

# History of the Basque Language

By Manfred Owstrowski, a German linguist and professor

## I. Language families and genetic language relationships in Europe

Most of the languages spoken in Europe belong to one single language family: Indo-European. Basque is the sole surviving non-Indo-European language in Western Europe, it is classified as a language isolate. Besides Indo-European, there are to be found languages of four other families in Europe; the Uralic family and the Altaic stock are represented, and we have to add two language families in the Caucasian area, namely South Caucasian and North Caucasian.

The Indo-European language family can be divided into 11 branches, consisting of living and/or extinct languages of Europe and parts of Asia: Indo-Iranian, with Sanskrit and modern representatives like Hindi and Punjabi on the Indic side and Persian, Kurdish, Pashto and many other languages on the Iranian side; Armenian; Classical and Modern Greek; Albanian, which presumably is a descendant of the ancient Illyrian language; Italic, originally consisting of Osco-Umbrian and Latino-Faliscan, today represented by the modern descendants of Latin, the Romance languages (Rumanian, Italian, French, Catalan, Spanish, Portuguese and others); Celtic, with Irish (= Gaelic), Welsh and Breton still spoken; Germanic, with the extinct Gothic language, North Germanic (Danish, Swedish, Norwegian, Icelandic) and West Germanic (German, Dutch, Frisian, English); Baltic, here we have to mention Lithuanian and Latvian; Slavic, with Russian, Ukrainian, Polish, Czech, Slovak, Serbo-Croatian, Bulgarian and some others; Tocharian, which is divided into two languages (called „Tocharian A“ and „Tocharian B“) once spoken in an area of western China; finally, Anatolian, a group of long extinct languages (e.g., Hittite and Luwian) of what is now Turkey. All these branches of Indo-European are believed to go back to a single proto-language, called Proto-Indo-European. The area where Proto-Indo-European was originally spoken (the Proto-Indo-European „homeland“) is still a matter of dispute, but various hints point to Eastern Europe, north and north-east of the Black Sea, and it seems to be rather clear that Indo-European languages are relatively late intruders in Western Europe. Concerning the time when Proto-Indo-European must have been in use, one may think of the end of the stone age in Europe. There is a vast literature on historical linguistics in general and on Indo-European in particular; as a starting-point one may take for example:

Anttila, Raimo, 1989, *Historical and Comparative Linguistics*, Amsterdam : John Benjamins,

Beekes, Robert, 1995, *Comparative Indo-European Linguistics*, Amsterdam : John Benjamins

Obviously, Basque is not an Indo-European language, as it can easily be seen if one compares the basic lexicon and grammar (morphology and syntax) of Basque with the basic lexicon and grammar of Indo-European languages. If one looks at the vocabulary of Basque, one finds the most fundamental words of the language strikingly different from those found in Indo-European languages, whereas a great number of other (less basic) words are loans from Latin or from Romance languages. For illustration, I give some basic words in Standard Basque (Euskara Batua) with their equivalents in Sanskrit, Latin, and Russian:

	Basque	Sanskrit	Latin	Russian
"I"	ni	aham	ego	ja
"you, thou"	hi	tvam	tu	ty
"we"	gu	vayam	nos	my
"you (plur.)"	zuek	yu:yam	vos	vy
"who?"	nor	ka-	quis	kto
"ear"	belarri	karna-	auris	uxo
"tongue"	mihi	jihva-	lingua	jazyk
"heart"	bihotz	hrd-	cors, cord-	serdce
"foot"	oin	pa:da-	pes, ped-	noga
"one"	bat	eka-	unus	odin
"two"	bi	dva	duo	dva
"three"	hiru	tri	tres	tri
"four"	lau	catva:r	quattuor	chetyre
"water"	ur	udaka-	aqua	voda
"fire"	su	agni-	ignis	ogon´

Whereas the vocabulary of modern Basque is a mixture of genuine („old“) Basque words and Indo-European (mainly Latin and Romance) loans, the grammatical elements (morphemes) of Basque are generally unrelated to Indo-European ones, so they constitute the non-Indo-European frame of the language. The inflexion of Basque nouns and verbs is totally different from that of Indo-European languages; just look at the following comparison of the Basque present tense forms of the verb for „to be“ with the corresponding forms of the Indo-European languages Sanskrit, Classical Greek, and Latin:

	Basque	Sanskrit	Class.	Greek	Latin
"(I) am"	naiz	asmi	eimi	sum	
"(you sg.) are"	haiz, zara	asi	ei	es	
"(he, she) is"	da	asti	esti	est	
"(we) are"	gara	smaH	esmen	sumus	
"(you pl.) are"	zarete	stha	este	estis	
"(they) are"	dira	santi	eisi(n)	sunt	

The Indo-European forms belong to a verbal root \*es-/\*s- „to be“. The irregular Basque paradigm allows at least one historical reconstruction: (zu) zara, which is today the unmarked equivalent for „(you sg.) are“, is obviously the old expression for \* „(you plur.) are“, which came to be used in the singular. Zarete, then, is a more recent formation, responding to the need to avoid ambiguity. The original paradigm of Basque certainly had the pairs naiz „(I) am“ : haiz „(you sg.) are“ and gara „(we) are“ : zara \* „(you pl.) are“. Zu „you (sg.)“ is the old plural pronoun \* „you (pl.)“, widely replacing hi „thou, you (sg.)“, which is nowadays only used in a very familiar style. Similarly, English you is a plural pronoun which spread into the singular, felt as more „polite“, and which came to be the unmarked form there. But English you is still singular and plural, whilst Basque created a new plural zuek „you (pl.)“, confining zu to the singular. So Basque went a step farther than English here.

