

# ‘Where there is fire, there is smoke’: A study in the Euskaro-Caucasian hypothesis<sup>1</sup>

JOHN D. BENGTON

**ABSTRACT:** The proposed Euskaro-Caucasian linguistic relationship between Basque and North Caucasian is demonstrated not only by basic words, grammatical morphemes, and phonetic correspondences, but by *lexical families*, complexes of semantically interrelated words, and by shared grammatical homologies that are unlikely to have arisen by chance. One such family includes the words for ‘smoke’ (Basque \**e=kē* ~ Proto-North Caucasian \**kwīnhV*) and ‘fire’ (Basque \**zu* / \**i=zu* ~ Proto-North Caucasian \**čǎjī* / \**čǎj-*), and other words relating to the production of fire (‘ember’; ‘firewood’; ‘tinder’; ‘ashes’; ‘soot’). The basic necessity of fire, and methods of kindling fire, did not essentially change until recent times. The analyses of these words also reveal traces of archaic Euskaro-Caucasian morphological features (class prefixes, oblique stem markers, and allomorphs) that are now lexicalized in Basque. “Where there is fire, there is smoke” can also serve as a metaphor for the observation that where there is a genetic linguistic relationship (“fire”), there are also lexical, phonological, and grammatical features (or their traces = “smoke”) that verify the existence of the former. This type of study, emphasizing interrelations among linguistic structures, and their relationships to social, anthropological and historical structures, can be traced back to influences from the Prague School.

**Key words:** Euskaro-Caucasian, Basque, North Caucasian, lexical family, fossilized morphemes, noun classes (genders), allomorphs, anthropology of fire, Prague School

It is often assumed among Vasconists, and among historical linguists in general, that Luis Michelena definitively dismissed the Euskaro-Caucasian hypothesis, but perhaps few know that he in fact held out some hope that it might someday be demonstrated. Michelena was keenly interested in the possibility of external relations of his native language, and several of his writings addressed these issues. The last of these was his *Sobre historia de la lengua vasca*, published in 1988, soon after his death. We may regard this book as his last word on the question of external relations of Basque.

In spite of some lexical comparisons he found “extremely tempting”, and some grammatical parallels that were “impressive”, Michelena concluded that the theory of “Basque-Caucasian linguistic kinship has been shown to be, up to now, singularly unfruitful”. Nevertheless, he was careful not to rule out the possibility of a fuller demonstration and held out some hope for future progress:

It is possible that we will always lack certain lost links that would be indispensable in order to reveal the mystery [of genetic affiliation of Basque]. It is also possible, but not very probable, that such links exist although we have not been able to recognize them [...] Also it is possible that an improvement in current methods or the discovery of new techniques might put more powerful instruments in our hands. [...] We hope that Fortune, always the friend of the bold and persevering, will see fit to smile upon us.<sup>2</sup>

---

<sup>1</sup> The title quotes Basque author Manuel de Larramendi (1690–1766): *Non sua dan, kea bada an*, roughly ‘where there’s fire there’s smoke’ (cited in OEH under *KE*). In modern standard Basque (*euskara batua*) the last word *an* is written *han* ‘there’.

<sup>2</sup> Some quotations in English included here are thanks to the translation by Linda White and M. Dean Johnson (in Salaburu 2008).

It is true that in Michelena's time, up to the 1980s, the evidence for the Euskaro-Caucasian hypothesis was, at best, fragmentary and disjointed. On the Caucasian side there were no detailed compendia of North Caucasian (NC) or Kartvelian etymologies, or of comparative grammar (other than Dumézil 1933) that were accessible to western scholars. Comparative dictionaries of North Caucasian (Abdokov 1981, 1983) and Kartvelian (Klimov 1964) existed, but only in Russian, and Abdokov's works were not available in time to be used by the most prominent Euskaro-Caucasian researchers, by then mainly René Lafon (†1974) and Karl Bouda (†1979). The genetic cleavage between North Caucasian and Kartvelian was not yet well known, so Lafon and Bouda freely compared Basque words and grammar with languages of both families.

The situation is much different today. A *North Caucasian Etymological Dictionary* (NCED), in English, with a comprehensive reconstruction of its subgroups and Proto-North Caucasian, has been available since 1994.<sup>3</sup> Michelena's hope for "improvement in current methods" and "discovery of new techniques" to "put more powerful instruments in our hands" has been fulfilled by a better understanding of basic vocabulary for use in genetic comparisons, such as the 50-item "ultra-stable list" (see G. Starostin 2010). As shown in this paper, the "lost links" mentioned by Michelena are gradually being discovered, and a comprehensive monograph, including comparative lexicon, phonology, and morphology, has recently been published by this author (Bengtson 2017a).<sup>4</sup> From this larger work I have excerpted the following discussion of words for 'smoke' (Basque \**e=kē*; PNC \**kwīnhV*) and 'fire' (Basque \**śu* / \**i=ću*; PNC \**čājī* / \**čūj-*), and other words relating to the production of fire.

The current paradigm of Euskaro-Caucasian (EuC), developed since the 1990s by this author, is a refined and methodologically updated version of the old hypothesis developed by C. C. Uhlenbeck, Alfredo Trombetti, Lafon, Bouda, Dumézil, and their predecessors (see Bengtson 2017a: 1–13). In this advanced version Basque is genetically linked specifically with North Caucasian, but excluding Kartvelian ("South Caucasian").<sup>5</sup> (See e.g. Bengtson 1996, 2004, 2008, 2009; Bengtson & G. Starostin 2015.) This model is the result of several scholarly streams directly or indirectly derived from the Prague Linguistic Circle. (See Appendix D for details.)

<sup>3</sup> As a pioneering effort, NCED has received some severe criticism. Nevertheless, lacking any better alternative, it is used by me as a working model. As S. A. Starostin himself stated, "Now, in December 1993, the work seems to be completed. There is still very much to be done, both in comparative phonology and in etymology of North Caucasian, and I am quite sure that 'Addenda et corrigenda' will follow, but I certainly feel that the publication of what we know so far is necessary" (NCED: 5). As to the North Caucasian family itself, still doubted by some Caucasologists, a prominent authority "deems the relationship between the WC and EC families as basically proven, thanks to works by N. Troubetzkoy, G. Dumézil, A. Shagirov, B. Balkarov, and especially A. Abdokov, S. Starostin and S. Nikolayev" (Chirikba 2016: 6).

<sup>4</sup> A corpus of more than 600 Basque etymologies, with external comparanda, has recently been revised and is available on the Tower of Babel / EHL webpages: <<http://starling.rinet.ru/main.html>> or <<http://ehl.santafe.edu/main.html>>.

<sup>5</sup> This model of EuC excludes Kartvelian languages (Georgian, Mingrelian, Laz, Svan). The North Caucasian languages included in EuC include WC (Abkhaz, Abaza; Adyge, Kabardian, †Ubykh) and EC (Batsbi, Chechen, Ingush; Avar; Andi, Akhwakh, Karata, Botlikh, Godoberi, Chamali, Bagwali, Tindi; Tsezi, Hinukh, Bezhta, Hunzib, Khwarshi; Lak; Dargi; Khinalug; Archi, Tabasaran, Lezgi, Agul, Udi, Kryz, Budukh, Rutul, Tsakhur) languages. Depending on the source consulted, spellings of the language names may differ.

## ‘Smoke’

The word denoting ‘smoke’ in Basque is universally some variant of /ke/, including EB *ke* ‘smoke’, with the aspirated form *khe* traditionally used by French Basque writers. In the far west (parts of Bizkaia [B] and Gipuzkoa [G]) and in the far east (the Vidángoz dialect of Roncalese [R], now extinct) the vowel is long: *kee* or *kē* ‘smoke’, which may become *kei-* before the definite article (*kei-a* ‘the smoke’); in more central parts of Basque Country the word is simply *ke* ‘smoke’, but in northern Bsq dialects (Lapurdian [L], Low Navarrese [BN], Zuberoan [Z]) the initial velar is strongly aspirated: *khe* ‘smoke’, and a variant with a voiced initial has been recorded, (G) *ge* ‘smoke’. In some central areas (parts of G and High Navarre [AN]) the vowel has been raised: *ki* ‘smoke’. Due to the peripheral attestation of *kee*, in this analysis it is considered to be the archaic state, thus we reconstruct Bsq *\*kē* ‘smoke’.<sup>6</sup> The long vowel also leads to a suspicion that a consonant was lost, in this case, based on external comparison, possibly a nasal, or *\*kē < \*\*kēe < [\*\*kenhe]?*, though there is no direct evidence in Basque for nasal vowels in this word. (But see some of the North Caucasian words cited below, and Bsq [R] *k̄ō* ‘[bad] odor’ = [L] *khino*, [Z] *khiño*, which seems to be a variant related to *\*kē*. [see Table 1])

Yet another peculiarity is attested in Baztanese (Bzt) and the Pyrenean dialects (Aezkoan [Azk], Salazarese [Sal], and R): the word has a prefix *e-*, i.e. *eke* ‘smoke’; this form has also been attested in adjacent parts of other provinces, AN (Erro, Ulzama, Bortzirirak) and BN (Amikuse, Baigorri). The Pyrenean dialects are otherwise known for having idiosyncratic and archaic forms. The coastal variety of Lapurdian has a form with an *i-*prefix, *ike* ‘smoke’. This prefix, which must be considered a fossilized morpheme, is common in other Bsq words denoting mass nouns (uncountables), e.g. *\*e=uli* ‘rain’, *\*e=thu-ř* ‘snow’, *\*e=thoři* ‘bramble’; the allomorph *\*i=* is also common in such nouns, e.g. *\*i=štu* ‘saliva’, *\*i=serdi* ‘sweat; sap’, *\*i=til* ‘pool’, *\*i=turi* ‘fountain’, *\*i=soc* ‘frost, ice’, *\*i=cal* ‘shadow’; see also below about Bsq *\*i=thinti* ‘ember’ (see Table 3) and the rare Bsq variant *itsu* (*\*i=ću*) ‘fire’. According to Trask, “very likely the widespread variant **eke** is conservative, and most varieties have undergone an irregular loss of the initial vowel” (EDB 248). If so, the most archaic form would be *\*e=kē*, with the fossilized class prefix. (See Appendix C, below, and Bengtson 2017a: 64–67 for an extensive discussion of the Bsq *\*e=*/*\*i=* fossilized prefix.)

In North Caucasian there is a widespread word for ‘smoke’ that in some of the languages is a rather close approximation of Bsq *\*kē*. In Avar and most Andian languages the word has a tense glottalized velar initial /kʰ:/, alternatively transcribed as  $\sqrt{k}ʰ$ ,  $\sqrt{k}/$ , or /k:/, thus Avar *k:uj* ‘smoke’, in the Chadakolob dialect *kʷi* (cf. central Bsq *ki*);<sup>7</sup> Andi *k:ʷoj*

<sup>6</sup> The principle that peripheral forms tend to represent earlier stages is well established in historical linguistics, thanks to insights gained from dialectology (e.g. Anttila 1972: 294–299). As used by Trask in relation to Basque, e.g. “The late origin of this gerund suffix [-tze] is consistent with its absence from the two peripheral dialects, B in the west and R in the east” (EDB 353); “it seems probable that the peripheral forms with /r/ are conservative, while the central forms with /l/ are innovations” (EDB 117); “The conservative longer variants are confined to peripheral varieties, east and west” (EDB 155).

<sup>7</sup> In regard to the citation of surface similarities between random forms from Bsq and modern NC languages, such as Central Bsq *ki* ‘smoke’ ~ Chadakolob *kʷi* ‘smoke’, and Bsq *su* (*\*śu*) ‘fire’ ~ Lak *çu* ‘fire’, this is,

‘smoke’, and some Andian languages have nasal vowels: Karata *k:ũji*, etc. (cf. Bsq [R] *kĩo* ‘[bad] odor’). In Tsezian languages the initial has regularly become the voiceless/aspirated uvular affricate *q*, phonetically [qʰ]: Bezhta, Hunzib *qo* [qʰo] ‘smoke’, with loss of the original nasal as in Avar, Andi, Lak, Dargi, WC, and Basque. In Lezgian languages only the peripheral Udi retains *k:uin* with the meaning ‘smoke’; the meaning is shifted in Archi *gum* ‘dust’. (See Table 1 for more details on attested forms.) From these forms Nikolayev & S. Starostin reconstructed PNC *\*k̥wĩnhV* ‘smoke’ (NCED 738). The Nakh languages, some Andian and Tsezian languages, and Lezgian languages (except Udi) have words denoting ‘smoke’ that cannot be derived from PNC *\*k̥wĩnhV* (see also SKJa 197).<sup>8</sup>

Both NCED and Chirikba (1996) agree that West Caucasian words for ‘smoke’ are cognate with PEC *\*k̥wĩnhV* ‘smoke’, though they are reconstructed somewhat differently: PWC *\*bʷV* ‘smoke’ in NCED (with a voiced labialized uvular fricative), but by Chirikba (with a voiced labialized pharyngeal fricative) PWC *\*ʕʷa* (= *\*ʕʷa*) ‘smoke’ (Č 393, 405). The recently extinct Ubykh had the simplex *bʷa* ‘smoke’, with compound words in other WC languages (as cited in NCED) Kabardian *ʔʷə-bʷa* ‘smoke’, Adyge *ʔʷə-bʷ*, Abkhaz *á-l-ʕʷa*, and Abaza *l-ʕʷa* ‘smoke’.

