g., Botlikh mić'i “tongue”: bući ‘cattle’). In Proto-Caucasic both words had an internal resonant (*mêlc'i: “blârc’w”) thus I postulate the Basque developments as follows: *mići: *berγi (with retroflex x, cf. Burushaski melc ‘jaw’) > *miși: *beγi > *mïxi: *behi (where x was a velar fricative < German ich-laut) > mithi: behi. The Zuberoan nasal in mithi is caused by the initial m; behi is not nasalized (J. LARRASQUET, Le Basque de la Basse-Soule oriental, Paris, 1939). The postulated stage *mïxi: *behi is verified by forms preserved in the western Basle Navarre dialect: [mïhçja] “the tongue” and [beççja] “the cow”; the same dialect also preserves [x] in [axhwa] = aho “the mouth” (N. MOUTARD, “Etude phonologique sur les dialectes Basques”, I, FLV 19 (1975), 5-42). The development s (s) > x is commonplace in historical phonology (e.g., Spanish, Slavic).

7. The dialectal Dargi forms are by metathesis, the Akusha dialect having also mez “tongue”.

8. The development was apparently from *berγi, where γ is the voiced velar fricative still heard in Basque (cf. the corresponding sound in Spanish vega, etc.). The correspondence of Basque b-’ Caucasian w- is regular (see below). *ber- (cf. Caucasian: Avar ber ‘eye’) is supported by the Zuberoan variant ber-phuru “eyebrow”: cf. Burushaski -1-pur “eyelash”. (For *berγ>beγi, cf. note 16).

9. γ is a reconstructed pharyngal stop, which develops to h (voiceless pharyngal fricative) or t (voiced pharyngal fricative) in the Dargi dialects, and ordinary h in Tsezian. In other Caucasian languages the labial element predominates: Avar ber ‘eye’, Lezgi vil, Udi pul id.

10. The reconstruction *be-lař’i is based on the postulated fossilized class/gender prefix be-/bi, cf. J.D. BENGTSON, “Postscript I” (in SHEVOROSHKIN, Op. cit., pps. 150-156). Note that Basque *lari- ‘ear’ is virtually identical to Proto-Nakh *lari- ‘ear’ > Chechen larig–lerig–lerig. The north eastern Basque form behari is apparently influenced by beha “to listen”.

11. Here the Basque velar fricative [r’] appears to correspond to a Caucasian lateral affricate [t’], as also with sagu “mouse” (no. 17), cf. note 12.

12. The correspondences of lateral affricates [t’, r’, l’] are very complex in Caucasian languages: in some languages they develop to velars (e.g. Lak dik “meat”, Tabasaran jik, khinalug lika id.) or even uvulars (Iwugšu duX “meat”); and conversely, original velars become laterals, cf. no. 17.

[8]

14. A Bizkaian form, aegari, seems to be very archaic, reflecting earlier *ašeγali, which is very close to the Caucasian *ceHwo¯le˘ (where E represents an unspecified front vowel, and H a laryngeal of uncertain quality). The derivation from the name Acenari (MICHELENA, Op. cit., ps. 119) is declined as both semantically and phonetically implausible. There is no evidence for -n-, since Zuberoan axeri [æɣε] “fox” has no nasal vowel (LARRASQUET, Op. cit.). A stem variant, azel-, is evidence that -z- is original. On the other hand, the Caucasian parallel is semantically exact, and phonetically plausible.

15. Other Caucasian forms retain 1 (e.g., Ingush cogal, Tabasaran sul, Khinalug pllä “fox”).

16. The development was probably hegi<*herj, parallel with no. 2 (begj<*berj) and no. 17 (sagu<*cerrj). The same r-elision takes place in some Caucasian languages (e.g., Tsakhur sok “weasel” <*caçwt: cf. Basque suga).

17. The correspondence of Basque -rd- to Caucasian lateral affricates (cf. note 11) is well documented: e.g., Basque adar (<*ardu) “horn” corresponding to Avar ċar “horn” (cf. Burushaski -lur “horn”).

18. This expressive word, cited from western Basque, also has the eastern variants tsipi~tti “small”, so this example may be doubtful.

19. *Hni- is based on the Zuberoan form, which seems the most archaic. Other Basque meanings, e.g., “spark; dead stalk of wheat; blight”, are probably related to an old fire-making complex (spark, tinder, ember), which is confirmed by more remote comparisons (e.g., in Na-Dene, Eyak lid “deadwood, firewood”). Cf. V. BLAZÉK & J.D. BENGTSON, “Lexica Dene-Caucasica” (to appear in Central Asiatic Journal).