To repeat it, one can state that Basque paradigms are quite different from Indo-European ones. Both the form of the involved elements and the underlying structure show that Basque is far away from Indo-European. The grammatical structure of Basque has many features which distinguish Basque from the Indo-European „standard“, simultaneously reminding of the situation in some other non-Indo-European languages. One of the most important features of Basque is its ergative case. In Basque, the subject in sentences with intransitive verbs and the object in sentences with transitive verbs appear in the same, morphologically unmarked case, which may be called absolutive or nominative. On the other hand, the subject (agent) in sentences with transitive verbs appears normally in the ergative case. There is, of course, no ergative in the bulk of Indo-European languages, here we find, instead, a special case for the object in sentences with transitive verbs: the accusative. The opposition between ergative:absolutive-structures and nominative:accusative-structures is a very important topic in typological studies of language diversity; I give just one example for a book dealing with this question: Dixon, Robert M.W., 1994, **Ergativity**, Cambridge : Cambridge University Press

To get an overview of Basque grammar in general, one may study the following book (taking in consideration that beside Standard Basque (Euskara Batua) there are several Basque dialects, differing in structural matters in varying degree; the book does not cover all the variation):

Saltarelli, Mario, 1988, **Basque**, London/New York : Croom Helm. Please be aware that this book is primarily written for linguists interested in language structure; you will not learn to speak Basque with its help. If you are interested in getting some command of Basque, you may consult: King, Alan R., 1994, **The Basque Language. A Practical Introduction**. Reno/Las Vegas : University of Nevada Press

Let us return to the classification of the languages of Europe. I should mention the non-Indo-European languages beside Basque now. The Uralic language family consists of two subfamilies, Finno-Ugric and Samoyed. There are 12 Finno-Ugric languages in Europe. Estonian, Finnish and Saami (Lapp) are languages belonging to the Finnic branch of Finno-Ugric, Hungarian represents Ugric. The common ancestor of the Finno-Ugric languages, Proto-Finno-Ugric, is said to have been spoken west of the Ural mountains, perhaps stretching into West Siberia. The Altaic language stock is represented in Europe by some Turkic languages. The Turkic group has its origins in Siberia. A very remote relationship between Uralic and Altaic seems to be possible, but is hard to prove in detail.

In the area of the Caucasus mountains, between the Black Sea and the Caspian Sea, we find a large number of non-Indo-European and non-Altaic languages. These languages can be grouped into two distinct language families, South Caucasian and North Caucasian. South Caucasian comprises four

languages: Georgian (the best known), Mingrelian, Laz, and Svan. North Caucasian is divided into two rather distinct subfamilies, West Caucasian and East Caucasian. West Caucasian shows two language groups, one including Abkhaz and Abaza, the other including Adygh and Kabardian. East Caucasian yields 29 different languages; as for the number of speakers, the most important ones are: Chechen and Ingush; Avar; Lak; Dargva; Lezgian and Tabasaran. Obviously, South Caucasian and North Caucasian are spoken in the Caucasus area since prehistoric times. Between South Caucasian and North Caucasian, no genetic relationship could so far be established. In spite of this, both language families have already been subject to comparisons with Basque. This leads us to the next topic: Basque in historical language comparison.

## II. Basque and other languages

As a matter of fact, Basque has been compared with a lot of different languages (and language families) all over the world. For many people it is not easy to accept that some language(s) should escape classification. And there are always some enthusiasts who are not willing to acknowledge that some dozens of „correspondences“ can be found between any two languages of the world, so they try to point out imaginary „macro-families“ or „super-stocks“. A serious research of the relations of Basque will start in South-Western Europe. Here we find first two areas of investigation: Basque and Aquitanian, and Basque and Iberian. Then we may ask: are there traces of extinct languages - which could have been relatives of Basque - to be found in modern languages of Western Europe? It seems to be quite possible that some words of an ancient language akin to modern Basque could have survived in Western Romance. As a next step, it surely is promising to look for elements common to Basque and Celtic. We know for sure that speakers of an early form of Basque and Celtic speakers must have been in contact for an important time-span. Perhaps it would not be unrewarding either to consider possible contacts between the ancestor of Basque and an early pre-Celtic dialect of Indo-European, which may have entered Western Europe.

In my opinion, the only rewarding field of study that will be left here after the treating of the mentioned fields of research is the problem of an assumable historical relation between Basque and North Caucasian, South Caucasian, or both. Are there remote connexions between Basque and the Caucasus?

At this point, one has to concede that the study of Basque language history cannot contribute substantially to our understanding of the origin of the Basque people, but for the formation of the Basque people, we dispose of informations from prehistory (archeology) and anthropology (genetics). These seem to show that the basic layer of the Basque population is a union of two components, one being the original local (mesolithic) hunter-gatherer population, the other being (neolithic) farmers who came from the Mediterranean area. That is, the Basques have never been homogeneous. The interested reader is referred to the following book:

Cavalli-Sforza, L.Luca, Paolo Menozzi, and Alberto Piazza, 1996, ***The History and Geography of Human Genes***. Abridged paperback edition. Princeton: Princeton University Press

For a long time, it had been necessary to evaluate books and articles written in Spanish to get an in-depth understanding of what was known about or postulated for the history of Basque. Many sources were not easily obtainable. At present the reader can also consult two books in English:

Hualde, Jose Ignacio, Joseba A. Lakarra and R.L. Trask (eds.), 1995, ***Towards a History of the Basque Language***, Amsterdam : J.Benjamins

Trask, Robert L., 1997, ***The History of Basque***, London and New York : Routledge

***Towards a History of the Basque Language*** is a collection of 12 rather independent articles; here I would like to draw your attention especially to „***The Basque Language and Its Neighbors in Antiquity***“, by J.Gorrochategui, and to „***The Ancient Basque Consonants***“, by Luis Michelena. Some other articles may be more controversial. ***The History of Basque*** represents a useful survey, combined with Robert Trask´s position on certain crucial problems. The book contains the following parts: 1 „Introduction“ (1-81), 2 „A thumbnail sketch of the language“ (82-123), 3 „Phonology“ (124-195), 4 „Grammar“ (196-247), 5 „Lexicon“ (248-357), 6 „Connections with other languages“ (358-429).