## ‘Fire’

The universal Basque word for ‘fire’ is conventionally written *su* (phonetic /s̺u/, with an apico-alveolar sibilant similar to Castilian *s*, denoted here as /s̺/). There is very little phonetic variation of this word in Bsq dialects, practically only the fronted vowel in Z *sü* /s̺y ~ s̺ø/, and an isolated case of diminutive *xu* /šy/ in the Amikuse dialect of BN (OEHS *su*). However, the early Basque dictionary by Landucci (1958 [1562]) bears witness to an interesting archaic variant (probably Araban) (iʃçu) = *itsu* ‘fire’, in (iʃçu arri) ‘flint’, lit. ‘fire-stone’ = *suharri*, *suarri*, or *sutarri* in other Bsq dialects (FHV 157; EDB 342). Thus in this model Bsq ‘fire’ is reconstructed as *\*s̺u* / *\*i=éu*, with an underlying morphological alternation analogous to that of Bsq *\*kē* / *\*e=kē* ‘smoke’ (see above).

In North Caucasian there is a word for ‘fire’ which is equally as universal as *\*s̺u* is in Basque. The form that most closely resembles Bsq *\*s̺u* / *\*i=éu* is Lak *ııy* ‘fire’, usually transcribed as *çu* by Russian Caucasologists, with an initial glottalized hissing affricate. Cognate words are used in all EC subgroups: Nakh, Avar, Andian, Tsezian, Dargi, Lezgian, and Khinalug. In West Caucasian the word has a nasal prefix *\*ma-* ‘‘of somewhat obscure origin’’ (NCED 354): Abkhaz *á-mca*, etc. (see Table 2 for more details on attest-

---

in principle, methodologically objectionable when the task is to demonstrate genetic relationship, for which only protoforms (as deep as possible) must be compared with each other. However, in the case of the words for ‘smoke’, for example, the juxtaposition of Bsq *\*(e=kē)* ‘smoke’ with PNC *\*k̥wĩnhV* ‘smoke’ is far from a transparent match. The plausibility of the match is enhanced when it can be shown that an array of NC forms, from Udi *k:uin* to Karata *k:ũji* to Chadakolob *k̥ʷi* ‘smoke’, may convergently replicate the historical stages leading from a form like *\*k̥wĩnhV* to the earliest known Bsq form *\*(e=kē)* ‘smoke’.

<sup>8</sup> Some Lezgian words, Lezgi *gum*, Tabasaran, Agul *kum* ‘smoke’, etc., look invitingly like they belong with Udi *k:uin*, etc., but because of phonetic developments (regular development of Proto-Lezgian *\*ʒʷ* > velars) they are derived instead from a PEC root with a lateral affricate, *\*Hʒwĩmā* ‘wind, air’ (NCED 590).

ed forms). An archaic form of ‘fire’ without the *\*ma-* prefix is found in Abkhaz *a-čmsá* ‘tinder’, Abaza *čwəmsa* id. < PWC *\*cʷə-məsa*, a compound with parallels in some Lezgian languages: Lezgi *çimaç*, Rutul *çimäs / çümäs*, Tsakhur *çimis* ‘tinder’ < PL *\*çoji-mäs* (NCED 800). On the basis of the above the PNC direct base is reconstructed as *\*çäj* ‘fire’, with oblique bases *\*çěj-* and *\*çaj-rV-* (NCED 354).

The comparison of Bsq *\*šu* ‘fire’ and Lak *çu* ‘fire’ was made more than a century ago by Trombetti (1903, §100), reiterated later in his 1925 book, and also, with other NC words, by Uhlenbeck and Bouda.

Another aspect of this etymology involves multiple noun stems (allomorphs) in East Caucasian, and their implications for West Caucasian and even Basque. “In most Daghestanian languages, nouns are characterized by multi-stem inflection. ... Besides the base stem, we need to distinguish an oblique (singular and plural) and a plural stem. The case-endings follow the oblique stems” (Berg 2005: 159). As regards words for ‘fire’ I would especially mention two of these stems. An archaic oblique stem for ‘fire’ is reconstructed as PNC *\*çěj-* on the basis of the Lak direct base *çu* ‘fire’, Lezgian oblique stem *\*çoji-* (e.g., Lezgi *çu-*) and the labialized forms in West Caucasian *\*mA-çʷa* (see the Adyge and Kabardian forms above) and PWC *\*cʷə-məsa* ‘tinder’. In a significant number of cases the archaic oblique base became the direct (nominative) base, as we have seen in Lak *çu* ‘fire’, and it seems that the Bsq word *\*šu* / *\*i=çu* ‘fire’ has an analogous origin.<sup>9</sup> Other examples of this kind of redistribution of allomorphs include the direct forms in Kryz *χʷar* ‘dog’, Budukh *χor* ‘dog’, and Khinalug *pχra* ‘dog’, all from the PEC oblique stem *\*χHwěj-rV-* ‘dog’ (NCED 1073); a cognate oblique stem seems to be reflected in Bsq *\*hor* ‘dog’.

Besides an ablaut alternation, as in PNC direct *\*çäj* / oblique *\*çěj-*, oblique stems may be formed by the addition of a consonantal marker, such as *\*-r-* in the words for ‘dog’ just mentioned. In the case of ‘fire’ “many [EC] languages reflect an old oblique base formed with the *\*-rV* suffix (*\*çaj-rV-*), cf. [Proto-Nakh] *\*çari-*, [Proto-Andian] *\*çari-* (reflected in Akhv. *çari*), Lak. *\*çara-*, [PL] *\*ça(j)-rV-*” (NCED 354).

In some earlier articles (e.g. Bengtson 2008: 89–90) I have proposed that some otherwise mysterious noun stem variants in Basque may be analyzed as preserving relics of archaic oblique stem markers, for example Bsq *\*-t-* corresponds to the PEC oblique stem marker *\*-d-*, and Bsq *\*-r-* corresponds to the PEC oblique stem marker *\*-r-* and/or *\*-l-*.<sup>10</sup> The Bsq *-a-* in words ending in a vowel could correspond to the PEC oblique stem marker *\*-yV-*, attested in Tsezian and Lezgian.<sup>11</sup> For example:

- Bsq *\*šu* ‘fire’ + *\*argi* ‘light’ > (B, BN, Sal) *su-t-argi* ‘firelight’
- Bsq *\*egun* ‘day’ + *\*aldi* ‘time’ > (c) *egu-r-aldi* ‘weather’

<sup>9</sup> I am thankful to G. Starostin for this observation.

<sup>10</sup> Bsq has two rhotic phonemes, flapped *r* [r] and trilled *rr* [r̄]; the latter is realized only in intervocalic positions, e.g. *lur* ‘earth’ but *lurra* ‘the earth’. Bsq trilled /t/ always corresponds to PNC *\*r*, while the flapped *r* usually corresponds to PNC *\*l* or *\*l̄*, except in a few cases where only Bsq and Yeniseian distinguish a “hard” PSC *\*r* from a soft or palatal PSC *\*r̄* (Bengtson & G. Starostin 2015: 8). The linking *r* in Bsq words like *egu-r-aldi* ‘weather’ is flapped, and could come from PSC *\*l*, *\*l̄*, or *\*r̄*.

<sup>11</sup> Hunzib, for example, also has the stem marker *-a-* (see examples).

- Bsq *\*hoin* ‘foot’ + *\*bide* ‘path’ > (B-arc) *o-r-bide* ‘(foot-)path’
- Bsq *\*e=će* ‘house’ + *\*-ko* [diminutive] > (R, Sal) *etx-a-ko* ‘outbuilding, pavilion’

Compare, for example, these examples of productive oblique stem markers in the EC language Hunzib (Berg 1995: 37–38):

- Hunzib *ože* ‘boy’: genitive *ož-du-s*
- Hunzib *koč* ‘bush’: gen. *koč-li-s*
- Hunzib *ko* ‘squirrel’: gen. *ko-ro-s*
- Hunzib *maru* ‘nose’: gen. *mar-a-s*

After the noun-class system lost its grammatical content (possibly rather early in the development of pre-Basque) the former class prefixes and oblique stem markers were lexicalized, and the stem markers were reorganized on a phonetic basis, so that *\*-t-* came to be preferred in nouns containing a plosive (usually *d* or *g*) or even *h* (e.g. BN-Aldude *bethalde* ‘herd of cows’ < *\*behi* ‘cow’ + *\*alte* ‘side’), and *\*-r-* became associated with nouns containing final *n*.

In this light it is interesting to consider a relic Bizkaian and High Navarrese form: *surtopil* ‘bread baked in embers’, more or less semantically equivalent to other forms like (B, Bzt, R) *sutopil* and *sukopil* ‘subcinericius panis’ (FHV 247-8). Michelena analyzes the latter as *\*suko opil* ‘torta del fuego’ (our *\*šu-ko \*og-t-bil*).<sup>12</sup> I would tentatively analyze *surtopil* as *\*šu-r-t-og-t-bil*, with two fossilized oblique stem markers, *\*-r-* and *\*-t-*, after *\*šu* ‘fire’. This kind of double marking is not unprecedented, and it also appears, for example, in Bsq (EB) *orkoi* ‘form, mold, shoe block’ (AW) = (Bzt, L) *orkoi* ‘shoe block, last’, (L, Z) *orkhei* id. (FHV 309; EDB 306–307). This word was recorded in a more archaic form, *hortkhei*, by Axular (17<sup>th</sup> c.), and this can be analyzed as *\*ho-r-t-gai* (*\*hoin* ‘foot’ + *\*gai* ‘thing, material’, with progressive umlaut in L and Bzt); cf. also (Z) *orpheko* ‘pedal’ < *\*ho-r-t-behe-ko* (*\*behe* ‘lower, below’, *\*-ko* [noun-forming suffix]).

Even more important is evidence that fossilized oblique stem markers show up in some relics of Basque noun inflection. Azkue (1905–1906: 969) cited Bsq “**SUR** : ... Var. de su, « fuego, » en la declinación de ciertos casos y en algún compuesto ...” The specific case forms cited by Azkue are *surtan*, *surten* ‘in the fire’ [locative] and *surtara* ‘to the fire’ [allative]. In the current model we can analyze this *sur* as *\*šu-r-*, historically a noun stem + (now fossilized) oblique stem marker. We are fortunate to have some notes about the oblique forms *\*šu-t-* and *\*šu-r-t-* by Michelena (see OEH: su ‘fuego’).<sup>13</sup>

In its main inessive [locative] use ..., the forms most used in the South (also by some High-Navarrese authors) are [locative] *sutan*, [ablative] *sutatik*, [allative] *sutara*, etc.; in the North these appear (alternating with *suán*, *sutik*, etc.) in Etcheberri Ciboure [17<sup>th</sup> c.], Pouvreau [17<sup>th</sup> c.], Gasteluzar [17<sup>th</sup> c.], Baratiart [18<sup>th</sup> c.] and Monho [18<sup>th</sup> c.]. *Surta-*, less frequent, is encountered in Gipuzkoan and Bizkaian authors, also in some northern ones such as Voltaire [17<sup>th</sup> c.], Etcheberri Ciboure [17<sup>th</sup> c.], or Mirande [20<sup>th</sup> c.]. The northern and High-Navarrese authors (including Mendiburu [18<sup>th</sup> c.]

<sup>12</sup> With the Bsq relational suffix *\*-ko* (HB 90, et passim). Bsq (EB) *opil* ‘bun, roll’, (Z) *óphil* ‘gâteau’ can be analyzed as *\*og-t-bil* (*\*ogi* ‘bread’ + *\*-t-* [fossilized stem marker] + *\*bil* ‘round thing’).

<sup>13</sup> The author’s translation. Michelena’s references to specific works have been removed and replaced by indications of the authors’ time span, e.g. 17<sup>th</sup> c., to emphasize the archaic nature of some of the forms.

generally used *suan*, *sutik*, *sura(t)*, etc., forms sometimes used by western writers such as Añibarro [17<sup>th</sup> – 18<sup>th</sup> c.], Iturriaga [17<sup>th</sup> – 18<sup>th</sup> c.], [etc.].

In other words, for the model proposed here, the postulated oblique stems *\*śu-t-*, *\*śu-r-*, and (double-oblique) *\*śu-r-t-*, alternating with the base stem *\*śu-*, have somewhat different dialectal distributions in case forms and in compound nouns. Bsq *\*śu-r-* ‘fire’ [oblique stem] agrees well with PEC *\*çaj-rV-* ‘fire’ [oblique stem] (see above); the Bsq vowel /u/ here is consistent with the PNC stem *\*çij-* (with ablaut), as mentioned above.