20. This must be admitted to be one of the most straightforward Vasco-Caucasian comparisons: semantically and phonetically perfect.
21. See notes 11, 12, and 16 for phonology. There is also a stem variant sat-(as in sat-(h) or "mole", lit. "mouse-dog") which is reminiscent of Chechen sat'q'a "weasel" (cf. note 22). Note also western Basque sagu~sat- "mouse".

22. The meanings "marten, mouse" are found in West Caucasic (e.g., Adyge c∂γwa "weasel"). Cf. also Burushaski cvharge "girder" (*Gwérć'V ~ *qwérć'V) 21.

23. This is a Bizkaian form, another meaning of which is "a shrub that resembles the juniper" (homonym? cf. R. M. de AZKUE, Diccionario Vasco-Español-Francés, Bilbao, 1905).

24. Other meanings include "beam" (Avar çarc'v') and "stick" (Batsbi γoc'v'); cf. Burushaski gache "twig, withe" (BLAŻEK & BENGTSON, Op. cit.).

25. Analyzed as u-k(h) ondo, u- being another fossilized class/gender prefix (cf. note 10). The analysis as uk(h)~ondo or uk(h)~a~ondo (TRASK, Op. cit., ps 45) is set aside, since *ukho- is only hypothetical, and ukalondo is a central Basque innovation, not a relic. The Basque-Caucasic comparison is semantically and phonetically precise.

26. Here NIKOLAEV & STAROSTIN (Op. cit., 1992) reconstruct a final front vowel. If so, the development to -o on both sides of the equation could be convergent, due to the influence of initial *b.-

27. E.A. BOKAREV, Czescie (Didzejkie) Jazyki Dagestana, Moscow, 1959, ps. 241. Cf. also Burushaski babar-un "hot (of taste, temper)".

The basic nature of these 23 comparisons disposes of the possibility that these resemblances could be the result of borrowing or diffusion. The remaining explanations are accidental resemblance or genetic inheritance. Regarding the first possibility, is it plausible that these many coincidences of sound and meaning, extending even to the same (or similar) stem vowel, could be accidental?

If this possibility of accidental resemblance is still entertained, consider that there are also regular sound correspondences running through these comparisons. Here I shall only point out a few of the most interesting non-trivial correspondences:

Basque -r- corresponds to Proto-Caucasic -1-:

(2) Basque *berγi = Caucasic *?willi
(6) Basque *-seγari = Caucasic *chwolč
(7) Basque [cori] = Caucasic *c’HwiiV
(23) Basque bero = Caucasic (Khwarshi) bobolu

These four are not the only examples of this correspondence. Consider also:

24. Basque ahur (*a-xur) : Khinalug, Archi, Udi kul “hand”; Avar kwer “hollow of the hand” “hand”; Hunzib koro, etc. (*kwilči).

It is obvious from these comparisons that the Basque development (*1>r) is parallel to developments within the Caucasian family: the Avar-Andi-Tsez languages have -r-, while Nakh, Lezgian, Dargi and Lak preserve -1-.

Another correspondence is that of Basque b to Proto-Caucasic *w:

(2) Basque *berγi = Caucasic *?willi
(22) Basque baso = Caucasic *wice

and possibly also:

(16) Basque habe = Caucasic *hwVwV

though here Nikolaev and Starostin allow for the alternative reconstruction *hwVbhV. Here again, the Basque development (*w > b) is shared with most of the Caucasian family. Only some Lezgian languages preserve w or v (cf. Lezgi vil “eye”, cognate with no. 2).

A further example of Basque b = Caucasian *w is probably the frequent Basque prefix be-/bi-, which appears to be cognate with the Caucasian gender marker *w- (third class singular, first and second class plural). One example is given above:

(3) Basque *be-lari = Caucasic *(w-)léHli

Additional examples include:

26. Basque *be-xacγ > : Basque [be]-xacγ, behatz-beatz “toe - paw - hoof - paw - thumb” “distance between thumb and forefinger”.

Caucasic *(w-)kwač ‘ė “paw”: Avar kwač’

kwač’a “paw”; Dargi k’ac’a “paw”; Hunzib k’ox’u
27. Basque *be-ašum: Caucasic *(w-)cwäjme “gall, anger”: Avar čin, Tindi, Botlikh, Godoberi šimi; Dargi sumi ~ himi time; Chechen stim “gall”.
28. Basque bi-zi [biši]: Caucasic *(w-)stįwV: Chechen sa “soul, breath”; Lak šiši “breath, vapor”, West Caucasic *p-sV “soul, breath”.