## II. 1. Basque and Aquitanian

We know several inscriptions from south-western France, pertaining to the time of the Roman Empire, that contain words which can be explained as Basque ones. These inscriptions are found in the territory of the ancient province of Aquitania. The pre-Roman, Basque-looking words of the inscriptions clearly belong to the native language of Aquitania, Aquitanian, which can best be explained as having been an old form of Basque. Relevant data are given in:

Michelena, Luis, 1964, *Textos Arcaicos Vascos*. Biblioteca Vasca VIII. Madrid : Ediciones Minotauro

According to Luis Michelena, the Aquitanian words constitute the most ancient testimony of the Basque language. I quote just one example from Michelena 1964 (p. 16): SEMBETTEN / BIHOSCIN / NIS . F(ilius) Here, SEMBE- is compared with Basque seme „son“, whereas BIHOS- could be explained by Basque bihotz „heart“. The Aquitanian form SEMBE- leads us to a morphological analysis for seme „son“; we can assume an old pair \*sen-be „son“ : ala-ba „daughter“, detaching a suffix -be/-ba.

## II. 2. Basque and Iberian

At the time of Roman conquest, the Iberian peninsula offered a complicated picture linguistically. A great number of inscriptions document the languages of the pre-Roman inhabitants. The inscriptions show that there were at least three, possibly four or five independent language groups in what is now Portugal and Spain. One group is clearly Indo-European, it is called Celtiberian. Another group, mainly in south-eastern Spain and along the Ebro valley, is decidedly non-Indo-European; it is called Iberian. Iberian inscriptions can nowadays be read, but we cannot translate them. The phonological values of the graphic symbols have been determined, but the meaning of the texts remains unknown.

There have been various attempts to use Basque for the analysis and interpretation of Iberian inscriptions, but these attempts have remained totally unsuccessful; it is not possible to translate Iberian with the help of Basque. Even so, researchers have repeatedly tried to compare Iberian fragments of unknown meaning with Basque words in order to prove a „relationship“ between Iberian and Basque. The results are completely unsatisfactory.

Today we still do not know if Iberian is related to Basque. On the other hand, that Iberian is the ancestor of Basque can safely be excluded. This does not mean that there could not be elements of Iberian origin in Basque; some contact between Iberian and the ancient form of Basque seems to be rather likely. To complete this section, I want to mention two Spanish titles on the subject:

Caro Baroja, Julio, 1988, *Sobre la lengua vasca*. Estudios vascos IX. San Sebastian : Editorial Txertoa

Tovar, Antonio, 1959, *El Euskera y sus parientes*. Biblioteca Vasca II. Madrid : Ediciones Minotauro

## II. 3. Basque and language elements in Western Europe without Indo-European (Latin, Celtic, Germanic) explanation

In Western Romance languages, we find many words which neither go back to Latin, nor can be explained by comparing Celtic, Germanic, or Greek for example. Some of these words have a counterpart in Basque. In a couple of cases, a Basque word may simply have been borrowed. In several cases, one could think of another historical development; a certain number of words may have survived from some pre-Indo-European language of Western Europe that is now extinct, and these words may have entered into Romance. If one finds counterparts in Basque, one could feel tempted to see Basque as the remainder of a prehistorical, more wide-spread language group. Is it possible that languages related to Basque were once spoken over a large area of Western Europe? Is there an old layer, a substratum, in languages of Western Europe, which is historically connected with Basque?

Linguists have offered some competing substratum theories to explain the situation in Western Europe. A rather elaborate hypothesis was proposed by the specialist Johannes Hubschmid. Two of his books shall be mentioned here, although they are written in German:

Hubschmid, Johannes, 1953, *Sardische Studien. Das mediterrane Substrat des Sardischen, seine Beziehungen zum Berberischen und Baskischen sowie zum eurafrikanischen und hispano-kaukasischen Substrat der romanischen Sprachen*. Bern : Francke Verlag

Hubschmid, Johannes, 1960, *Mediterrane Substrate mit besonderer Berücksichtigung des Baskischen und der west-östlichen Sprachbeziehungen*. Bern : Francke Verlag

The Basque scholar Luis Michelena has objected to existing substratum theories that every Basque word with an corresponding entity in Romance remains a possible loan from Romance, even if there is no known source for the Romance word in question. We are not able to prove that a Basque word with a Romance match is really an original Basque element, since we do not know the ancestor of Basque.

As linguistic fields of study, we'll have to distinguish sharply between the problem of ancient substrata in modern languages (and the use of Basque in substratum theories) on the one hand - and the search for loanwords from Basque or „Proto-Basque“ in languages of Western Europe (mainly in Western Romance) on the other hand. Such loanwords would have been adopted from the Basques at a later stage and would not constitute a basic layer in the languages concerned. The existence of loanwords from Basque and/or „Proto-Basque“ in languages of the Iberian peninsula is beyond doubt and beyond dispute. As rather clear cases I may give the following two:

(a) Basque *ezker* „left (part), left hand“ explains Spanish *izquierdo* „left“ and *izquierda* „left hand, (political) left“ (see also *esquerre* in Catalan and Gascon)

(b) Basque *txapar* „brush, shrub“ and *sapar* „thicket, underbrush, brambly ground“ point to a Basque or „Proto-Basque“ origin of Spanish *chaparro* „dwarf evergreen oak“ and *chaparral* „oak bush, thicket“ (from Spanish: N.Amer. English *chaparral* „dense tangled brushwood; undergrowth“)

Some examples indicate that one can also posit a Basque etymology for a couple of Western Romance words even if there is an Italian counterpart. Of course, such cases will be somewhat controversial.