This is the clearest evidence that Basque not only has relic oblique stem markers in compound words, but that there are also some traces of oblique stem markers in the declension of Basque nouns that are parallel to EC declensional forms. Trombetti (1925: 73–75) discussed the morph *-t-* in Bsq noun and pronoun declension, pointing out forms like Bsq *mendi-e-ta-n* ‘in montibus’ [locative plural: ‘in the mountains’]; *ni-ta-z* ‘(done) by me’ [instrumental singular]; and in composition he cited the Bsq form *su-t-opil* ‘pane cotto nella cenere’ (see above). Trombetti did not connect this *-t(a)-* with any NC forms, but in the same book Trombetti (1925: 76–79) noticed the parallel between Basque *-r-* and EC *-r-*, specifically mentioning forms like Batsbi *pñu* / [oblique] *pñar-* ‘dog’, *çe* / [oblique] *çar-* ‘fire’, in connection with the Bsq *-r-* that appears in singular case forms like *mendi-a-r-i* ‘to the mountain’ [dative], *mendi-a-r-en* ‘of the mountain’ [genitive], and also in compound nouns like *egu-n* ‘day’ / *egu-r-aldi* ‘weather’; as far as I know he did not address the topic of Bsq *\*śu-r-* ‘fire’. Indeed, Bsq *\*śu-r-* is never found by itself, but only in the compounded form *\*śu-r-t-*. Possibly the original Bsq case paradigm of *\*śu* ‘fire’ included forms with the *\*-r-* morph and others with the *\*-t-* morph, and eventually, as the grammatical meaning of the oblique stem markers was lost, the compound morpheme *\*-r-t-* arose in some (mainly western) Bsq dialects as a compromise or blend-form.

### Other words pertaining to ‘fire’:

- Basque *\*i=łhinti* ‘firebrand, ember’ (EB *ilinti*, NBsq *ilhinti*): putative East Caucasian cognates mean ‘firewood’: PEC *\*łwindV* > Andi, Godoberi *łudi* ‘firewood, wood’; underlying *\*-nd-* is reflected as nasal vowels in Tsezian: Bezhta *hũdo*, Hunzib *hũdu* id. (NCED 764). See Table 3 below for more details.<sup>14</sup>
- Basque *\*sendo-ř* ‘load of firewood, fagot’ (EB *zendor* ‘wooden pyre for making charcoal’ [AW]): cf. PEC *\*zindV* ‘fuel’ > Avar *ç:ada-ri* ‘fuel’, etc. (NCED 366). See Table 4 for more details.
- Basque *\*hardo* ‘tinder’ (EB *ardai* < *\*harda-gai*): putative EC cognates denote various herbs: PEC *\*HwVrdV* ‘a kind of herb’ > Dargi *ħorda* ‘henbane’, etc. (NCED 623). See Table 5 for more details.
- Basque *\*hauć* ‘ashes, dust, powder’ (EB *hauts*): cf. PEC *\*jōmćV* ‘earth’ (NCED 684). See Table 6 for more details.
- Basque *\*keda-ř* ‘soot’ (EB *gedar*): cf. Akhwakh *q:ęta* ‘soot’, etc. < PEC *\*qđidV* ‘soot, dust’ (NCED 927). See Table 7 for more details.

<sup>14</sup> Note that *\*ł* is NCED’s notation of voiceless lateral fricative, phonetically identical with the symbol [ɬ] commonly used in transcriptions of Na-Dene, sometimes also [ɬ] or [ɬ].

## Discussion

Although there is ample evidence that the separate development of Basque in Western Europe and North Caucasian languages in the Caucasus has led to extensive lexical differences (e.g. numerals and kinship terms), the lexical family revolving around fire and making fire has remained relatively stable over nine or ten millennia.<sup>15</sup> This is not surprising, since regardless of geographical migrations, influences from neighboring languages, and social and cultural upheavals, the basic need to make and use fire, and methods of kindling fire (until the recent adoption of matches), have changed very little over tens of millennia. This is in accord with the 50-item word list of the most stable lexical meanings, in which ‘fire’ is #7 in stability, ‘smoke’ is #36, and ‘ashes’ is #38. (G. Starostin 2010: 112–113).

Along with the etyma themselves, evidence is presented here that Basque preserves traces of the old Euskaro-Caucasian morphological system. These include fossilized class prefixes, some of them historically separable, as in the Basque variants *\*kē* / *\*e=kē* ‘smoke’ and *\*śu* / *\*i=ću*, ‘fire’, as discussed above. Especially in the case of ‘fire’ we find traces of fossilized oblique stem markers, namely those found in Bsq forms containing *\*śu-t-* and (double-oblique) *\*śu-r-t-*, the latter of which has direct parallels in, for example, Lak *çu* ‘fire’, [oblique] *ça-ra-*; Ingush *çi* ‘fire’ / *çe-r-a* [genitive], *çe-r-aga* [allative]; Archi *oç* ‘fire’ / *çe-re-q* [locative], etc. In Akhwakh *çari* ‘fire’ (*ç-a-ri*) and Udi *aruχ* ‘fire’ (*a-r-uχ*: see Table 2) the oblique stem has been incorporated into the direct form (analogous to the formation of Bsq *\*hor* ‘dog’ ~ Budukh *χor* ‘dog’: see above).

Thus, “where there is fire, there is smoke” can also serve as a metaphor for the observation that where there is a genetic linguistic relationship (“fire”), there are also lexical, phonological, and grammatical features (or their traces = “smoke”) that verify the existence of the former.

This type of study, emphasizing linguistic structures and their relationships to social, anthropological and historical structures can be traced back to influences from the Prague School (see Appendix D).

**Acknowledgments:** I am grateful for helpful comments from Václav Blažek, James Harrod, Peter Norquest, and Juho Pystynen (thanks to a discussion session on Academia.edu); and from anonymous reviewers for *Slovo a slovesnost*.

---

<sup>15</sup> “The split of the Basque–NCauc. proto-language into the Basque and NCauc. branches glottochronologically occurred in the first half of the 7<sup>th</sup> millennium BC” (Kassian 2010: 425, citing G. Starostin, using calculations based on the 50 most stable basic lexemes). See also the tree on TOB: <<http://starling.rinet.ru/images/globet.png>>.



APPENDICES

Appendix A: Etymologies discussed

Table 1: Euskaro-Caucasian ‘smoke’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
smoke	*(e=)kē	*k̄wĩnhV ‘smoke’ NCED 738	Bsq (B, G, R-Vidángoz) <i>kee</i> ‘smoke’, <i>kei-</i> (before vowels, as <i>keia</i> ‘the smoke’), (B, G, AN) <i>ke</i> ‘smoke’, (G, AN) <i>ki</i> , (BN, L, Z) <i>khe</i> ‘smoke’; a few northeastern dialects (AN-Ulzama, BN-Baigorri, Azk, Bzt, Sal, and general R) have the prefixed form: <i>eke</i> ‘smoke’, (L) <i>ike</i> (Azkue 478; FHV 67, 156; EDB 248, 250; OEH). A related word seems to be <i>*kino</i> ‘stench, odor’ > (BN) <i>kino</i> , (L) <i>khino</i> , (Z) <i>khĩño</i> , (R) <i>kĩo</i> (but not related to Bsq <i>*kelu</i> ). <sup>16</sup> In this analysis Bsq <i>keru</i> < <i>*kelu</i> ‘stench, bad taste, grudge’ (attested in AN-Irun as <i>kelu</i> ) is related to Bsq <i>*ke[l]-ac</i> ‘bitter, sour, stench’, <i>*kel-de-r̄</i> ‘bile, gall’ and to PNC <i>*q̄ēhIV</i> ‘bitter; salty; sweet’ (NCED 912). Bsq <i>*kel-de-r̄</i> ‘bile’ is not related to <i>*keda-r̄</i> ‘soot’ except by lexical blending (contamination).	<i>Avar-Andian</i> : Avar <i>κ:uj</i> ‘smoke’, Chadakolob <i>κ̄<sup>wi</sup></i> , Andi <i>κ:woj</i> , Akhwakh <i>κ:ṽṽ</i> , Tindi <i>κ:ṽṽ</i> , Karata <i>κ:ũji</i> , Bagwali <i>κ:ũj</i> id. <i>Tsezian</i> : Bezhta, Hunzib <i>qo</i> ‘smoke’. <i>Lak</i> : Lak <i>kuw</i> ‘soot’, Khosrekh <i>κ<sup>w</sup>a</i> id. <i>Dargi</i> : Dargi Akusha, Urakhi <i>gaw</i> , Gapshima <i>κ:aw</i> ‘smoke’. <i>Lezgian</i> : Udi <i>κ:uin</i> ‘smoke’, Archi <i>gun</i> ‘dust’. <i>West Caucasian</i> : Abkhaz <i>â-l-ṽ<sup>w</sup>a</i> , Abaza <i>l-ṽ<sup>w</sup>a</i> ; Adyge <i>ṽ<sup>w</sup>ə-Ḃ<sup>w</sup></i> , Kabardian <i>ṽ<sup>w</sup>ə-Ḃ<sup>w</sup>a</i> ; Ubykh <i>Ḃ<sup>w</sup>a</i> ‘smoke’.

Table 2: Euskaro-Caucasian ‘fire’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
fire	*śu ~ *i=ću	*ć:ājī / *ć:ūj- ‘fire’ <sup>17</sup> NCED 354	Bsq (c) <i>su</i> ‘fire’, (Z) <i>sū</i> ‘fire’. Arabian had the exceptional prefixed form <i>itsu</i> ‘fire’ (in <i>itsu arri</i> ‘flint’ = <i>*i=ću hari</i> ‘fire-stone’). In other Bsq dialects words for ‘flint’ reflect <i>*śu-hari</i> , <i>*śu-t-hari</i> (with a fossilized oblique stem marker <i>*-t-</i> ), or <i>sukarri</i> (< <i>*śu-ko-hari</i> (FHV 157; EDB 342; OEH). §. In oblique case forms some Bsq writers (and speakers) have employed stems reflecting fossilized stem markers <i>*śu-t-</i> or <i>*śu-r-t-</i> (with double marking). In this model Bsq <i>*śu-r-</i> is deemed cognate with the PL oblique stem <i>*ć:o(j)-rV-</i> (see the main text). This is one of the most straightforward and plausible Basque-North Caucasian etymologies. Trombetti (1903 §100), Orig 306; U 581: Bsq + Lak, etc.	<i>Nakh</i> : Chechen, Batsbi <i>će</i> ‘fire’, Ingush <i>ć:i</i> id. <i>Avar-Andian</i> : Avar <i>ć:a</i> ‘fire’, Chadakolob <i>ća</i> ; Andi <i>ć:a</i> , Karata, Botlikh, Bagwali <i>ć:aj</i> , Godoberi <i>ć:aji</i> , Akhwakh <i>ćari</i> , Chamali, Tindi <i>ć:ā</i> id. <i>Tsezian</i> : Tsezi <i>ć:i</i> ‘fire’, Inkhokwari, Bezhta, <i>ć:o</i> , Khwarshi <i>ć:a</i> , Hinukh <i>će</i> Hunzib <i>ć:ə</i> id. <i>Lak</i> : Lak <i>ć:u</i> ‘fire’. <i>Dargi</i> : Dargi <i>ć:a</i> ‘fire’. <i>Lezgian</i> : Lezgi <i>ć:aj</i> (ergative <i>ć:u</i> ) ‘fire’, Rutul <i>ć:aj</i> , Tabasaran, Agul, Tsakhur <i>ć:a</i> , Kryz <i>ć:ā</i> , Budukh <i>ć:ə</i> , Archi <i>oć:</i> , Udi <i>a-r-uḡ</i> id. <sup>18</sup> <i>Khinalug</i> <i>ćā</i> ‘fire’. <i>West Caucasian</i> : (with <i>*m</i> =prefix) Abkhaz <i>â-mça</i> , Adyge <i>māśā</i> , Ubykh <i>māśā</i> ‘fire’.

