Comparing Basque and Proto-Indo-European,  
a preliminary phonetic survey

Arnaud Fournet

Abstract: The paper describes the phonetic correspondences that can be found between Basque and Proto-Indo-European (PIE). More than 30 items of the basic vocabulary are presented together with a table of sound correspondences. Basque is usually considered to be an isolate but it would appear that it actually contains words that have clear PIE affinities.

Keywords: Basque, Proto-Indo-European, Nostratic.

1. Introduction

So far Basque has most often been compared with the so-called “Caucasian” languages and some people have claimed it is indeed related to these languages. For example Bengtson (2003:23) quite boldly states: “Basque and the languages that are most closely related to it, namely the (North) Caucasian languages, and Burushaski.” Trask (1997:35), a well-known specialist of Basque, was equally assertive in the opposite and negative direction: “Basque is a genetically isolated language: there is not the slightest shred of evidence that it is related to any other living language.” On the whole the idea that Basque and Indo-European languages have about nothing in common but loanwords of Latin and late Romance dating is widespread. Bengtson (2005:33) again: “families and isolates of northern Eurasia (and extending into North America), negatively defined as those that did not fit into the developing hypotheses of Afro-Asiatic (Hamito-Semitic) and Nostratic macro-families. These entities are Basque, Caucasian, Burushaski, Yeniseian, Sino-Tibetan, and Na-Dene.” In other words if Basque happens to be related to some other linguistic group then it is preferably or probably not the Indo-European family. In spite of this widespread belief the article tackles the issue of comparing Basque and PIE, knowing that a body of potential cognates has already been proposed between Basque and Caucasian. As a preliminary step, the article deals with words and sound correspondences and leaves aside morphological and grammatical issues.

2. Proto-Indo-European (PIE)

From a descriptive point of view, PIE is a set of graphemes, rules and items that account for the recurrent phonetic isomorphisms that pervade the vocabulary and morphology of some languages like Latin, Greek, Sanskrit, English, etc., that are classified as Indo-European for that specific reason. In a fictionalist approach, PIE could just be considered as a convenient model that generates a large stock of words and grammatical paradigms of the Indo-European languages. But most people, of whom I am one, hold PIE to be a prehistorical language which must have existed sometime and somewhere in the past and which thereafter split and evolved into the attested Indo-European languages. The most simple set of graphemes necessary to describe PIE in terms of correspondences is:

- labials *(b) *p *bh *m
- dentals *(d) *t *dh *n

1 The original reads <Caucasian>.
Comparing Basque and Proto-Indo-European

- palatals  *g *k *gh
- velars  *g *k *gh
- labio-velars  *gw *kw *ghw
- non-obstruents  *w *s *y *r *l
- laryngeals  *H1 *H2 *H3 *H4
- vowels  *e *o *i *u

This set of graphemes requires some comments, the first of which is the nature of these graphemes from a linguistic and phonological point of view. I deliberately used the word grapheme in order to emphasize that I do not accept the usual approach, implicit in the Indo-Europeanist communis opinio, that each grapheme of the PIE modelisation should be bijectively seen as one single proto-phoneme. Each grapheme contains a hypothesis about the phoneme(s) it is supposed to stand for. As regards the phonetic values of some graphemes, the following comments can be made:

- the distinction between *k, *g, *gh and *ğ, *ğh is considered by some people like Bomhard (2008) to be a secondary development of Proto-(Central)-IE. In all cases there is no need for a distinction between palatals and velars when PIE and Basque are compared. They are undistinguishable as is the case for most IE languages.

- the vowel traditionally written PIE *e was phonemically /a/ in my opinion. In the article, the grapheme PIE *e is to be understood as standing for the proto-phoneme */a/. The vocalic phonemes hypothesized for PIE and Basque are therefore *i, *u, *o and *a (which is written <e> in the PIE tradition). As will be shown below Basque <e> most often comes from an original *a inflected by a neighboring i or j.

- It has been proposed that the voiced phonemes of traditional PIE were glottalized and that the so-called voiced aspirate series was marked by plain voice: for example *b stands for [pˀ], *g for [kˀ], etc. This is known as the Glottalic Theory. Whether this approach true or not, it can be noted that the phonetic reality of PIE graphemes is irrelevant as the comparative method is about comparing items no matter what the exact phonetic substance is or is theorized to be.