But consider the relation between Basque *zango* „leg“ and Spanish *zanca* „leg of a bird, slim leg“, *zanco* „stilt“. In my opinion, it is quite probable that the Spanish words come from Basque, especially for semantic reasons. Now there is an Old Italian counterpart, too: *zanca* „leg, foot“. Can we assume that Old Italian has borrowed the word from the west? I think this to be more likely than the hypothesis traditionally put forward, namely, that the Romance words perpetuate *zanca* „kind of shoe worn by the Parthians (an Iranian tribe)“, occurring as a „foreign word“ in some Latin texts. Of course, the problem cannot be settled at this point.

For a substratum theory to be convincing, linguists will expect that the Basque data are correctly analyzed and constitute reliable proof, and that the impact and contribution of Celtic can be determined sufficiently. Celtic languages were spoken in the main part of Western Europe in pre-Roman times, but the vocabulary of Continental Celtic (Celtiberian, Gaulish) is very imperfectly known. So in many cases we only can speculate if a non-Latin word in Romance is a Celtic one or not. This leads us to the next point, to the problem of Basque-Celtic contacts.

## II. 4. Basque and Celtic

Prior to the romanization of Western Europe, an early Form of Basque and Celtic languages must have been in some contact. We can assume that this time of contact may still be reflected in the vocabulary of Basque (words with Celtic etymology) and perhaps also in the vocabulary of modern Celtic languages (words with Basque or „Proto-Basque“ etymology). Of course, we must be aware that many words coming from Latin appear both in Basque and in modern Celtic languages, and that a number of Celtic words entered Basque via Romance; these words obviously must be discarded in the search for evidence of ancient Basque-Celtic contact.

Actually, we find a set of correspondences between Basque and Celtic suggesting the exchange of words between early Basque (or „Proto-Basque“) and Celtic languages. In those cases where the involved Celtic words have got a good Indo-European etymology, Basque has certainly been the receiving language. In other cases, however, one can imagine that early Basque, „Proto-Basque“ or a language closely related to Basque has been the „donor“ language (the source). I'll give 6 examples of Basque-Celtic comparisons now:

(a) Basque *mendi* „mountain“ : Welsh *mynydd* „mountain“, Cornish *meneth*, Breton *menez* „mountain“. The Celtic words are cognates of Latin *mons*, *montis* „mountain“, so we should consider Basque *mendi* to be a loanword from Celtic.

(b) Basque maite „beloved, dear“ : Irish maith „good“. This etymology has been suggested repeatedly. Nevertheless the relationship remains dubious, since the old Celtic form must have been \*mati-, which should have given \*madi in Basque, but not maite. Irish maith < \*mati- is said to be cognate with Latin maurus „ripe“.

(c) Basque harri < \*karri „rock, stone“ : Irish carrac „rock, large stone“, Welsh carreg „stone“. The Celtic words probably stem from a Pre- Indo-European source, which could have been a relative of Basque. One cannot explain the words with the help of Indo-European.

(d) Basque adar „branch“, „horn“ : Irish adarc „horn“. The Irish word is not an Indo-European formation; it is obviously a non-Indo-European element which has replaced the old Indo-European term. In Breton we still find karn „horn“, cognate with Latin cornu.

(e) Basque andere, andre „woman, lady“ : Irish ainder „woman, young lady“. Irish ainder has not got an Indo-European explanation either. Basque andere is already documented in Aquitanian inscriptions. Probably Celtic has borrowed the word.

(f) Basque izokin „salmon“ : Irish eo „salmon“, Welsh eog „salmon“, Breton eog „salmon“. For Celtic, we can reconstruct an original \*esok-. Spanish yields esguin „young salmon“, perhaps from Celtiberian. The Celtic word is documented in Latin sources as esox, esocis „a fish of the Rhine“. Celtic \*esok- is another stem which lacks a plausible Indo-European interpretation and analysis. On the other hand, Basque izokin could be analyzed as \*iz-o-ki- : \*iz- appears in some Basque words as a first element with the meaning of „water“; the Basque-French dictionary by Pierre Lhande documents itze as a word for „sea“ (can we reconstruct Basque itsaso „sea“ as \*itz-a-so?). Basque -ki is a suffix with the meaning of "food (of)", compare basa-ki „wild animal meat“ from baso „forest“. For -o-, one could think of the -o- in adiskide-o-k „my/our friends“ beside adiskide-a-k „the friends“. Is, then, Basque izokin simply \*„food of our water“? This would mean that already Common Celtic possessed words derivable from „Proto-Basque“, a rather surprising possibility.

There are some cases where Basque and Celtic words seem to belong together, but where sound laws appear to forbid a direct connection. The Basque word hartz „bear“ resembles Irish art „bear“ and Welsh arth „bear“ (Proto-Celtic probably \*arkto-s). The Celtic words are of Indo-European origin and cognate with Classical Greek arktos, Latin ursus, and Sanskrit rksa- „bear“. We would expect \*arto or \*arta in Basque, if the Celtic word were represented here. There is no transition -t- > -tz- in Basque. If Basque hartz is not a Celtic loanword, should we then assume (Pre-Roman, Pre-Latin) Pre-Celtic Basque-Indo-European contact?