**Table 3:** Euskaro-Caucasian ‘firebrand ~ firewood’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
firebrand, ember	*i=θinti	*λwindV [hwindV] ‘firewood, wood’ <sup>19</sup> NCED 764	Bsq (Z) <i>ilhinti</i> ‘firebrand, ember’, (L) <i>ilhindi</i> , <i>ilhendi</i> , (B, R-Uztárroz) <i>ilinti</i> , (R-Vidángoz) <i>illindi</i> , (B) <i>ilindi</i> , <i>illentí</i> , (G) <i>illanti</i> , (B, G, AN, L) <i>illeti</i> , (Azk) <i>eileti</i> , (B-Bergara, Leintza, G-Tolosa, Betერი) <i>illati</i> , (B-Gernika) <i>illunti</i> , (Sal) <i>illendi</i> id.; (with transposed /h/) (Z) <i>hilindi</i> (Oihenart, 17 <sup>th</sup> c.); (with expressive /ç/): (BN-Amikuse) <i>itxendi</i> , (BN, L) <i>itxindi</i> , (BN-Baigorri) <i>itxindu</i> id. Cf. (AN-Elkano, Bzt, Sal, L) <i>ilindi</i> ‘dead ear of wheat’ (< ‘dry thing’?) (FHV 188, 190; MDE XV 910; EDB 223; OEH: ILINTI; ITXINDI).	<i>Avar-Andian</i> : Andi, Godoberi <i>hudi</i> ‘firewood, wood’, Akhwakh, Karata <i>huda</i> , Chamali <i>hunni</i> , Tindi <i>tuji</i> , Botlikh <i>hudi</i> , Bagwali <i>hur</i> id. <i>Tsezian</i> : Tsezi <i>redu</i> ‘firewood, wood’, Hinukh <i>rede</i> , Khwarshi <i>lida</i> , Inkhokwari <i>lido</i> , Bezhta <i>hūdo</i> , Hunzib <i>hūdu</i> id. <i>Lak</i> : Lak (Khosrekh) <i>x:undu-ti</i> ‘a stick for working leather’. <sup>20</sup>

**Table 4:** Euskaro-Caucasian ‘firewood’, ‘fuel’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
load of firewood	*sendo-ř	*zindV ‘fuel’ <sup>21</sup> NCED 366	Basque (G) <i>zendor</i> ‘load of firewood (for making charcoal)’, (B) <i>sendor</i> ‘load of wood’. §. Note that <i>s</i> [ʃ] and <i>z</i> [s] have merged in B (HB 138). OEH cites other related meanings: ‘bundle, bunch, bouquet, handful (of wood, wheat)’, etc. There are some northern Bsq variants with <i>-ld-</i> rather than <i>-nd-</i> : (BN, L) <i>seldor</i> ~ <i>zeldor</i> (OEH: SENDOR); by dissimilation, or confusion with a different etymon?	<i>Avar-Andian</i> : Avar <i>ç:adā-ri</i> ‘fuel’. <i>Lak</i> : Lak <i>çinna</i> ‘fuel (of dry stalks)’, Lak Khosrekh <i>çajnda</i> ‘straw’. §. A recessive etymon, with attested reflexes only in Lak and Avar.

<sup>16</sup> Azkue gives only the meaning ‘bad taste’ (‘mal gusto / mauvais goût’); FHV (67) has ‘hedor’ (> ‘mal sabor’); Larrasquet (1939) has lower Zuberoan as KHINO [khiño] n. *mauvais goût* (et non par *mauvaise odeur*); ‘Olor, gusto; mal olor, mal gusto’ (OEH: KINO).

<sup>17</sup> \*çij- is the reconstructed PNC oblique base, attested in PL \*çoji- (Lezgi *çu-*, etc.), Lak *çu*, and PWC \**ma-ç<sup>va</sup>*.

<sup>18</sup> Udi *a-* (?*a-*) is the regular reflex of PL \*çaj ‘fire’, followed by an oblique stem marker *-r-* and a plural affix *-uç*.

<sup>19</sup> NCED cites alternative reconstructions \*λwindV or \*λwāndV: for the latter cf. the Bsq dialectal forms *illanti*, *illati*.

<sup>20</sup> “A perfect phonetic match is Lak. (Khosr.) *x:undu-ti* ‘a stick for working leather’, but the meaning raises some doubts” (NCED). However ‘stick’ would be a natural development of ‘wood’.

<sup>21</sup> Because of the scarce attestation this reconstruction is tentative. The first consonant could also be tense \*ç, and the first vowel could also be \*ř or \*ā.

**Table 5:** Euskaro-Caucasian ‘tinder ~ herb’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
<b>tinder</b>	*hardo	*HwVrdV ‘a kind of herb’ NCED 623	Basque (L) <i>hardo</i> ‘tinder (made from amadou fungus)’, (with expressive palatal) (BN-Ustaritz, Mendionde) <i>harddo</i> [har̥do], (B-Bergara, Leintza, G-Azpeitia, Goierri) <i>arda-gai</i> , (G-Beterri) <i>arta-gai</i> , (AN-Larraun) <i>ardan-gai</i> , (BN, Z) <i>ardai</i> , (B-Arratia) <i>erdai</i> id. (some of the words are compounded with *gai ‘thing, material’). (FHV 220, 250; EDB 103; OEH). <sup>22</sup>	<i>Avar-Andian:</i> Akhwakh <i>ʃidi-ra</i> ‘young grass’. <i>Lak-Dargi:</i> Lak <i>urt:u</i> ‘grass’; Dargi Akusha <i>ħorda</i> ‘henbane’. §. A scantily attested Dagestanian etymon.

**Table 6:** Euskaro-Caucasian ‘ashes ~ dust ~ earth’<sup>23</sup>

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
<b>ashes, powder</b>	*hauć	*jōmćV ‘earth’ NCED 684	Basque (L) <i>hauts</i> , <i>auster</i> ‘ashes’, <i>herr-auts</i> ‘powder’, (Z) <i>hauts</i> , <i>hautx</i> ‘ashes’, <i>er-háuts</i> , <i>er-háutx</i> ‘powder’, (R) <i>auts</i> ‘ashes’, <i>err-áuts</i> ‘powder’, (B) <i>auts</i> ‘ashes, powder’, <i>err-auts</i> ‘ashes’, (G) <i>auts</i> ‘ashes, powder’, <i>auster</i> , <i>austerre</i> ‘ashes’, etc. (FHV 93; EDB 119; OEH). <sup>24</sup>	<i>Nakh:</i> Batsbi <i>jobst</i> ‘earth’, Ingush <i>jost</i> ‘loose earth’. <i>Dargi:</i> Dargi Akusha <i>ʃanži</i> ‘earth’, Kubachi <i>ʃanć:a</i> id., Chiragh <i>ʃanć:i-ya</i> , Urakhi <i>ʃanži</i> , Tsudakhar <i>ʃanć:i</i> , <i>ʃanć:a</i> id. <i>Lezgian:</i> Udi <i>oć:i</i> ‘dirt, mud’, <i>oć:a-l</i> ‘ground, land’. <sup>25</sup> <i>Khinalug</i> <i>inć:i</i> ‘earth’. §. A Nakh-Eastern Dagestanian isogloss.

<sup>22</sup> This word has secondarily been confused and blended with Bsq *gardu* ~ *kardu* ‘thistle’, which is distinct and a loanword from Romance (Latin *carduus*). See the additional notes in Appendix B.

<sup>23</sup> For semantic typology of ‘ashes’ ~ ‘dust’ ~ ‘earth’ cf. Gk κότις ‘dust’ ~ Lat *cinis* ‘ashes’; Goth *mulda* ‘dust’ ~ OE *molde* ‘loose earth, soil’, etc. (Buck 1949: 1.213).

<sup>24</sup> Some of the Bsq words are apparently compounds with \**eře* ‘to burn’. In an earlier analysis I proposed an element \*-*teře*, but it now seems that forms such as *austerre* should be analyzed as \**hauś-t-eře*, with the fossilized oblique marker \*-*t-* and verb \**eře* ‘to burn’ or ‘burnt’.

<sup>25</sup> The two Udi forms have different affricates (*oć:al* ‘ground, land’ vs. *oć:i* ‘dirt, mud’) and may be unrelated, the semantic shift ‘dirt’ > ‘ground, land’ is also unlikely. Lezgian words for ‘earth’ have been extensively treated in Kassian (2011–2012).

**Table 7:** Euskaro-Caucasian ‘soot’

General gloss (Basque)	Proto-Basque	Proto-North Caucasian	Basque attested forms and notes	North Caucasian attested forms and notes
soot	*keda-ř	*q̇idV ‘soot, dust’ NCED 927	Basque (B, G, Bzt, AN-Esteribar, BN-Aldude) <i>keđar</i> ‘soot’, (B) <i>keđarra</i> , <i>geđarra</i> , (G) <i>kear</i> , <i>kelar</i> ; (BN-Hazparren) <i>keđer</i> , <i>kerrada</i> , (L) <i>kheder</i> , (L-Ainhua, Senpere, BN-Mugerre) <i>kelder</i> , (Z) <i>khedárre</i> , <i>kherrade</i> , (R) <i>geđar</i> , (Sal) <i>geđarre</i> , etc. §. Forms with <i>g-</i> by voicing assimilation. Any association of Bsq * <i>keda-ř</i> (PEC * <i>q̇idV</i> ) with * <i>kē</i> ‘smoke’ (PNC * <i>kw̃nh̃V</i> ) is secondary, since the original PEuC forms were quite distinct ( <i>contra</i> FHV 255, EDB 249; cf. MDE XIII 339; OEH).	<i>Avar-Andian</i> : Akhwakh <i>q̇:ēřa</i> ‘soot’, Tsegob <i>q̇:eře-n</i> , Ratlub <i>q̇:ařa-ni</i> , Karata <i>q̇:ēřa-n</i> , Tokita <i>q̇:iřa-n</i> ‘soot’. <i>Tsezian</i> : Hunzib <i>řiřo</i> ‘mould’. <i>Lak</i> : Lak <i>q̇it</i> ‘soot; flour dust’. §. A recessive etymon, attested in only four Dagestanian languages.

## Appendix B: Phonological and semantic notes

Phonetic correspondences between Basque and PNC and other phonetic and semantic changes are explained as follows. The order of presentation is the same as for the etymologies in Appendix A.

It should be borne in mind that Euskaro-Caucasian phonology is still in its relative infancy. Apart from some observations by earlier scholars like Bouda and Lafon, a systematic inventory was begun by the author in the 1990s, and has continually been refined and expanded since then (e.g. Bengtson 1992, 1996, 2004, 2008, and most recently, 2017a). As can be seen below, some parts of the system are well established and relatively transparent (see 1.a, 2.a, 3.a, 4.a, etc.), while other issues will require more work in the future to untangle regular sound changes from the cumulative effects of assimilations, contaminations and other countervailing forces. Euskaro-Caucasian ablaut of noun stems (see 1.b, 2.b, 4.b, 7.b) is a recent discovery (within the past two years), and is only beginning to be understood (Bengtson 2017a: 106–110).

**Table 1:** Basque \*(e=)kē ‘smoke’ ~ PNC \**kw̃nh̃V* ‘smoke’:

**1.a.** The initial correspondence, Bsq \**k* = PNC \**k̃*, is supported by a significant number of etymologies:

- Bsq \**kain* ‘fog, clouds’ ~ PNC \**kw̃nh̃V* ‘cloud, fog’ (NCED 737)
- Bsq \**kako* (~ \**gako*) ‘hook, point, key’ ~ PNC \**kw̃ăkē* ‘edge, point, corner’ (NCED 733)
- Bsq \**kankano* ‘large fruitstone, kernel, almond’ ~ PNC \**k̃řřk̃ř(-nV)* ‘small stone, grain, egg’ (NCED 730)
- Bsq \**koce* ‘nape’ < PEuC \**k̃ōce* or \**k̃ōce* ~ PEC \**k̃ōce* ‘a kind of bone’ (NCED 698)
- Bsq \**koipe* (~ \**goipe*) ‘rendered fat, grease’ ~ PEC \**k̃āmpē* ‘butter; sour cream’ (NCED 716)
- Bsq \**kokoř* / \**kor-ka* ‘throat’ ~ PNC \**kw̃řřV* ‘craw, crop’ (NCED 740)
- Bsq \**koroc* (~ \**goroc*) ‘dung, manure’ ~ PEC \**k̃[u]řřV* / \**k̃[u]řřV* ‘dung, excrement’ (NCED 728)
- Bsq \**kuko* ‘bug, louse’, \**kuku-šo* ‘flea’ ~ PEC \**k̃ăkV* ‘a kind of insect’ (NCED 715)

In some of the Bsq words there is dialectal vacillation between the voiceless initial *\*k-* and the voiced initial *\*g-*. Some cases may be explained by assimilation or dissimilation, or by “voicing crossover”, a phenomenon observed in Austronesian and other language families.<sup>26</sup>

**1.b.** The correspondence of the stem vowels, Bsq *\*-ē-* = PNC *\*-i-*, may be explained by Proto-Euskaro-Caucasian *\*ə* / *\*i* **ablaut**. PEuC *\*ə* (> Bsq *\*e*) originally belonged to the direct noun stem, *\*i* (> Bsq *\*i*) to the oblique stem, and over time direct and oblique allomorphs became redistributed. In this and the following cases Bsq may have retained the *\*e* [direct] grade (regularly = PNC *\*ə*) vs. PNC [oblique] *\*i*:

Bsq *\*esko* < *\*e=Seko* ‘(bees-)wax’ ~ PNC *\*ĉHġwǎ* ‘dung, ordure, dirt’ (NCED 387)<sup>27</sup>

Bsq *\*hesu-r* ‘bone’ ~ PEC *\*HġšiwV(-tī)* ‘rib; side’ (NCED 584)