We have seen that these lexical parallels between Basque and Caucasic are far more than mere “look-alikes”. There are several recurrent sound correspondences, and the words even agree in their final (stem) vowels. I suggest that resemblances such as:

(1) Basque mihi : Andi mič’i
(4) aragi : rit’i
(14) ilhinti : ludi
(17) sagu : sart’u
(18) [garc’u] : q’erc’u

are every bit as significant as those between Latin poti-, pecu, and Sanskrit pāti-, pāśu, respectively. These are cognate words of a language family that should, however belatedly, be recognized as valid.

These close correspondences could even give us some clues about the time depth of this family, Macro-Caucasic (or Vasco-Caucasic). This evidence suggests to me that the differences between Basque and Caucasic are comparable to those between different subgroups of Indo-European, especially those that have developed at some distance from one another, say Germanic and Armenian. The etymologies (5) “cow, cattle” and (25) “wheat” imply that the pre-Basques and the pre-Caucasics shared a simple Neolithic farming and herding culture before their dispersal. This is completely in accord with René Lafon’s conclusion that the Basque language was brought to Iberia by immigrants from northern Anatolia, around the end of the Third Millennium B.C.

The linguistic evidence supports the conclusion that the relationship between Basque and Caucasic is not “extremely remote”, as some scholars even among those who accept the relationship have assumed. To be sure, much that the languages once held in common has long since been swept away by millennia of attrition, but many clear traces still remain, as has been shown in this essay.

TRANSCRIPTION GUIDE:

ä, é, etc. nasalized vowel
ã, ë, etc. pharyngalized vowel
a, e, etc. prosodic condition of vowel (tense voice?)
c, c’, 3= [ts, ts’, dz] dentalveolar affricates
c̃ = [t̃] Basque dorso-alveolar affricate (orthographic ťz)
c̃ = [t̃] Basque apico-alveolar affricate (orth. ts)

CORRESPONDENCES OF BASQUE AND CAUCASIC FINAL STEM VOWELS

c = [ts] Basque prepalatal affricate (orth. tx)
c c retroflex affricate (Burushaski)
G = 9 voiced uvular stop
γ (Basque) voiced velar fricative; (Caucasic and Burushaski) voiced uvular fricative
h glottal fricative (audible in “French” Basque, silent in “Spanish” Basque)
H = h voiceless pharyngal fricative
η = ’ glottal stop
ξ pharyngal stop
c voiced pharyngal fricative
H “laryngeal” of undetermined quality
i high mid vowel
l lateral sonant (possible velar or “dark” l)
l voiceless lateral fricative (hl)
d’ = 8 voiced lateral affricate
t’, t’ = 3, 8 voiceless lateral affricates
p, t, k, s, etc. = p, t, s, etc. fortis (“emphatic”) consonants
p’, t’, k’, s’, etc. glottalized consonants
q, q’ voiceless uvular stops
r Basque trilled resonant (orth. rr)
s Basque dorso-alveolar fricative (orth. z)
s Basque apico-alveolar fricative (orth. s)
s prepalatal (postalveolar) fricative (Basque orth x)
s retroflex fricative (Burushaski)
x (Basque) voiceless velar fricative; (Burushaski) voiceless uvular fricative
X voiceless uvular fricative (= Burushaski x)
ü “French” Basque (and Chechen, Lezgi) fronted u
V reconstructed vowel of undetermined quality

DEFINITIONS:

Burushic = a small family in the far northwest of greater India, now represented by the Hunza-Nagir and Yasin dialects of Burushaski. A Burushic substratum can also be traced in certain Indo-European languages, e.g. Khowar and Shina.

Caucasic = native languages of the Caucasus region, exclusive of Indo-European, Altaic (Turkic, Mongolic), Semitic, and Kartvelian languages. They can be positively described as (North) Caucasic, i.e. Abkhazo-Adygan and Nakho-Dagestanian, as well as the extinct Hurrian, Urartian, and Hatti.

Dene-Caucasic = a macrophylum consisting of the Vasco-Caucasic (see below), Sino-Tibetan, Yeniseian, and Na-Dene phyla or families. The possible inclusion of Sumerian and Kusunda is less certain. (= Dene-Caucasian, Sino-Caucasian)

Vasco-Caucasic (= Macro-Caucasic, Macro-Caucasian) = a phylum level subdivision of Dene-Caucasic, consisting mainly of Basque, Caucasian, and Burushic, as well as extinct languages such as Aquitanian, and possibly Iberian.