3. The basic features of Basque when compared to PIE

As a general rule words underwent a number of phonetic changes which account for the typical outlook of Basque. The apparent opacity of Basque words results from the following changes:

- Sibilantization of liquids: PIE *l and PIE *r > Basque <z>, but PIE *s > Basque <s>,
- Complete shift of the dentals: PIE *d > Basque /l/ (or /r/ when preceded by a back-vowel), PIE *dh > Basque /rr/, but the cluster PIE *n-t- > Basque /dl/,
- Devoicing (or deglottalization) : PIE *b > Basque /p/, PIE *g > Basque /k/ (or rarer /g/),
- (Aspirate) voiced stops: PIE *gh(w) > Basque /h(u)/, PIE *dh > Basque /rr/. There is no clear instance of PIE *bh.
- Spirantization of voiceless stops: PIE *p ~ t ~ k > Proto-Basque *β ~ *ð ~ *ġ. The spirants later evolve either to stops: *b ~ *d ~ *g or to *u ~ *(t)z ~ *h.
- A frequent feature of Basque is the metathesis of CvCv into vCCv. In general this metathesis applies last, after the regular phonetic changes but there are some exceptions.
- Not infrequently laryngeals have been preserved in Basque (dialects). Basque /h/ is either PIE *H or PIE *k > . Needless to say that aspiration is not a recent development of Basque dialects.
These features account for the typical look of many Basque lexemes and they make it clear which words are inherited and which are borrowed. Anytime a Basque word has *r or *l where PIE also has *r and *l, it can only be a borrowing or a wanderwort.

4. Examples of cognates

1. Basque *adar ‘horn’ ~ PIE *H₂-ent ‘before, front’ (Pokorny 48-50). Cf. Bomhard (2008:II.665): ‘Root 629. Proto-Nostratic *h₁ent- ‘the most prominent or foremost (person or thing), front, front part’.” The reconstruction of *r instead of *n is based on the (possibly erroneous) addition of Uralic data. The Basque word has often been compared with Old Irish adarc, for which no PIE etymology exists but the potential connection does not account for Irish final -c.


8. Basque *edan ‘to drink’ ~ (?) PIE *H₁egw ‘to drink’ (Pokorny 808). Not in Bomhard (2008). The suffix -an also appears in *erran ‘to say’. The word would be the only example of *gw > Basque *d, which an unusual and hardly acceptable sound change: *i-Hagw-i- > *i- günd-i > (?) ed-. A better etymology for edan is probably *im-t-.

9. Basque *i/e-kusi ‘to see’ ~ PIE *Hokw ‘eye’ (Pokorny 808). Not in Bomhard (2008). The initial stop in Basque is a problem as -hu- is expected. This problem can be addressed if one accepts that -k- and -H- were metathesized in *i-H(a)ku- > *i-kHu-. The phonetic structure of the root probably explains why the phonetic development is irregular. The regular change would have yielded *i-hāhu.


13. Basque *eseri ‘to sit’ ~ PIE *sed (Pokorny 732), *H4es-. Cf. Bomhard (2008:762): “Root 723. Proto-Nostratic *#es- (~ *#es-) ‘(vb.) to put, to place, to set; to sit, to be seated; (adj.) put, placed, set, established; (n.) place, seat.’” Bomhard (2008) does not in fact include *sed in this root. The derivation is *i-H#es-i > *i-iës-i > ese(ri). A very old verb with the native Basque conjugation.


17. Basque *gibel / bigel ‘liver’ ~ (?) PIE *yekw (Pokorny 504). Not in Bomhard (2008). The derivation might be *pi-yäkw > *bi-yähu- > (?) *biña-. The phonetic details are unclear.


21. Basque *hauts ‘to break, tear’ ~ PIE *(s-)kē-(t) ‘to cut, hack’ (Pokorny 954). Cf. Bomhard (2008:493): “Root 463. Proto-Nostratic *k`w- (~ *k`w-₅d) ‘(vb) to strike, to beat, to smash, to pound; (n.) knock, stroke, thrust.’” This root should be resegmented *k`w₅d. The derivation is *kaHu-t- > haut-s- > hauts. This item is in semantic interference with item 21 ‘ash, powder’, which also explains a phonetic mutual contamination.

22. Basque *hau ‘this, that’ ~ PIE *ko (or *gho) (Pokorny 609). Cf. Bomhard (2008:762): “Root 347. Proto-Nostratic *k`u- (~ *k`u) *k`i- (~ *k`e) *k`u- (~ *k`o) demonstrative pronoun stem (*k`a- appears to have been proximate, *k`/i-intermediate, and *k`u- distant).”