Bsq *\*lerde-n* ‘straight, slender’ < PEuC *\*lǎV-n* ~ PNC *\*=iǎiV* ‘thin’ (NCED 639)

In some other cases there may have been an opposite redistribution, with Bsq retaining the *\*i* [oblique] grade (regularly = PNC *\*i*) vs. PNC [direct] *\*ə*:

Bsq *\*biha-* ‘next day, tomorrow’ < PEuC *\*biká-* ~ PEC *\*bǎgə* / *\*bǎgo* ‘morning, evening’ (NCED 292)

Bsq *\*gibe-l* / *\*bigē-l* ‘liver, back side’ < PEuC *\*gibV-* / *\*bigV-* ~ PNC *\*gǎbǎ* / *\*bǎgwǎ* ‘side’ (NCED 292)

Bsq *\*higuni* ‘hate, hatred’ ~ PNC *\*=HǎχwAn* ‘to quarrel, fight’ (NCED 565)

Bsq *\*mġhu-ri* ‘seed, kernel’ ~ PEC *\*mǎhwV* ‘grain, seed’ (NCED 806)

**1.c.** Loss of an original internal nasal in medial *\*-NH-* clusters, as in Bsq *\*(e=)kē* ‘smoke’ ~ PNC *\*kŵĩnhV*, is also a recurrent phenomenon. Note that there have been convergent deletions of nasals in some NC reflexes, and in some an intermediate stage of nasalized vowel remains:

Bsq *\*a=tha-ba* ‘daughter’ ~ Udi *χġy-ār* ‘daughter, girl’, etc. < PEC *\*liñhV* ‘woman, female’ (NCED 762)

Bsq *\*kaio* ‘seagull var.’ ~ Akhwakh *q<sup>a</sup>a* ‘jackdaw, rook’, etc. < PEC *\*qǎmhǎ* / *\*qǎnhǎ* ‘raven, crow’ (NCED 889)

Bsq *\*šaie* ‘vulture’ ~ Tseji *ce(j)* ‘eagle, vulture’, Karata *ç:ũji* ‘eagle’, etc. < PEC *\*çwǎmǎV* (NCED 370)

Bsq *\*sihi* ‘wedge, skewer, spit’ ~ Bagwali *çi* ‘arrow’, Tindi *ĉihu* ‘arrow’ etc. < PNC *\*çǎñHV* ‘arrow, bow’ (NCED 358)

See also below (Table 6: Bsq *\*hauć* = PEC *\*jǎmćV*) for the frequent deletion of PEuC internal *\*-m-* in clusters.

**1.d.** There seems to be an etymological relationship between Bsq *\*(e=)kē* ‘smoke’ and Bsq *\*kino* ‘(bad) odor, (bad) taste’ that is parallel to that between Bsq *\*minhi* ‘tongue’ and *\*minco* ‘speech, voice’, i.e. an ‘active’ noun with a front final vowel vs. a ‘stative’ noun ending in *\*-o*. Cf. also Bsq *\*anci* / *\*anco* ‘likeness, manner, skill’, etc. ~ Andi *anži* ‘mark, sign’, etc. < PEC *\*hŵǎlǎzi* ‘mark, sign’ (NCED 534), though here the semantic opposition is not so clear-cut.<sup>28</sup>

<sup>26</sup> Blust (1996) has discussed a typologically similar situation in the Austronesian languages, in which “the reflex of *\*g* or *\*k* in one member of a cognate set exhibits a contrary value for voicing”. Blust calls this phenomenon “voicing crossover”.

<sup>27</sup> Budukh ‘ear-wax’, etc.

<sup>28</sup> Bsq (B) *antzi* ‘likeness, appearance, skill, knack, cunning’ vs. (Z, AN-Irún) *antzo* ‘proportion, manner, measure’.

**Table 2.** Basque \**śu* / \**i=ću* ‘fire’ ~ PNC \**čǎjř* / \**čǎj-* ‘fire’:

**2.a.** The match of Bsq apico-alveolar fricative \**ś* or the corresponding affricate \**ć* with PNC/PEC laminal affricate \**č* is recurrent, and seems to be associated with adjacent or proximal labial (rounded) vowels or consonants:<sup>29</sup>

- Bsq \**auśi-ki* ‘to bite’ ~ PEC \*=*VčV* ‘to drink; to gulp, to eat’ (NCED 1017)
- Bsq \**ośo* ‘whole, complete, entire’ / \**aśe* ‘(to be) full, sated’ ~ PNC \*=*hōčV* ‘full, to fill’ / \*=*h[a]čV* ‘to be filled, full’ (NCED 525)
- Bsq \**śaie* ‘vulture’ ~ PEC \**čwām?V* ‘eagle, vulture’ (NCED 370)
- Bsq \**śami-n* ‘bitter, sour, pungent’ ~ PEC \**čěmhV* / \**čwěnhV* ‘salt, salty’ (NCED 371)
- Bsq \**śapa-r* ‘thicket, bramble’ ~ PEC \**čibV* ‘plant with berries’ (NCED 367)
- Bsq \**śoin* ‘shoulder, (upper) back’, etc. ~ PEC \**Hčwěymǎ* ‘leg bone’ (NCED 555)
- Bsq \**a=čo* ‘old woman’ ~ PNC \**čwōyV* ‘woman, female’ (NCED 374)
- Bsq \**e=će* ‘house’ ~ PNC \**č[ř]yu* ‘house’ (NCED 364)
- Bsq \**huć* ‘empty, bare’, \**hućal* ‘poor, barren’ ~ PNC \*=*HřčĀl* / \*=*HřčV-* ‘naked, bare’ (NCED 567)
- Bsq (western) \**o=koć* ‘chin; snout; nape’ ~ PEC \**qǎčř* / \**qǎčřü* ‘bite; mouth’ (NCED 907)

**2.b.** Basque \**śu* (~ \**i=ću*) ‘fire’, with its rounded vowel, does not match the PNC direct form \**čǎjř* ‘fire’, but the postulated oblique base \**čǎj-*, with \**a* / \**u* ablaut, attested in Lak *ču* ‘fire’, the Lezgi oblique base *ču-* (< Proto-Lezgian \**čoji-*), and the labialized forms in West Caucasian, PWC \**mAč<sup>w</sup>a* ‘fire’ and Abaza *č<sup>w</sup>əmsa* ‘tinder’ < PWC \**c<sup>w</sup>ə-məsa*.

**Table 3:** Basque \**i=thinti* ‘firebrand, ember’ ~ PEC \**λwindV* [†windV] ‘firewood, wood’:

**3.a.** The match of medial Bsq \**-th-* with the PNC/PEC voiceless lateral fricative \**λ* is recurrent:<sup>30</sup>

- Bsq \**a=tha-ba* ‘daughter’ ~ Udi *χǎj-är* ‘daughter, girl’, etc. < PEC \**λinhV* ‘woman, female’ (NCED 762)
- Bsq \**a=thar-gun* ‘widower, widow’ ~ PEC \**λ(w)řr-(dV)* ‘woman, female’ (NCED 764)
- Bsq \**a=the* / \**a=tho-* ‘seed, grain’ ~ PEC \*=*VλwV-n* ‘to sow’, \**λwin?ř* ‘seed’ (NCED 1021)
- Bsq \**b=ithu-r* ‘binding, tether’ ~ PNC \*=*VλwE(n)* / \*=*VλwE(n)* ‘to bind, knit, plait’ (NCED 1021)
- Bsq \**e=thori* ‘thorn, spine, hawthorn’ ~ PNC \**λĀl(H)V* ‘arrow’ (NCED 758)
- Bsq \**e=thu-r* ‘snow’ ~ PEC \**λřwV* / \**jřwλV* < \**jř=λřwV* ‘snow’ (NCED 684)
- Bsq \**othe* ‘meek, mild’ ~ PEC \*=*uλwV* ‘to be silent, quiet’ (NCED 1011)
- Bsq \**otho* ‘oats, wild oats’ ~ PNC \**λwřłwV* ‘millet’ (NCED 763)

**3.b.** In several cases Bsq \**i* matches PNC \**wi* or \**wř*. It is so far uncertain whether this represents an innovation in PNC (insertion of \**w*) or metathesis of a class prefix \**w=*, or something else.<sup>31</sup>

- Bsq \**kino* ‘bad odor, bad taste’ (cf. \*(*e=*)*kē* ‘smoke’, above) ~ PNC \**kwinřV* ‘smoke’ (NCED 738)
- Bsq \**kiri-ki* ‘hedgehog’ ~ PEC \**qHwir-dV* ‘hedgehog’ (NCED 926)
- Bsq \**lincu-ra* ‘marsh, bog’, \**lisun* ‘mold, mildew; moldy, musty’ (< \**lincu-n*) ~ PEC \**λwilčwV* ‘bog, marsh’ (NCED 770)

<sup>29</sup> On the other hand, there are also examples of Bsq laminals \**s* and \**c* with PNC/PEC \**č*. These cases involve especially the unrounded vowels /a/ and /e/.

<sup>30</sup> Aspirated /lh/ is attested only in NBsq dialects (spoken and literary); /lh/ is absent from standard Bsq (EB), thus EB *ilinti*, *alaba*, etc. In absolute initial position the Bsq reflex is ordinary (unaspirated) \**l-* in all attested dialects, e.g. Bsq \**irdi* ‘drool, saliva, slime’ ~ PEC \**λwirdi* ‘pus; dung’ (NCED 763).

<sup>31</sup> There are also some cases in which Bsq \**u* matches PNC \**wi* / \**wř*.

Bsq *\*lindila* ‘lizard’ (in compounds) ~ PEC *\*λwitλwiV* ‘lizard’ (NCED 763)

Bsq *\*lirdi* ‘drool, saliva, slime’ ~ PEC *\*λwirdi* ‘pus; dung’ (NCED 763)

Bsq *\*sirsu-* / *\*sincu-r̄* ‘throat’ ~ PEC *\*čwĩrsV* ‘gullet’ (NCED 337)

**3.c.** As noted in the main text, the initial element *\*i=* is analyzed here as a fossilized class prefix, common in other Basque words denoting mass and uncountable nouns, and is thought to be cognate with the North Caucasian prefix *\*j=* / *\*i=* [II-class animate female]. The fossilized prefix *\*e=*, as in Bsq *\*e=kē* ‘smoke’, is apparently an allomorph of *\*i=* (see Appendix C).

**Table 4:** Basque *\*sendo-r̄* ‘load of firewood’ ~ PEC *\*zindV* ‘fuel’:

**4.a.** The match of initial consonants, Bsq *\*s* = PNC/PEC *\*ʒ* [d<sup>2</sup>] is well supported:

Bsq *\*sethai* ‘plain, prairie’, etc. ~ PEC *\*ʒəlV* ‘plain, plateau’ (NCED 1092)

Bsq *\*sesen* ‘bull’ ~ Proto-Avar-Andian *\*zin-HV* ‘cow’ < PNC *\*ʒVn-* (NCED 262)

Bsq *\*siki-ro* ‘gelded ram’, *\*siki-te* ‘gelded goat’ ~ PNC *\*ʒikV̄* / *\*kĩʒV̄* ‘goat, kid’ (NCED 1094)

Bsq *\*simi(n)c(a)* bedbug ~ PNC *\*ʒimiʒĀ* / *\*miʒĀ* ‘a kind of stinging insect’ (NCED 823)

Bsq *\*sinaguri* ‘ant’ (insect) ~ PNC *\*ʒHĒmVĵĀ* ‘ant’ (NCED 1093)

Bsq *\*sinca*, *\*since-ri*, *\*sinci-la* ‘bell, cowbell’ ~ PNC *\*ʒi[l]ʒV* ‘brass; item made of brass’ (NCED 1095)

Bsq *\*solhi* ‘ripe, mature’, etc. ~ PNC *\*ʒōlV* ‘healthy, whole’ (NCED 1095)

Bsq *\*suřin* (in *\*gult-suřin* ‘kidney’) ~ PEC *\*ʒwěre* / *\*rěʒwe* ‘urine’ (NCED 1097)

Bsq *\*susa* / *\*susta* (in *\*mar-sus-(t)a* ‘blackberry, mulberry’) ~ PEC *\*ʒimʒV* ‘a kind of berry (or flower, blossom)’ (NCED 1095)

Bsq *\*susen* ‘right, correct, just’, *\*sin* ‘oath, truth; true’ ~ PNC *\*HĵĒm-* ‘clean, to clean’ (NCED 552)

**4.b.** The match of Bsq *\*e* with PNC *\*i* is provisionally explained as due to EuC ablaut of *\*e* / *\*i*. The regional distribution of Western Bsq *negar* ‘tears, weeping’ vs. Eastern Bsq *nigar* id., an otherwise unexplained variation, seems to correspond to PEC *\*nēwqū* [direct] vs. *\*nĩwqV̄-* [oblique] ‘tear; pus’ (NCED 848). In this case the originally morphological allomorphs were apparently redistributed in Bsq as regional variants. Other cases fitting the same pattern (Bsq *\*e* = PNC *\*i*) include:

Bsq *\*a=mec* ‘dream, sleep’ ~ PNC *\*fnićwV̄* ‘night, evening’ (NCED 524)

Bsq *\*egu-ski* ‘sun’, *\*egu-* ‘day’; *\*eki* ‘sun’ ~ PEC *\*HwĩqV* ‘day’ (NCED 622)

Bsq *\*e=henši* ‘fleeing, to flee’ ~ PNC *\*hilćwĒ* ‘to run’ (NCED 490)

Bsq *\*e(=)reka* ‘ravine, stream’ ~ PNC *\*rĩqwĀ* ‘mountain, rock; cave’ (NCED 953)<sup>32</sup>

Bsq *\*erho* ‘to kill’ ~ PNC *\*HĩrĵA* ‘to beat, hit, throw’ (NCED 581)

Bsq *\*(es)-pel* ‘box (tree)’ ~ PEC *\*phĩlV* ‘a kind of deciduous tree’ (NCED 870)

Bsq *\*hega-n* ‘to fly’ ~ PEC *\*=HiĵV* ‘to fly’ (NCED 582)

Bsq *\*herći*, *\*herš-tu* ‘to close, squeeze’, etc. ~ PNC *\*HićĀn* / *\*HićĀn* ‘to press, squeeze’ (NCED 568)

Bsq *\*keda-r̄* ‘soot’ ~ PEC *\*qĩdV* ‘soot, dust’ (NCED 927); see 7.b.