## II. 5. Basque and Indo-European

In Basque, we find a number of words which cannot be traced back to Latin and which cannot be explained with the help of Celtic, but which still remind one of well-known Indo-European word-families. Perhaps some (or all) of these cases point to prehistoric relations between „Proto-Basque“ and early representatives of Indo-European. I would like to mention three examples:

(a) Basque urki, burkhi „birch tree“ : Sanskrit bhu:rja- „kind of birch tree“, Icelandic björk and German Birke „birch tree“, Lithuanian berzhas and Russian berjoza „birch tree“. In Georgian (South-Caucasian) the word for „birch tree“ is arqi, also resembling Basque urki.

(b) Basque arrano „eagle“ : Hittite hara-s, haran-as „eagle“, Class.Greek ornis and orneon „bird“, Gothic ara „eagle“, Lithuanian erelis „eagle“. There may be a connection between Basque arrano „eagle“ and Basque arraultza, arrautza „egg“; this would indicate that arrano once meant „bird“. The Indo-European words surely belong to an old lexical layer, since they contain the suffix \*(e/o)n-, often found in Indo-European words for animals.

(c) Basque zilar „silver“ (zidar in Biscayan) : Gothic silubr and German Silber „silver“, Lithuanian sidabras „silver“, Russian serebro „silver“. The sound-correspondences in Indo-European are irregular here; the alternation Germanic -l- / Baltic -d- / Slavic -r- is completely unusual. Probably this points to an ancient borrowing.

Furthermore, we can state some interesting lexical parallels between Basque and Classical Greek. There are Greek words without an explanation within the frame of Indo-European that resemble Basque words; we come across cases like the following two:

(a) Basque apar „foam“ : Greek aphro-s „foam“

(b) Basque zitu „fruit (of a harvest)“ : Greek sito-s „corn, wheat“

Of course, one could think of chance resemblances. It remains to be determined if the quality and quantity of such parallels makes them something to take serious. In any case, isolated similarities between Basque words and some Indo-European ones (as the pair Basque mutur „snout, muzzle“ : Lithuanian mute and muturze „mouth, muzzle“) will probably forever escape judgement.

## II. 6. Basque and „Caucasian“

The comparison of Basque with the various Caucasian languages looks back to a long history. The eminent Basque scholar Luis Michelena remained sceptical about the chance to prove a Basque-„Caucasian“ relationship. The Basque linguist and author Jose Luis Alvarez Enparantza („Txillardegi“) has taken a rather optimistic standpoint in this question. I expect that the discussion will continue. My own impression is that only a extremely remote connection is possible; one would have to fix a supposed relationship for paleolithic times, i.e. at least ten thousand years ago. This means that the traditional methods of historical linguistics cannot apply.

Remember there is no single language family „Caucasian“, but that we have to distinguish sharply between the South Caucasian language family on the one hand and the North Caucasian family on the other hand. If one wants to link Basque with the area of the Caucasus, he/she must consider that a relationship between North Caucasian and South Caucasian ist not proven, so Basque should be compared with North Caucasian and with South Caucasian separately. I would like to add that a possible genetic connection between Basque and one or both of the Caucasian language families presupposes not only a basic common lexical stock, but also an ancient set of common grammatical elements (morphemes). So a researcher would not only need to discover a sufficient number of word parallels, but also morphological evidence.

Most of the comparisons between Basque words and words from Caucasian languages that have been proposed over the years will surely have to be rejected. Some connections, however, seem to be interesting. I mention four cases of resemblances in the lexicon between Basque and North Caucasian languages first:

(a) Basque argi „light“, „clear“ : Lak b-argh „sun“, Avar b-aq'q' „sun“

(b) Basque su „fire“ : Lak c'u „fire“, Avar c'a „fire“

(c) Basque baso „forest“ : Lak wac'a „forest“

(d) Basque otso „wolf“ : Lak b-arc' „wolf“, Avar b-ac' „wolf“

Please note that the symbol c in the transcription of North Caucasian represents ts as in English rats.

Let us now have a brief look at four possible matches between Basque and Georgian (South Caucasian). The first three etymologies were set up by the linguist Karl Bouda, who wrote amply about Basque and „Caucasian“.

(a) Basque gau „night“ : Georgian gham-e „night“

(b) Basque sagu „mouse“ : Georgian tagv-i „mouse“

(c) Basque so „look, glance“ : Georgian tval-i „eye“

(d) Basque mihi „tongue“ < \*b-ini : Georgian en-a „tongue“, „language“

Some of those interested in the origin of Basque may wish to study facts from Caucasian for themselves. I recommend especially one historical treatise concerning West Caucasian (Chiribka 1996), the volumes of *The Indigenous Languages of the Caucasus*, and four monographic descriptive grammars, three for North Caucasian languages (Kabardian, Godoberi, Lezgian), one for Georgian:

Chiribka, Viacheslav A., 1996, *Common West Caucasian. The Reconstruction of its Phonological System and Parts of its Lexicon and Morphology*. Leiden, The Netherlands : Research School CNWS.