23. Basque *haz ‘seed, to grow’ ~ PIE *gher / *ghreH ‘to grow, green, grain’ (Pokorny 454). Cf. Bomhard (2008:343): “Root 324. Proto-Nostratic *gar – (~ *gər) ‘(vb.) to swell, to increase, to grow.’” The root is only attested in Dravidian and PIE. The derivation is *gar > haz. Initially the word had no final -i. Cf. heze ‘green’.

Initially the word must have had a final -i. It is possible that the verb was originally *i-häl-i and was reshaped as *hel-du.

25. Basque *hel ‘to come, to arrive’ ~ PIE *H₁ed ‘at, toward’ (Pokorny 3). Cf. Bomhard (2008:241): “Root 224. Proto-Nostratic *dʰʰi₃- (~ *dʰʰe₃-)'(vb.) to reach, to arrive at, to come to; to surpass, to exceed’.” Bomhard (2008) does not list the PIE word. The verbal and prepositional uses probably explain the morphological complications of this root. This root is homophonous with item 25 and displays a similar phonetic development.


30. Basque *(h)il ‘moon’ ~ Latin iːdus ‘mid-month’ (not in Pokorny). Not in Bomhard (2008). The Latin word is isolated and could be held to be a loanword but Egyptian jâḥ and Coptic 𝑤 ʿ moon(-god) show that this root can in fact be old.


34. Basque *ontsa ‘well, good’ ~ (?) PIE H₁esu (Pokorny 342). Not in Bomhard (2008). The derivation is unclear.


37. Basque *larru ‘skin’ ~ PIE *der- (Pokorny 206). Not in Bomhard (2008). The derivation is *dar-dhu > *laʾrru > larru (with deletion of *z).

38. Basque *luze ‘long, tall’ ~ PIE *dol- (Pokorny 196). Cf. Bomhard (2008:216): “Root 196. Proto-Nostratic *tʰal- (~ *tʰəl-) ‘(vb.) to stretch out, to extend; (n.) length; height; (adj.) long, tall; high’.
The derivation is *doli > *lözi > luze. The vocalism is irregular. Basque larri < *del-dh ‘big, large’ may also belong here.

39. Basque *{(h)orri ‘leaf’ ~ PIE *H2-endh ‘flower, grass’ (Pokorny 40). Not in Bomhard (2008). The derivation is *Hondhi > *ho²-rii > (h)orri. Initially the word had no final -i.


41. Basque txori ‘bird’ ~ PIE lēu- ‘to sing’ (Pokorny 683). The derivation is *laHu-d-i > *zauli > *zori > txori (with expressive emendation).


43. Basque *ziho ‘fat’ ~ (?) PIE *leip (Pokorny 670). Not in Bomhard (2008). The derivation is easier if Greek lipos is assumed to be *likwos. In that case Basque is: *likwo > *zihuo > ziho.

44. Basque *zorri ‘louse’ ~ PIE *lus (Pokorny 692). Not in Bomhard (2008). The derivation is *lus-dh > *zu’rr > zorri. The vowel seems to be irregular. Initially the word had no final -i.