Bsq *\*negu* ‘winter’ < PEuC *\*nĵěwV* ~ Chechen *ʒa-* / *ʒāna-* ‘winter’ < PNC *\*Bwĩn?V* ‘name of a season’ (NCED 482)

See also 1.b. (with putative EuC *\*ə* / *\*i* ablaut) and 2.b. (with *\*a* / *\*u* ablaut), and 7.b. (with *\*e* / *\*i* ablaut).

<sup>32</sup> ‘ravine, valley’ in Andian.

**Table 5:** Basque *\*hardo* ‘tinder’ ~ PEC *\*HwVrdV* ‘a kind of herb’:

Due to scant attestation of this root in NC (only three Dagestanian languages) it is impossible to reconstruct the PEC vowels, and the quality of the initial laryngeal is unclear; there is otherwise no phonetic difficulty with the comparison. However, the semantic connection between Bsq ‘tinder’ (generally *amadou* fungus) and PEC ‘a kind of herb’ is rather tenuous, though clearly a variety of dried and shredded herbs may be used as tinder. Another complication is the contamination or blending of the Bsq word, in some dialects, with reflexes of Latin *carduus* ‘thistle’ (see OEH: ARDAI, HARDDO, KARDO, KARJO), so much so that Michelena and other Vasconists have derived Bsq *\*hardo* from Latin *carduus*. Overall, this comparison must be reckoned as less convincing than the other six in this paper.<sup>33</sup>

**Table 6:** Basque *\*hauć* ‘ashes, dust’ ~ PEC *\*jōmćV* ‘earth’:

**6.a.** The greatest phonetic difficulty seems to be the initial match of Bsq *\*h* = PEC *\*j*, otherwise unprecedented. However, there may be some clues in the EC etymology, namely (a) the initial “voiced emphatic laryngeal fricative” /ʕ/ in Dargi, and (b) the pharyngealized affricate /čʰ/ in Udi (written /čI/ in NCED), both of which suggest a laryngeal in the proto-language. The normal reflex of PEC *\*j* in Dargi is the glottal stop /ʔ/ (NCED 44), and the normal reflex of PEC *\*ć* in Lezgian languages is /č/; the reflexes /ʕ/ and /čʰ/ are not explained in the NCED etymology of *\*jōmćV*; however, it is stated that “in some cases a new pharyngealization, missing in other languages, appears [in Udi]. The reason for this ... [is] not yet clear” (NCED 143). Thus I would venture to suggest that the Dargi /ʕ/ and Udi /čʰ/ could possibly reflect an old form with an incorporated class prefix, *\*i=HōmćV* (> later *\*jōmćV*, with deletion of the laryngeal), which could explain the reflexes in Dargi and Udi, and the match with Bsq *\*h* in *\*hauć*.

**6.b.** For the match of Bsq *\*-au-* with PEC *\*-ōm-*, in a significant number of cases Bsq diphthongs (*au*, *ai*, *ei*, *oi*) were created when PEuC internal *\*m* disappeared, leaving a trace as Bsq *\*u* or *\*i*:

- Bsq *\*e=aici* ‘to milk (a cow)’ ~ PNC *\*=āmzŪ* ‘to milk’ (NCED 262)
- Bsq *\*eiho* ‘to grind’, *\*eihera* ‘mill’ ~ PNC *\*HēmχwV* ‘to grind’ (NCED 559)
- Bsq *\*e=roic* ‘cliff, fissure’ ~ PEC *\*hrōmćwe* ‘stone’ (NCED 495)
- Bsq *\*gau* / *\*gab-* ‘night’ ~ Proto-Lezgian *\*χI:am*: ‘evening, night’ (TOB; Bengtson 2015)<sup>34</sup>
- Bsq *\*haic* ‘stone, rock, crag’ ~ PNC *\*hāmVćō* ‘stone’ (NCED 516)<sup>35</sup>
- Bsq *\*haise* ‘wind’ ~ PNC *\*ǰāmsā* ‘breath, soul; cloud, sky; god’ (NCED 243)
- Bsq *\*haušin* / *\*hašuin* ‘nettle’ ~ PEC *\*hωāmšwi* ‘a kind of herb or weed’ (NCED 536)
- Bsq *\*kaio* ‘seagull var.’ ~ PEC *\*qāmħā* ‘raven, crow’ (NCED 889)
- Bsq *\*kain* ‘fog, clouds’ ~ PNC *\*kωāmħV* ‘cloud, fog’ (NCED 737)
- Bsq *\*koipe* ‘rendered fat, grease’ ~ PEC *\*kāmǰā* ‘butter; sour cream’ (NCED 716)

<sup>33</sup> Alternatively, a case could perhaps be made for comparing Bsq *\*hardo* with PEC *\*hōrtV* ‘a kind of grass [i.e. herb]’: Avar *het* ‘feather-grass’; Dargi *ʕat-uri* ‘ramson’; Tabasaran *urt-im* ‘fern, moss’, Archi *hoti* ‘grass’ (NCED 527).

<sup>34</sup> In an alternative notation: *\*χʰ:am*: / *\*χʰ:an*: (tense pharyngealized voiceless uvular fricative).

<sup>35</sup> Here we assume the Bsq word came from a syncopated form such as *\*hVmćV*.



Bsq *\*paita* ‘duck’ < PSC *\*pHām[t]V̄* ~ PEC *\*bHāmV* ‘duck’ (NCED 301)  
 Bsq *\*šaie* ‘vulture’ ~ PEC *\*čwāmV̄* ‘eagle, vulture’ (NCED 370)

**6.c.** The match of the Bsq apico-alveolar affricate *\*č* with the nearly identical PNC/PEC alveolo-palatal affricate *\*č* is also well supported, in cases that involve the proximity of labial vowels or consonants (cf. 2.a, above):

Bsq *\*a=meč* ‘dream, sleep’ ~ PNC *\*fničwV̄* ‘night, evening’ (NCED 524)  
 Bsq *\*iču* ‘blind; blind person’ ~ PEC *\*VčV* ‘blind, dark’ (~ *\*=VčV*) (NCED 1017)<sup>36</sup>  
 Bsq *\*mahač* ‘grape(s)’ ~ PNC *\*šāmčō* ‘apple; medlar’ (NCED 237)  
 Bsq *\*oroc* ‘male animal’ ~ Agul *urč* ‘calf’, etc. < PEC *\*wHilčV* (NCED 1048)  
 Bsq *\*pač* ‘pomace (of apples, grapes)’ ~ PEC *\*bHačwe* ‘bark, skin, leaf’ (NCED 297)

**Table 7:** Basque *\*keda-ř* ‘soot’ ~ PEC *\*qidV* ‘soot, dust’:

**7.a.** The match of initials, Bsq voiceless velar stop *\*k* = PNC/PEC glottalized uvular affricate *\*q̣*, is supported by multiple examples:

Bsq *\*kelu* ‘stench, bad taste’, *\*ke[l]jač* ‘bitter, sour; stench’ ~ PNC *\*q̣čhilV* ‘bitter; salty; sweet’ (NCED 912)  
 Bsq *\*kiriki-* ‘hedgehog’ ~ PEC *\*q̣Hwir-dV* ‘hedgehog’ (NCED 926)  
 Bsq *\*koco* ‘male quadruped’ ~ PEC *\*q̣HōcV* ‘kid, hornless goat’ (NCED 920)  
 Bsq *\*koču* ‘bowl, vessel’ ~ PEC *\*q̣HečwV* ‘vessel, jar’ (NCED 920)  
 Bsq *\*koko* ‘egg; food’ (children’s word) ~ PEC *\*q̣wāq̣wV(-tV)* ‘egg, grain, seed’ (NCED 932)  
 Bsq *\*kone* in *\*hars-kone* ‘badger’ ~ PNC *\*h̄q̣wēnē* / *\*f̄nēq̣wē* ‘mouse, rat’ (NCED 523)  
 Bsq *\*konko-ř* / *\*kunku-ř* ‘hump, joint, bone’ ~ PNC *\*q̣ām̄q̣(w)ā* ‘knee, leg-bone’ (NCED 907)  
 Bsq *\*kuku-* (with various suffixes) ‘peak, top, crest’ ~ PEC *\*q̣HwVq̣V* ‘hill; pillar, post’ (NCED 927)  
 Bsq *\*kuřu(-to)* ~ *\*kuři(-to)* ‘crane (bird)’ ~ PNC *\*q̣wVřV* ‘crane’ (NCED 914)

**7.b.** As to the root vowels, Bsq *\*keda-* ~ PEC *\*qidV* fits the postulated Euskaro-Caucasian ablaut pattern *\*e / \*i* (see 4.b).

**7.c.** Finally, there is the trivial match of Bsq *\*d* = PNC/PEC *\*d*, which holds in Bsq medial position:<sup>37</sup>

Bsq *\*hodoē* ‘cloud’ < *\*HVdwi* (*vel sim.*, with metathesis) ~ PEC *\*dwiHV* ‘wind’ (NCED 407)  
 Bsq *\*i=di* ‘ox’ ~ Abkhaz *a-tš* ‘ram’, etc. < PNC *\*dwānV* / *\*dwānhV* ‘sheep, ram’ (NCED 405)  
 Bsq *\*i=dul-ki* ‘block’ ~ Ingush *tāl-g* ‘chock’, etc. < PEC *\*dwāh̄i*, *\*dwāh̄i-kV* ‘stick’ (NCED 405)  
 Bsq *\*sede* ‘boundary, marker’ ~ PEC *\*šwādV* ‘rock, stone; tombstone’ (NCED 975)  
 Bsq *\*u-dagera* ‘otter’ ~ PEC *\*darq̣wV* ‘weasel, ermine’ (NCED 399)<sup>38</sup>

## Appendix C: A note about the hypothesis of fossilized noun class prefixes in Basque

As developed in detail in my earlier work, and summarized most recently in Bengtson 2017a (see especially pp. 58–71), it is hypothesized that the phonetic forms of certain

<sup>36</sup> Due to paucity of attestation (only Avar-Andian-Tsezian) the original vowels are uncertain, as well as the internal consonant (tense *\*c* or *\*č*).

<sup>37</sup> However, in initial position all PEuC coronal consonants merge in Bsq as *\*t-*, e.g. Bsq *\*toš-ko* ‘block of wood’ ~ PEC *\*dūčV* ‘wood, firewood’ (NCED 401); Bsq *\*tanka*, *\*tanko* ‘drop (of liquid)’ ~ PNC *\*tHān̄kō* ‘drop, spray’ (NCED 1000), etc.

<sup>38</sup> Bsq *\*u-dagera* ‘otter’ seems to reflect *\*hur-dagera* (with *\*hur* ‘water’), i.e. ‘water-weasel’.

Basque nouns with otherwise unexplained initial vowels, and the frequent initial element *b-/be-/bi-*, can be explained on the basis of an underlying noun class (gender) system, which is still fully productive in most East Caucasian languages, and in part of the West Caucasian branch (Abkhazian); grammatical gender was lost in Circassian, Ubykh, Lezgian, Agul, Udi, and some dialects of Tabasaran.