Harris, Alice C. (ed.), 1991, *The Indigenous Languages of the Caucasus*, Volume 1: The Kartvelian Languages. Delmar, New York : Caravan Books

Hewitt, George (ed.), 1989, *The Indigenous Languages of the Caucasus*, Volume 2: The North West Caucasian Languages. Delmar : Caravan Books

Smeets, Rieks (ed.), 1994, *The Indigenous Languages of the Caucasus*, Volume 4: North East Caucasian Languages, Part 2. Delmar : Caravan Books

Colarusso, John, 1992, *The Kabardian Language*. Calgary, Canada : University of Calgary Press

Kibrik, Alexandr E. (ed.), 1996, Godoberi. *Lincom Studies in Caucasian Linguistics 02*. München/Newcastle : Lincom Europa

Haspelmath, Martin, 1993, *A Grammar of Lezgian*. Mouton Grammar Library 9. Berlin/New York : Mouton de Gruyter

Hewitt, B.G., 1995, Georgian. *A structural reference grammar*. Amsterdam : John Benjamins

We will now turn to the field of Basque historical grammar. What do we know about the development of sounds and of formal elements (morphemes) in Basque?

### III. Internal features of Basque language history

I'll divide this account into three parts. First we will deal with the history of Basque sounds. The description of the history of sounds is the task of historical phonology. Then we will treat the history of bound meaningful elements (meaningful elements that are part of a word) in Basque. Here we are in the field of historical morphology. Finally the historical evaluation of classical texts in Basque will be introduced as a promising object of study; one may discover trends of change in inflectional verb morphology and in syntax here.

#### III. 1. On Basque historical phonology

Where do the informations we have got about the history of Basque sounds come from? There are two main sources of information: First, we can study the development of words borrowed from Latin in Basque. Many of these words have changed over time; one can observe certain regular changes in sound shape. Secondly, we can compare the existing Basque dialects. The differences in sound shape between corresponding words in the individual Basque dialects may eventually be explained historically. We distinguish seven main dialects within Basque: Biscayan (bizkaiera), Guipuzcoan (gipuzkera), High Navarrese (nafarrera garaia), Low Navarrese (nafarrera beherea), Labourdin (lapurtera), Souletin (zuberoera), and Roncalese (erronkariera). Roncalese has disappeared recently as a spoken language.

Basic work on Basque historical phonology we owe to Luis Michelena. His *Fonética Histórica Vasca* should be read by everyone who intends to do historical research on Basque. Many of Michelena's articles have been made available again in *Sobre historia de la lengua vasca*.

Michelena, Luis, 1990 (first edition 1961), *Fonética Histórica Vasca*. New edition. Donostia : Publicaciones del Seminario de Filología Vasca „Julio de Urquijo“

Michelena, Luis, 1988, *Sobre historia de la lengua vasca*. Edited by Joseba A.Lakarra.. Anejos ASJU 10. Donostia : Seminario de Filología Vasca „Julio de Urquijo“

We know now that Basque originally possessed two n-phonemes (/n/ and /N/ or /nn/) and two l-phonemes (/l/ and /L/ or /ll/); these pairs were parallel to the opposition between /r/ and /rr/ (/R/) still existing in Basque today. An original /n/ got lost between vowels, an original /l/ developed into /r/ in this position. These changes are reflected clearly in old loanwords from Latin:

Basque ahate „duck“ < Latin anas, anatis, acc. anate-m „duck“

Basque ohore „honor“ < Latin honor, honoris, acc. honore-m „honor“ Basque harea „sand“ < Latin (h)arena „sand, sandy place“

Basque soro „field,cultivated land“ < Latin solum „ground, soil“

Basque borondate „will“ < Latin voluntas, voluntatis, acc. voluntate-m „will“

Basque aingeru „angel“ < Latin angelus, acc. angelu-m „angel“

The ancient phonemes /N/ and /L/ became /n/ and /l/ respectively between vowels. We find Latin -nn- and -ll- represented by -n- and -l- in Modern Basque:

Basque ano „portion, food supply“ < Latin annona „food, supplies“

Basque gela „room“ < Latin cella „(small) room“

Basque makila „stick“ < Latin bacillum „small stick“

In native Basque words, an original /n/ between vowels can be reconstructed for a number of cases. An /n/ between vowels in native words of modern Basque cannot represent an original /n/, it must go back to /N/ (/nn/):

(a) Basque mihi, Biscayan min, High Navarrese and Roncalese mi (-i is nasalized in Roncalese!) „tongue“ < Old Basque \*bini

(b) Basque zain, Guipuzcoan, High Navarrese and Roncalese zai (-i again nasalized in Roncalese) „keeper, custodian“ < Old Basque \*zani

(c) Basque anaia, anai „brother“ < Old Basque \*annai(-a) The opposition between the Old Basque phonemes /b/, /d/, /g/ and the

Old Basque phonemes /p/, /t/, /k/ obviously did not correspond fully to the opposition between Latin /b/, /d/, /g/ and Latin /p/, /t/, /k/. It is striking that in most ancient Basque loanwords from Latin the voiced stops /b/, /d/, /g/ replace the Latin initial voiceless stops /p/, /t/, /k/:

Basque bake „peace“ < Latin pax, pacis, acc. pace-m „peace“

Basque denbora „time“ < Latin tempora „times“

Basque gorputz „body“ < Latin corpus „body“

We will now turn to the phonological variation between modern Basque dialects. An important source on dialectal variation is Azkue's Basque-Spanish-French dictionary: Azkue, Resurreccion Maria de, 1905, **Diccionario vasco-español-frances**. Bilbao. This dictionary is currently available as a reprint by the „Academy of the Basque Language“, Euskaltzaindia.

In some cases, Biscayan /a/ replaces an /e/ of the other dialects:

Basque berri „new“ : Biscayan barri „new“

Basque (h)ezur „bone“ : Biscayan azur „bone“ (Since /e/ is nasalised in Roncalese, we may perhaps reconstruct an original \*enazur)

Sometimes we find Western Basque /u/ beside Eastern Basque /i/:

Western Basque uri „town“ : Eastern Basque (h)iri „town“

Western Basque zulo „hole“ : Eastern Basque zilo „hole“

The word for „ear“ is belarri in Western Basque and be(h)arri in Eastern Basque. In my opinion, one could think of \*benarri as being the underlying form. Luis Michelena has proposed \*berarri.