5. Table of sound correspondences

<table>
<thead>
<tr>
<th>Basque</th>
<th>PIE</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>b</td>
<td>*ezpain ‘lip’ ~ PIE *leb</td>
</tr>
<tr>
<td>b / u</td>
<td>p</td>
<td>*begi ‘eye’ ~ PIE *s-pék- ‘to see’ ; *i-bil-i ‘to walk’ ~ PIE *ped ‘foot’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*belhaun ‘knee’ ~ PIE *ped + *genu ; *bide ‘road’ ~ PIE *pent</td>
</tr>
<tr>
<td>?</td>
<td>bh</td>
<td></td>
</tr>
<tr>
<td>o / u</td>
<td>w</td>
<td>*hobi ‘gums’ ~ PIE *gheu</td>
</tr>
<tr>
<td>l</td>
<td>d</td>
<td>*i-bil-i ‘to walk’ ~ PIE *ped ‘foot’ ; *belhaun ‘knee’ ~ PIE *ped + *genu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*hil ‘moon’ ~ Latin i:dus ‘middle of month’ ; *luze ‘long, tall’ ~ PIE *dol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*hel ‘to come, to arrive’ ~ PIE H₂ed ‘at, toward’</td>
</tr>
<tr>
<td>*l &gt; r</td>
<td>d</td>
<td>*euri ‘rain’ ~ PIE *(H)wed ; *hur ‘water’ ~ PIE *(H)wed</td>
</tr>
<tr>
<td>*l &gt; r</td>
<td>d</td>
<td>*haritz ‘oak-tree’ ~ Greek ked-ros</td>
</tr>
<tr>
<td>d (tz)</td>
<td>t</td>
<td>*bide ‘road’ ~ PIE *pent ; *adar ‘horn’ ~ PIE H₂ent ‘head’</td>
</tr>
<tr>
<td>rr</td>
<td>dh</td>
<td>*erre ‘to burn’ ~ PIE H₃aiddh ; *(h)erro ‘udder’ ~ PIE H₂eudh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*herri ‘inhabited place’ ~ PIE *gherdh ; *(h)orri ‘leaf’ ~ PIE H₂endh ‘flower’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*erran ‘to say’ ~ PIE *menth ‘mouth’</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>*eseri ‘to sit’ ~ PIE H₂es/sed ; *hauts ‘ashes’ ~ PIE H₂es</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>*elhur ‘snow’ ~ PIE *(s-)neighw</td>
</tr>
<tr>
<td>z (tx)</td>
<td>l</td>
<td>*ezpain ‘lip’ ~ PIE *leb ; *zorri ‘louse’ ~ PIE *lus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*haize ‘wind’ ~ PIE H₂ew-il-a ; *luze ‘long, tall’ ~ PIE *dol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*txiki ‘small’ ~ PIE Heig ; *zanko ‘leg’ ~ PIE *legh</td>
</tr>
<tr>
<td>(t)z</td>
<td>r</td>
<td>*bi-hotz ‘heart’ ~ PIE *kor-d- ; *haritz ‘oak-tree’ ~ Greek ked-ros</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*hazi ‘seed, to grow’ ~ PIE *gher / *ghreH₂-no ; *heze ‘green’ ~ PIE *gher</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*hertz ‘tooth’ ~ PIE H₁od-nt-</td>
</tr>
<tr>
<td>k</td>
<td>g</td>
<td>*txiki ‘small’ ~ PIE Heig</td>
</tr>
<tr>
<td>?</td>
<td>gw</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>k</td>
<td>*bi-hotz ‘heart’ ~ PIE *kor-d- ; *haritz ‘oak-tree’ ~ Greek ked-ros</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*hauts ‘to break, tear’ ~ PIE *(s-)keu- ‘to cut, hack’</td>
</tr>
<tr>
<td>nk</td>
<td>nk</td>
<td>*zanko ‘leg’ ~ PIE *legh</td>
</tr>
<tr>
<td>ku gu</td>
<td>kw</td>
<td>*i/e-kusi ‘to see’ ~ PIE *Hokw ‘eye’</td>
</tr>
</tbody>
</table>
In the Swadesh 100-word list 30 items are based on roots shared between Basque and PIE. They are underlined in the following list (unclear cases are omitted):

all, ashes, bark, belly, big, bird, to bite, black, blood, bone, breast, to burn, claw, cloud, cold, to come, to die, dog, to drink, dry, ear, earth, to eat, egg, eye, fat-grease, feather, fire, fish, to fly, foot, full, to give, good, green, hair, hand, head, to hear, heart, horn, to kill, knee, to know, leaf, liver, long, louse, lying, man-male, many, me I, meat-flesh, moon, mountain, mouth, name, neck, new, night, nose, not, one, rain, red, road, root, round, sand, to say, to see, seed, sitting, skin, to sleep, small, smoke, standing, star, stone, sun, to swim, tail, that, this, thou, three, tongue, tooth, tree, two, to walk, warm, water, we us, what, white, who, woman, yellow.

It can be noted that a handful of these 30 items have simultaneous cognates in Caucasian. There are three possible conclusions to be drawn from the present survey:

1. These correspondences are due to chance coincidences.

2. Basque contains a significant input of Indo-European words, that entered the language at a very early period, so that they underwent typically Basque opaque changes.

3. Basque is much closer to PIE than has ever been considered so far.

Option 1 seems hard to believe. Historical linguistics is about “chance coincidences” that have a genetic relevance. Option 2 is also hard to believe. How come the Indo-European input does not sound Celtic or Italic but is coherent with PIE itself? So my personal choice is the last option. Contrary to a hackneyed cliche, Basque is related to its present-day neighbors. It can be added that Basque shares morphological features with PIE:


(2) nasal infix *-n-: *(H)wonda > hodai ‘cloud’, *lanko > zanko ‘leg’.

(3) participle *-nt-: *(H)od-nt- > hortz ‘tooth’.

So far Basque has been preferably compared to Caucasian assuming the hypothesis that PIE was not at all a possibility, but this premise can be shown to be unacceptable.
Comparing Basque and Proto-Indo-European

References