Already in 1927 Uhlenbeck noticed the initial *b-* elements in Basque nouns, and suggested that they could be typologically compared with the class prefix systems exemplified in the Bantu languages, but the idea was not developed any further. In a brief paper Chirikba (1985) adopted this proposal, as seen in lexical comparisons like that of Bsq *belarri* (*be=larri*) ‘ear’ with Proto-East Caucasian *\*lérhIV* ‘ear’: Dargi *lihi*, Chechen *ler-g*, Batsbi *lar-k* ‘ear’.<sup>39</sup>

Since North Caucasian languages also share numerous lexical and grammatical homologies with Basque, it is reasonable to posit that the unexplained initial segments are “stage III articles” (Greenberg 1978), that is, morphemes that formerly marked class distinctions, analogous to those in East Caucasian languages, and at some time in the past became degrammaticalized, i.e. lexicalized and fused with the noun stem (Croft 2003: 260). Besides the *\*b(e/i)=* segment, I have identified other initial segments that can be designated as *\*e=/\*i=*, *\*o=/\*u=*, and *\*a=*, with weaker evidence for *\*m(a)=* and perhaps other prefixes.<sup>40</sup> Based on the patterning of the prefix *\*b=/\*be=/\*bi=*, with the mid-high vowel alternation /e/ ~ /i/, I have proposed that *\*e=/\*i=* and *\*o=/\*u=* each represent allomorphs of functionally identical prefixes.

I have further suggested that Bsq *\*b=/\*be=/\*bi=*, corresponds formally to the East Caucasian (EC) inanimate class marker *\*b-* (Deeters 1963) or *\*w-/\*b-* (S. Starostin 2002) [III-class singular]; *\*e=/\*i=* to EC *\*j-* [= *\*y-*: II-class singular] (S. Starostin 2002) or [feminine rational] (Deeters 1963); and *\*o=/\*u=* to EC *\*w-* (Deeters 1963) or *\*u-* [I-class singular] (S. Starostin 2002). The Bsq prefix *\*a=* cannot, so far, be linked to any of the EC class markers, but is similar to Abkhazian *\*a-/0-* [third singular inanimate oblique].

Now as to whether there is any isomorphism between the Bsq fossilized prefixes and the NC class markers, there are strong phonetic resemblances, but imperfect semantic matches. However, here we must remember that historically preceding the synchronic EC and WC class systems we are thinking about at least two deeper levels, (a) the Proto-North Caucasian (PNC) noun class system, and (b) the putative Proto-Euskaro-Caucasian noun class system. The Proto-Basque noun class system was derived from the system at level (b), and it would be entirely likely that noun class systems should have undergone extensive changes between stages (b), (a), and the existing systems in modern NC languages.

<sup>39</sup> The PNC protoform was later changed to *\*lérHhe* (~-i) ‘ear’ (NCED 756), and then again to *\*leHe* ‘ear’. The rhotic common to Basque and the Nakh words can be attributed to generalization of the oblique stem; cf. Proto-Nakh direct stem *\*la* / oblique stem *\*lari-*. The latter, with a diminutive suffix (*\*lari-k*) replaced the old direct forms in Chechen, Ingush and Batsbi (Bengtson & G. Starostin 2015: 21).

<sup>40</sup> Cf. Bsq *\*mo=kol(o)* ‘shell (of egg, nut), husk (of maize)’, beside (Bzt) *akal* ‘empty (of a chestnut shell)’, (B) *kakol* ‘shell’ ~ PEC *\*qwālV* ‘bark, crust’ > Akhwakh *qoli* ‘crust, rind’, Tsezi *q’ul* ‘bark’, Bezhta *qeqel*-ba ‘birch bark’, etc. (NCED 931).

Matasović (2017) has recently tried to reconstruct the class system at level (a), i.e. the PNC level. Regarding EC  $*y-$  (=  $*j-$ ) he concludes that “[PNC] 3 sg. inanimate  $*y-$  became the new [PEC] 3 sg. feminine human form (class II; but it also continued to be used for some inanimates ...” (Matasović 2017: 16). This would be consistent with the description of the Bsq  $*e=/*i=$  prefix, which includes a significant number of inanimate nouns involving masses and uncountables (e.g.  $*e=thu-r$  ‘snow’,  $*e=sene$  ‘milk’,  $*i=sar$  ‘star’, and  $*e=kē$  ‘smoke’,  $*i=ču$  ‘fire’, discussed above), and also at least one noun with a clearly feminine meaning:  $*i=so$  /  $*i=se-ba$  /  $*i=se-ko$  ‘aunt (maternal or paternal)’ (Bengtson 2017a: 64–67).<sup>41</sup>

The rules governing the distribution of /e/ versus /i/ in nouns with the  $*e=/*i=$  prefix are not yet fully understood, but seem to be related to rules of vowel harmony, and raising (closing) and lowering (opening) of vowels, which differ in the various Basque dialects; Michelena discusses these in the chapter “Apertura y cierre” (FHV 59–71). Regarding the phonetically similar Bsq verbal prefix  $*e-$ , which likewise varies between /e/ and /i/ in Bsq dialects, Trask offers this rule: “When the /e/ is followed by a consonant followed by a high vowel /i/ or /u/, and a third syllable exists, /e/ is raised to /i/: *ebili* (attested) > *ibili* ‘be in motion’; *ekusi* (attested) > *ikusi* ‘see’” (EDB 32). This rule works for some nouns like Bsq  $*i=turi$  ‘fountain, spring’,  $*i=thinti$  ‘firebrand, ember’, and  $*i=dul-ki$  ‘block of wood’; on the other hand, when the only following vowel is /e/, the prefix remains /e/, e.g. Bsq  $*e=sene$  ‘milk’,  $*e=ce$  ‘house’,<sup>42</sup>  $*e=rece$  ‘corner, edge, side’,  $*e=the$  ‘word, speech’,  $*erle$  ‘bee; beehive’ <  $*e=rethe$ , etc. Nevertheless, as noted by Trask, exceptions that violate his rule (*ebili*, *ekusi*) are also attested; likewise, for example, in a Bsq word for ‘buttocks, arse’, which synchronically varies between a prefix /e/ in (AN) *epurdi*, (B) *eperdi*; /i/ in (L) *iphurdi*, (Z) *iphürdi*, (A, G, Sal, B-Markina, Oñate) *ipurdi*, (B-arc) *ipirdi*; or, with no prefix, (Bzt, BN-Aldude) *purdi*. From among these the Basque Academy selected *ipurdi* as the standard (EB) form.

## Appendix D: The Prague School and the Euskaro-Caucasian hypothesis

One of the founders of the Prague School, Nikolai Trubetzkoy, produced pioneering studies (1922, 1926) in the comparative phonology of the NC languages, upon which Abdokov and the authors of NCED built decades later. Dumézil, who was also instrumental in demonstrating the underlying grammatical unity of the North Caucasian languages (see the note about NCED, above), and also in his 1933 book directly compared the grammatical systems of NC and Basque,<sup>43</sup> was clearly deeply influenced by the Prague School, and the emphasis on structure and “le système”, in his linguistic work and

<sup>41</sup> EB *izeba* ‘aunt’, western Bsq *izeko* id. (with a different suffix); (Azk) *izo*, (Bzt) *izoa* ‘aunt’ with the simple stem  $*i=so$ ; cf. the cognate Chechen *dēca* ‘paternal aunt’ < PEC  $*r=ičwīV$ , with the EC feminine class prefix  $*r=$ , < PEC  $*(-i)čwīV$  ‘girl, woman’.

<sup>42</sup> There is a secondary change to /i/ in some southern Bsq dialects, *itxe* ‘house’, caused, according to Michelena, by the following palatal affricate /č/ (FHV 67; EDB 36).

<sup>43</sup> “Mais avec le caucasien du Nord, la coïncidence [avec basque] est, sur tous les points importants, presque complète” (Dumézil 1933: 123–125 & note 5).

studies of mythology. Another founding member of the Prague School, Roman Jakobson, was, among other things, a keen student of the so-called “Paleosiberian languages” (e.g. Jakobson, Hüttl-Worth & Beebe 1957), later a favorite subject of the Moscow School (see below) in their studies on Yeniseian, Chukchi-Kamchatkan, Nivkh and Yukaghir languages.

The Prague Circle of the 1920s and 1930s profoundly impacted scholars in North America, notably the “First Yale School” associated with Edward Sapir and Morris Swadesh, as described by Dell Hymes (1971). It also influenced the Moscow School, led by V. M. Illich-Svitych, et al.<sup>44</sup>

In 1988 I met Václav Blažek, who had studied in Prague at Charles University and is now a professor at Masaryk University. From 1978–1990 Blažek had been in frequent contact with Vladimír Skalička, a well-known member of the Prague School, and the two shared an interest in the Basque language (e.g. Skalička 1984; Blažek 1992, 1995, 1999; Blažek & Bengtson 1995). Blažek, who has collaborated with me several times, is also closely associated with the Moscow School.

These international contacts, together with discussions with North Americans influenced by the First Yale School, have helped me to formulate strategies for genetic linguistics combining the best elements of each school’s methods. These have been applied to producing the current model of Euskaro-Caucasian, and may briefly be summarized as:<sup>45</sup>

- Only linguistic evidence, and only specific resemblances involving both sound and meaning are relevant to genetic classification. Resemblances in typology alone are not relevant to genetic classification.
- Multilateral comparison of languages is more effective for genetic classification than comparisons between pairs of languages.
- Evidence should be drawn from both lexicon (basic vocabulary) and grammar (morphology), and the conclusions from both should lead to the same results.
- Grammatical paradigms, or parts of paradigms, and especially suppletive paradigms, are especially convincing in genetic linguistics.
- Semantic changes must be expected, but they should be plausible, and whenever possible typologically similar changes should be cited.
- Comparative phonology is a subsidiary but important component of etymology that helps the linguist to test etymologies, to detect false cognates and distinguish loanwords from genuine cognates.
- The goal of genetic linguistics is to provide the *best explanation* possible consistent with the facts, rather than to attain some arbitrary threshold of absolute “proof”.

---

<sup>44</sup> See G. S. Starostin, Zhivlov & Kassian (2016). This late 20<sup>th</sup> century Moscow School (or Nostratic School) is to be distinguished from the Moscow Linguistic Circle of several decades earlier, associated with Jakobson and Trubetzkoy, and a forerunner of the Prague School, though naturally there are threads connecting the early Moscow Circle with the later one.

<sup>45</sup> In compiling this list of principles I have relied especially on Greenberg (1966, 1987, 1995) and Newman (1995, 2000).

## ABBREVIATIONS

Basque dialects often exhibit extremely diverse forms of an etymon, thus sub-dialectal indications such as AN-Ulzama (the community of Ulzama in High Navarre), BN-Baigorri (the community of Baigorri in Low Navarre); similarly for internally diverse NC languages, e.g. Dargi-Akusha, Dargi-Kubachi, etc.

### 1. Languages and dialects

A	Araban: Bsq dialect of Araba (Álava) (= Michelena's <i>meridional</i> ['southern'])
Azk	Aezkoan (Bsq dialect)
AN	Alto Navarro = High Navarrese (Bsq dialect)
arc	Archaic or obsolete form
B	Bizkaian = Biscayan (Bsq dialect)
Bsq	Basque
Bzt	Baztanese (Bsq dialect)
BN	Bas-navarrais = Low Navarrese (Bsq dialect)
c	Common Basque: the form in all dialects / or the most widespread form
EB	<i>euskara batua</i> (standard Basque)
EC	East Caucasian (a.k.a. Northeast Caucasian or Nakho-Dagestanian)
G	Gipuzkoan (Bsq dialect)
L	Lapurdian = Labourdin (Bsq dialect)
NBsq	Northern Basque (literary)
NC	North Caucasian (= EC + WC)
PEC	Proto-East Caucasian (see NCED)
PEuC	Proto-Euskaro-Caucasian
PL	Proto-Lezgian
PNC	Proto-North Caucasian (see NCED)
PSC	Proto-Sino-Caucasian
PWC	Proto-West Caucasian
R	Roncalese (Bsq dialect)
Sal	Salazarese (Bsq dialect)
WC	West Caucasian (a.k.a. Northwest Caucasian or Abkhazo-Adyghean)
Z	Zuberoan = Souletin (Bsq dialect)

### 2. References

ASJU	<i>Anuario del Seminario de Filología Vasca 'Julio de Urquijo'</i>
AW	Aulestia & White (1992)
Č	<i>Common West Caucasian</i> (Chirikba 1996)
EDB	<i>Etymological Dictionary of Basque</i> (Trask 2008)
EHL	Evolution of Human Language (Project files). An International Project on the Linguistic Prehistory of Humanity coordinated by the Santa Fe Institute. <sup>46</sup> Available online at: < <a href="http://ehl.santafe.edu/main.html">http://ehl.santafe.edu/main.html</a> >.
FHV	<i>Fonética Histórica Vasca</i> (Michelena 1961)
HB	<i>History of Basque</i> (Trask 1997)
MDE	<i>Materiales para un diccionario etimológico de la lengua vasca</i> (Agud & Tovar 1988-)
NCED	<i>North Caucasian Etymological Dictionary</i> (Nikolayev & S. Starostin 1994)

---

<sup>46</sup> In general the data obtained from EHL and TOB (see below) are equivalent. Many EHL/TOB files are available on the websites cited, while some others are privately used and continually edited by EHL researchers.