LABURPENA

Euskal-kaukasiar (Euskera + (Ipar) kaukasiar hizkuntza + Burushic hizkuntza) hipotesiari buruzko ikerketa berri batek kidetasun lexikoak,
erro-bukaerako bokalak motako egokitzapen zehatzak euskera eta kaukasia hizkuntzaren artean azaltzen dituztenak, erakutsi ditu. Adibidez:

- mihi (euskera) : Botlikh mič’i “mihi”
- behi (euskera) : Botlikh buč’i “abere”
- beso (euskera) : Bezhta bico “beso”
- baso (euskera) : Bezhta bizo “mendi”

Egokitzapen hauek Euskera, Kaukasia hizkuntza eta Burushic hizkuntza Neolitiko (8.000-10.000 urte K.A.) Aroan kokatzen den aurrehizkuntza amankomun batetik datozelako teoriarri ematen diote sostengu.

RESUMEN

Un reciente estudio sobre la hipótesis Vasco-Caucásica (Vascuence + Cáucaso (Septentrional) + Burushic) reveló afinidades léxicas que dan fe de correspondencias precisas de raíz-vocales finales entre el Vascuence y el Caucásico. Por ejemplo:

<table>
<thead>
<tr>
<th>Vasco</th>
<th>Vascuence</th>
<th>Botlikh</th>
<th>Cáucaso</th>
</tr>
</thead>
<tbody>
<tr>
<td>mihi</td>
<td>“lengua”</td>
<td>mič’i</td>
<td>“lengua”</td>
</tr>
<tr>
<td>behi</td>
<td>“vaca”</td>
<td>buč’i</td>
<td>“ganado”</td>
</tr>
<tr>
<td>beso</td>
<td>“brazo”</td>
<td>bico</td>
<td>“brazo”</td>
</tr>
<tr>
<td>baso</td>
<td>“bosque”</td>
<td>bizo</td>
<td>“montaña”</td>
</tr>
</tbody>
</table>

Estas correspondencias apoyan la teoría de que el Vascuence, el Cáucaso y el Burushic proceden de un protolenguaje común datado en el Neolítico (edad cronológica 8.000-10.000 antes del presente).

RÉSUMÉ

Une étude récente sur l’hypothèse Basque-Caucasienne (Basque + (Nord) Caucasien + Burushic) a révélé des affinités lexiques qui font foi de correspondances précises de racines-voyelles finales entre le Basque et le Caucasicien. Par exemple:

<table>
<thead>
<tr>
<th>Basque</th>
<th>Vascuence</th>
<th>Botlikh</th>
<th>Caucasic</th>
</tr>
</thead>
<tbody>
<tr>
<td>mihi</td>
<td>“tongue”</td>
<td>mič’i</td>
<td>“tongue”</td>
</tr>
<tr>
<td>behi</td>
<td>“cow”</td>
<td>buč’i</td>
<td>“cattle”</td>
</tr>
<tr>
<td>beso</td>
<td>“arm”</td>
<td>bico</td>
<td>“arm”</td>
</tr>
<tr>
<td>baso</td>
<td>“forest”</td>
<td>bizo</td>
<td>“mountain”</td>
</tr>
</tbody>
</table>

Ces correspondances confirment la théorie sur quoi le Basque, le Caucasicien et le Burushic proviennent d’un proto-langage commun daté du Néolitique (entre 8.000 et 10.000 ans avant J.C.) 2.º

ABSTRACT

Recent research on the Vasco-Caucasic hypothesis (Basque + [North] Caucasican + Burusichic) has revealed lexical cognates which attest to precise correspondences of stem-final vowels between Basque and Caucasic, e.g.

<table>
<thead>
<tr>
<th>Basque</th>
<th>Vascuence</th>
<th>Botlikh</th>
<th>Burushic</th>
</tr>
</thead>
<tbody>
<tr>
<td>mihi</td>
<td>“tongue”</td>
<td>mič’i</td>
<td>“tongue”</td>
</tr>
<tr>
<td>behi</td>
<td>“cow”</td>
<td>buč’i</td>
<td>“cattle”</td>
</tr>
<tr>
<td>beso</td>
<td>“arm”</td>
<td>bico</td>
<td>“arm”</td>
</tr>
<tr>
<td>baso</td>
<td>“forest”</td>
<td>bizo</td>
<td>“mountain”</td>
</tr>
</tbody>
</table>
These correspondences support a hypothesis that Basque, Caucasic, and Burushic derive from a common proto-language dated in the Neolithic (ca. 8,000-10,000 years B.P.).