Biscayan narru „hide, skin, leather“ contrasts with larru of the other Basque dialects. I suppose narru to represent the old sound shape. In other cases, Biscayan seems to show the more recent sound shape: Biscayan edur, erur „snow“ : remaining Basque dialects elur „snow“ Biscayan bedar, berar „grass“ : remaining Basque dialects belar „grass“

I want to mention one interesting alternation more; „to say“ is expressed either by erran or by esan in Basque dialects. esan is more at home in Western Basque, erran is more typical of the eastern dialects. Obviously, /rr/ (/R/) can alternate with /s/. The direction of change cannot be determined definitively. s > r is more common in Indo-European languages, but in Turkic we find z < \*r, for example. In any case, the existence of an alternation /rr/ : /s/ allows us to compare the roots of Basque j-os-i „to sew“ and of Basque orr-atz „needle, pin“. j-/i-/e- is an old Basque verbal prefix, -atz must be a noun suffix, so we can isolate \*-os-/\*orr- „to sew“.

### III. 2. On Basque historical morphology

Those who wish to get information concerning the productive or semi-productive bound meaningful elements (morphemes) of Basque may consult for example the following work:

"Haize Garbia", 1978-1981, *Grammaire basque pour tous*. I-III. Hendaye : Editions „Haize Garbia“

Here we will look at some of those bound morphemes of Basque which are no longer productive, but which can be ascertained either by comparison of Basque dialects or by internal reconstruction.

A basic distinction can be made between elements forming verb stems and elements creating noun stems. I'll start with the Basque verb.

Many Basque verbs contain an old prefix -ra-. We observe pairs like the following ones:

i-bil-i „to walk, to move“ : e-ra-bil-i „to manage, to use; to move something“

i-kas-i „to learn“ : i-ra-kats-i „to teach“

i-kus-i „to see, to look“ : e-ra-kuts-i „to show“

e-ntzu-n „to listen, to hear“ : e-ra-ntzu-n „to answer, to reply“

Not that obvious is the derivation of Basque e-ra-ma-n „to carry, to transport“. According to Luis Michelena, e-ra-ma-n does not belong to e-ma-n „to give“, as one might think, but to j-oa-n „to go“, so we have to posit a development \*e-ra-oa-n > \*e-ra-ba-n > e-ra-ma-n. We find e-r-oa-n „to take, to carry, to transport“ in Biscayan.

-ra- is the only prefix that was used in Old Basque to form new verbs; beside -ra-, we only know of suffixes as former means to create new verbs. Two of such suffixes must have been -ki and -o, as far as we can judge from pairs of verbs like these:

Basque j-ai-ki „to get up, to rise“ : Basque j-ai-o „to be born“; Guipuzcoan i-ze-ki „to burn, to burn up“ : Biscayan i-zi-o „to light, to kindle“; Basque j-arrai-tu „to follow“ : Low Navarrese and Souletin j-arrai-ki „to follow“ (j-arra(-)i is used as a noun for „continuation“ in modern Basque); Classical Eastern Basque j-ar(-)i „to spring, to gush“ : Basque j-ar(-)i-o „to spill, to drip, to ooze, to flow“.

There can be isolated by far more old noun forming elements (suffixes) than former verbal morphemes. I'll mention just some interesting cases.

We surely are right in comparing Basque behi „cow“ with Basque behor „mare“; this gives us a root beh- and the two suffixes -i and -or. In fact, words ending in -i are very frequent in Basque, and we may consider -i to be an old semantically unmarked formant. The noun forming suffix -or might reoccur in alor „cornfield“ beside ale „grain, seed“, or perhaps in sa-t-or „mole“ (sa- here comes from sagu „mouse“ obviously).

Well established is the ancient morpheme -ba forming kinship terms, see ala-ba „daughter“, ilo-ba „nephew, niece“, osa-ba „uncle“, ize-ba „aunt“. In my opinion, it is not definit that ugaz-aba „employer, boss, master“ belongs here, too; this may be an old compound with j-abe (j-aba in Old Biscayan) „owner, proprietor“ as a second part. In any case do neither osa-ba nor ugaz-aba prove an ancient Basque word \*aba „father“ („father“ is aita in all dialects of Basque), as Azkue has assumed. Since an alleged \*aba „father“ obviously must be discarded as a real Basque word, the neologisms ab-erri „fatherland“ and ab-er-tzale „patriot“ are questionable formations from a linguistic standpoint.

I would like to suggest a comparison between es-ne „milk“ and e-ra-its-i < \*e-ra-eits-i „to milk“. This leaves us with a root \*(-)eits- and a noun forming suffix -ne. I have looked for other examples of an element -ne in modern Basque and I have found at least one, I believe: zor-ne „pus, infection“ should be joined etymologically with zaur-i „wound“ in my opinion. Traditionally Basque zor-ne is compared with Spanish sarna „itch, scabies, mange“, so one could ask if Spanish sarna is a very early loanword from Basque. Serna and sarna „mange“ occur als „Iberian“ words in Latin texts; they may well represent the Basque word.

Finally I want to give three examples of how dialect differences within Basque offer insights into Basque historical morphology.

(a) Bizcayan ira „fern“ requires an analysis of Eastern Basque iratze, iratz „fern“ as ira-tz(e); this means that there exists no relationship with Irish raith „fern“ as could otherwise be suspected.