OEH	<i>Orotariko Euskal Hiztegia</i> (Michelena 1987)
SKJa	<i>Slovar' kavkazskix jazykov</i> (Klimov & Xalilov 2003)
TOB	Starostin, S. A., et al. (2005—): Tower of Babel: International, Web-based project on historical and comparative linguistics developed by S. A. Starostin & Yu. Bronnikov. Available online at: < <a href="http://starling.rinet.ru/main.html">http://starling.rinet.ru/main.html</a> >.

## REFERENCES

- ABDOKOV, A. I. (1981): *Vvedenie v sravnitel'no-istoričeskiju morfologiju abxazsko-adygskix i naxsko-dagestanskix jazykov* [Introduction to the Comparative-Historical Morphology of the Abkhaz-Adygean and Nakh-Daghestanian Languages]. Nal'čik: Kabardino-Balkarskij Gosudarstvennyj universitet.
- ABDOKOV, A. I. (1983): *O zvukovyx i slovarnyx sootvetstvijax severokavkazskix jazykov* [On Sound and Lexical Correspondences of the North Caucasian Languages]. Nal'čik: El'brus.
- AGUD, M. & TOVAR, A. (1988–1994): *Materiales para un diccionario etimológico de la lengua vasca*. Published in fascicles in *ASJU*. I. 22 (1988): 253–312; II. 23 (1989): 463–532; III. 24 (1990) 111–202, 615–668, 819–870; IV. 25 (1991): 255–314, 543–622, 805–864; V. 26 (1992): 281–340, 645–694, 825–914; VI. 27 (1993): 321–360, 613–692; VII. 28 (1994): 263–332, 631–682, 915–992.
- ANTTILA, R. (1972): *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan.
- AULESTIA, G. & WHITE, L. (1992): *Basque-English English-Basque Dictionary*. Reno: University of Nevada Press.
- AZKUE, R. M. DE (1905–1906): *Diccionario vasco-español-francés / Dictionnaire basque-espagnol-français*. Bilbao, Paris: The Author, Paul Geuthner.
- BENGTSON, J. D. (1992): Macro-Caucasian phonology. In: V. Shevoroshkin (ed.), *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum: Brockmeyer, 342–351.
- BENGTSON, J. D. (1996): Correspondences of Basque and Caucasic final vowels: -i/-e, -u/-o. *Fontes Linguae Vasconum* 71, 7–15.
- BENGTSON, J. D. (2004): Some features of Dene-Caucasian phonology (with special reference to Basque). *Cahiers de l'Institut de Linguistique de Louvain* 30 (4), 33–54.
- BENGTSON, J. D. (2008): Materials for a comparative grammar of the Dene-Caucasian (Sino-Caucasian) languages. In: A. V. Dybo, V. A. Dybo, A. O. Mudrak & G. S. Starostin (eds.), *Aspects of Comparative Linguistics* 3. Moscow: RSUH Publishers, 45–118.
- BENGTSON, J. D. (2009): Basque and the other Mediterranean languages. *Mother Tongue* 14, 157–175.
- BENGTSON, J. D. (2015): Basque Etymology Database. EHL/TOB. Available online at: <<http://starling.rinet.ru/main.html>>, <<http://ehl.santafe.edu/main.html>>.
- BENGTSON, J. D. (2017a): *Basque and its Closest Relatives: A New Paradigm*. Cambridge: Mother Tongue Press.
- BENGTSON, J. D. & STAROSTIN, G. (2015): The Dene-Sino-Caucasian hypothesis: state of the art and perspectives. Discussion draft posted on Academia.edu. Available online at: <[https://www.academia.edu/13072385/The\\_Dene-Sino-Caucasian\\_hypothesis\\_state\\_of\\_the\\_art\\_and\\_perspectives](https://www.academia.edu/13072385/The_Dene-Sino-Caucasian_hypothesis_state_of_the_art_and_perspectives)>.
- BERG, H. VAN DEN (1995): *A Grammar of Hunzib (with Texts and Lexicon)*. München, Newcastle: Lincom Europa.
- BERG, H. VAN DEN (2005): The East Caucasian language family. *Lingua* 115, 147–190.
- BLAŽEK, V. (1992): Basque and North Caucasian or Afroasiatic? In: E. Ebermann (ed.), *Komparative Afrikanistik: sprach-, geschichts- und literaturwissenschaftliche Aufsätze zu Ehren von Hans G. Mukarovskij anlässlich seines 70. Geburtstags*. Wien: Afro-Pub, 21–30.
- BLAŽEK, V. (1995): Towards the position of Basque: A reply to Trask's critique of the Dene-Caucasian hypothesis. *Mother Tongue* 1, 104–110.
- BLAŽEK, V. (1999): Some remarks on 'A Comparison of Basque and (North) Caucasian Basic Vocabulary' by John D. Bengtson. *Mother Tongue* 5, 59–61.

- BLAŽEK, V. & BENGTSON, J. D. (1995): Lexica Dene-Caucasica. *Central Asiatic Journal* 39 (1), 11–50, 39 (2), 161–164.
- BLUST, R. (1996): The neogrammarian hypothesis and pandemic irregularity. In: M. Durie & M. Ross (eds.), *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. New York, Oxford: Oxford University Press, 135–156.
- BUCK, C. D. (1949): *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. Chicago: University of Chicago Press.
- CHIRIKBA [ČIRIKBA], V. (1985): Baskskij i severokavkazskie jazyki [Basque and North Caucasian languages]. In: B. B. Piotrovskij (ed.), *Drevnjaja Anatolija*, 95–105. Moskva: Nauka.
- CHIRIKBA, V. A. (1996): *Common West Caucasian: The Reconstruction of its Phonological System and Parts of its Lexicon and Morphology*. Leiden: Research School CNWS.
- CHIRIKBA, V. A. (2016): From North to North West: How North-West Caucasian evolved from North Caucasian. *Mother Tongue* 21, 1–28.
- CROFT, W. (2003): *Typology and Universals*. Cambridge: Cambridge University Press.
- DEETERS, G., SOLTA, G. R. & INGLISIAN, V. (1963): *Armenisch und kaukasische Sprachen*. Leiden, Köln: Brill.
- DUMÉZIL, G. (1933): *Introduction à la grammaire comparée des langues caucasiennes du Nord*. Paris: Champion.
- GREENBERG, J. H. (1966): *The Languages of Africa*. Bloomington, The Hague: Indiana University, Mouton & Co.
- GREENBERG, J. H. (1978): How does a language acquire gender markers? In: J. H. Greenberg (ed.), *Universals of Human Language*. Stanford: Stanford University Press, 47–82.
- GREENBERG, J. H. (1987): *Language in the Americas*. Stanford: Stanford University Press.
- GREENBERG, J. H. (1995): The concept of proof in genetic linguistics. *Mother Tongue* 1, 207–216.
- HYMES, D. (1971): Morris Swadesh: From the first Yale school to world prehistory. In: M. Swadesh, *The Origin and Diversification of Language*. Edited by J. Sherzer. Chicago, New York: Aldine Atherton, 228–270.
- JAKOBSON, R., HÜTTL-WORTH, G. & BEEBE, J. (1957): *Paleosiberian Peoples and Languages: A Bibliographical Guide*. New Haven: HRAF Press.
- KASSIAN, A. (2010): Hattic as a Sino-Caucasian language. *Ugarit-Forschungen* 41, 309–448.
- KASSIAN, A. (2011–2012): Annotated 100-item wordlists for the Lezgian group, North Caucasian family. In: G. Starostin (ed.), *The Global Lexicostatistical Database*. Moscow, Santa Fe: Center for Comparative Studies at the Russian State University for the Humanities, Santa Fe Institute. Available online at: <<http://starling.rinet.ru/new100>>.
- KLIMOV, G. A. (1964): *Ėtimologičeskij slovar' kartvel'skix jazykov* [Etymological Dictionary of the Kartvelian Languages]. Moskva: Nauka.
- KLIMOV, G. A. & XALILOV, M. Š. (2003): *Slovar' kavkazskix jazykov: Sopostavlenie osnovnoj leksiki* [Dictionary of Caucasian Languages: Comparison of the Principal Lexis]. Moskva: Vostočnaja literatura.
- LANDUCCI, N. (1958 [1562]): *Dictionarium linguae cantabricae*. Edited by M. Agud & L. Michelena. San Sebastian: Diputación Provincial de Guipuzcoa.
- LARRASQUET, J. (1939): *Le Basque de la Basse-Soule orientale*. Paris: Klincksieck.
- MATASOVIĆ, R. (2017): The origin of gender in Northwest Caucasian. A revised version of a paper read at the conference Historical linguistics of the Caucasus held at the Ecole pratique des hautes études in Paris, April 2017. Draft posted on Academia.edu. Available online at: <[https://www.academia.edu/34519267/The\\_origin\\_of\\_gender\\_in\\_Northwest\\_Caucasian](https://www.academia.edu/34519267/The_origin_of_gender_in_Northwest_Caucasian)>.
- MICHELENA, L. [MITXELENA, K.] (1961): *Fonética Histórica Vasca*. San Sebastián: Diputación Provincial de Guipuzcoa.
- MICHELENA, L. [MITXELENA, K.] (1987–): *Orotariko Euskal Hiztegia / Diccionario General Vasco*. Bilbao: Euskaltzaindia. [Edited since Michelena's death by I. Sarasola.] Available online at: <[http://www.euskaltzaindia.eus/index.php?option=com\\_oeh&view=frontpage&Itemid=413&lang=eu](http://www.euskaltzaindia.eus/index.php?option=com_oeh&view=frontpage&Itemid=413&lang=eu)>.

- MICHELENA, L. [MITXELENA, K.] (1988): *Sobre historia de la lengua vasca*. Edited by J. A. Lakarra. Donostia/San Sebastián. Anejos del ASJU 10.
- NEWMAN, P. (1995): *On Being Right: Greenberg's African Linguistic Classification and the Methodological Principles which Underlie it*. Bloomington, Indiana: Institute for the Study of Nigerian Languages and Cultures, Indiana University.
- NEWMAN, P. (2000): Comparative linguistics. In: B. Heine & D. Nurse (eds.), *African Languages: An Introduction*. Cambridge: Cambridge University Press, 259–271.
- NIKOLAYEV, S. L. & STAROSTIN, S. A. (1994): *A North Caucasian Etymological Dictionary*. Moscow: Asterisk Publishers.
- SALABURU, P. (2008): *Koldo Mitxelena: Selected Writings of a Basque Scholar*. Trans. by L. White & M. D. Johnson. Reno: Center for Basque Studies, University of Nevada, Reno.
- SKALIČKA, V. (1984): Die Ergativität und die Typologie. In: *Acta Universitatis Carolinae: Philologica 1: Linguistica Generalia 3* (1981). Praha: Univerzita Karlova, 59–63.
- STAROSTIN, G. S. (2010): Preliminary lexicostatistics as a basis for language classification: A new approach. *Journal of Language Relationship / Voprosy jazykovogo rodstva 3*, 79–116.
- STAROSTIN, G. S., ZHIVLOV, M. & KASSIAN, A. (2016): The “Nostratic” roots of Indo-European: from Illich-Svitych to Dolgopolsky to future horizons. *Slovo a slovesnost 77*, 392–415.
- STAROSTIN, S. A. (2002): *North Caucasian Morphology*. Unpublished manuscript.
- TRASK, R. L. (1997): *The History of Basque*. London, New York: Routledge.
- TRASK, R. L. (2008): *Etymological Dictionary of Basque*. Edited by M. W. Wheeler. University of Sussex.
- TROMBETTI, A. (1903): Delle relazioni delle lingue caucasiche con le lingue camito-semitiche e con altri gruppi linguistici. *Giornale della Società Asiatica Italiana 16*, 145–175.
- TROMBETTI, A. (1925): *Le origini della lingua basca*. Bologna: Azzoguidi. [Reprinted, with the author's annotations and introduction by C. Tagliavini. Bologna: A. Forni, 1966.]
- TRUBETZKOY, N. S. (1922): Les consonnes latérales des langues Caucasiques-Septentrionales. *Bulletin de la Société de Linguistique 23*, 184–204.
- TRUBETZKOY, N. S. (1926): Studien auf dem Gebiete der vergleichenden Lautlehre der Nordkaukasischen Sprachen I. *Caucasica 3*, 7–36.
- UHLENBECK, C. C. (1927): Die mit *b-* anlautenden Körperteilnamen des Baskischen. In: H. Dehérain (ed.), *Festschrift Meinhof: sprachwissenschaftliche und andere Studien*. Hamburg: L. Friederichsen & Co., 351–357.

*Association for the Study of Language in Prehistory, c/o Department of Sanskrit & Indian Studies, Harvard University, 1 Bow Street, Cambridge, MA 02138, U.S.A.*

*Evolution of Human Languages, an International Project on the Linguistic Prehistory of Humanity, coordinated by the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, U.S.A.*

*<palaeojdb@hotmail.com>*