(b) Western Basque ego „wing“ demands that Eastern Basque (h)egal „wing“ must be segmented as (h)ega-l; -l can then be connected with -le as in egi-le „author“ or ikas-le „student, pupil“. It is surely incorrect to compare (h)egal with words like Chukchee galga „bird“, as it was tried.

(c) From Western Basque orr-i „leaf“ cannot be separated Basque (h)osto „leaf“. The Basque author Oihenart still shows or-sto, so we get a suffix -sto. Speakers of English now have at their disposal a valuable tool for the study of historical and synchronic features of the Basque lexicon: Aulestia, Gorka, 1989, *Basque-English dictionary*, Reno and Las Vegas : University of Nevada Press

### III. 3. Materials for the study of language change: Basque classical texts in comparison with Modern Basque

The Basque language is documented by lengthy texts since the 16th century. Of course, changes can be observed between classical Basque literature and the modern varieties of the language. The development of the inflectional verb morphology may be of interest, and especially important are syntactical changes.

French linguist Rene Lafon contributed a detailed account of the Basque verbal system of the 16th century. As an example of syntactical research I may mention an interesting contribution by Victor Hidalgo Eizagirre written in Basque.

Lafon, Rene, 1944 / 1980, *Le systeme du verbe basque au XVIe siecle*. Donostia/Baiona : Elkar

Hidalgo Eizagirre, Victor, 1995, „*Ohar estatistiko garrantzitsuak euskararen hitz ordenaren inguru*. Euskara, S.V.O.?” , *Fontes Linguae Vasconum* 70 p. 401-420

Those who want to get an overview of Basque classical literature may read either Villasante's or Michelena's book on the subject:

Villasante, Fr. Luis, 1979, *Historia de la literatura vasca*. 2.a edicion. Aranzazu : Editorial Aranzazu

Michelena, Luis, 1988, *Historia de la literatura vasca*. 2.a edicion. Donostia : erein

Many of Basque's classics are available again as reprints. These offer the chance to explore the peculiarities of important Basque texts for a wide readership. The first book which was printed in Euskara was *Linguae Vasconum Primitiae*, by Bernard Etxepare. This is now re-edited by Euskaltzaindia. Etxepare, Bernard, *Linguae Vasconum Primitiae*. Edizio kritikoa, Patxi Altunak paratua. Euskararen Lekukoak 2. Bilbao : Euskaltzaindia (1987)

o this we have to add the work of Ioannes Leizarraga, which consists of a translation of the New Testament, a calendar, and a primer or catechism. All were published in La Rochelle in 1571. Again Euskaltzaindia offers us a reprint.

Leizarraga, Ioannes, *Iesus Christ Gure Iaunaren Testamentu Berria*. Kalendrera. ABC edo Christinoen instrukzioea. Bilbao : Euskaltzaindia (1990)

Another fundamental work to be edited by Euskaltzaindia has been the old collection of Basque proverbs and sayings from 1596 (author unknown): *Refranes y sentencias* (1596). Ikerketak eta edizioa. Joseba Lakarra Andrinua. Bilbao 1996 : Euskaltzaindia

A celebrated Basque author is Pedro de Axular (1556-1644). He wrote the famous book *Gero*, the best known of all Basque classics. There are two re-editions:

Axular, Pedro, *Gero* (despues). Argitalpena: Aita Luis Villasante, O.F.M. Arantzazu : JAKIN (1976)

Axular, Pedro de, *Guero*. Bilbao 1988 : Euskaltzaindia

I have to mention now Arnaud d'Oihenart (1592-1667). He published a book in Basque in the year of 1657. This work also is available again: d'Oihenart, Arnaud, *Proverbes et poesies basques* (1657-1664). Edition trilingue (basque, francais, espagnol) presentee par Jean-Baptiste Orpustan. Baigorri : Editions Izpegi (1992)

An important Basque author is then Juan Antonio de Moguel (1745-1804). He finished his work *Peru Abarka* in 1802; the book was first published in 1881. I give the re-edition: de Moguel, Juan Antonio, *Peru Abarka*. Octava edicion bilingüe. Bilbao : Editorial La Gran Enciclopedia Vasca (1981)

Finally I want to call attention to the classics of Jean Pierre Duvoisin (1810- 1891). He was the first to translate the whole Bible into Basque. Well-known is also his *Laborantzako liburua* (1858). Both publications represent a very elaborated form of Basque.

Duvoisin, Jean, *Bible Saindua, edo Testament Zahar eta Berria*. I-III. Bilbao : Editorial La Gran Enciclopedia Vasca (1972)

Duvoisin, Jean Pierre, *Laborantzako liburua*. Klasikoak 3. Euskal Editoreen Elkarte. Donostia : Kriselu (1978)

These are some basic examples of Basque classical literature. The language of the mentioned books may be compared with Modern Standard Basque (Euskara Batua) on the one hand and with modern Basque dialects on the other hand. We still are in need of detailed descriptions for modern dialects; it is hoped that there will be extensive research on Basque dialects in the near future. This research could follow the lines of the study presented by Jose Ignacio Hualde, Gorka Elordieta and Arantzazu Elordieta on the dialect of Lekeitio. I recommend this study to all those interested in modern spoken Basque:

Hualde, Jose Ignacio, Gorka Elordieta and Arantzazu Elordieta, 1994, *The Basque dialect of Lekeitio*. Supplements of Anuario del Seminario de Filologia Vasca „Julio de Urquijo“, XXXIV. Bilbo: Euskal Herriko Unibertsitatea and Donostia : Gipuzkoako Foru Aldundia.

Source: <http://www.ehj-navarre.org/blessons/mowstr.html> (3/22/2006)