Basque in Western Europe: some arguments for a Vasconic substratum

Rebecca Tollan

0. Abstract
This paper discusses the possibility of a Vasconic substratum in Western Europe. In the past, many historical linguists have attempted to prove that Basque is related to other languages, or families of languages, largely without success. I do not wish to infer that Basque has any surviving relatives, or that it is genetically linked to Indo-European in any way. Instead, this discussion is intended to instigate debate regarding the degree of influence, if any, of Basque on Indo-European peoples and languages due to contact as Indo-European spread to Western Europe. I begin with a brief overview of some structural and lexical features of Basque, before mentioning some of the links between Basque and other languages that have previously been suggested. This will lead to discussion of some of the evidence in favour of a Vasconic substratum in Western Europe, particularly that which is put forward by German linguist Theo Vennemann. Eventually, I shall review the data and discuss how it may suggest that Basque was once spoken over a larger stretch of Europe than it is today.

1. Introduction
The Basque language (native name euskara) is spoken in the Basque region of northern Spain and south-western France. This region, known as Basque country or Euskal Herria runs along the Bay of Biscay at the western end of the Pyrenees. Although geographically surrounded by Indo-European languages, Basque is considered a language isolate and is widely believed to be the only surviving pre-Indo-European language in Western Europe.

2. The structure of Basque

2.1 Noun morphology
Basque morphology is strongly agglutinating. Articles, demonstratives most quantifiers must follow the noun they determine, and all noun phrases must contain a determiner. This is illustrated in (1) with the NP ‘kale’ (street):

1. **kalea** – ‘the street’ (street-DET)
   - **kaleak** – ‘the street’ (street-DET-PL)
   - **kale hau** – ‘this street’ (street this)
   - **kale hauek** – ‘these streets’ (street this-PL)
   - **kale bat** – ‘one street (street one)’

(King, 1994:22)

However, where a noun precedes an adjective, it is the adjective that must take the article or determiner. See (2):

2. **kale txikia** – the small street (street small-DET)
   - **kale txikiak** – the small streets (street small-DET-PL)
   - **kale txiki bat** – one small street (street small one)

2.2 Case

Basque has a rich and complex case system, consisting of the case suffixes in (3):

3. Dative (‘to’, ‘for’, ‘from’): -i
   - Genitive (possessors): -en
   - Locative (place of rest ‘in’, ‘at’, ‘on’): -n
   - Allative (goal of motion ‘to’): -ra
   - Ablative (source of motion ‘from’): -tik
   - Instrumental (‘by, ‘of’, ‘about’): -z
   - Comitative (‘with’): -ekin
   - Benefactive (‘for’(person)): -entzat
   - Destinative (‘for’(thing)): -rako
   - Directional (direction of motion ‘towards’): -rantz
   - Terminative (‘as far as’): -raino

(Trask, buber.net)
One of the most renowned features of Basque is the use of ergative-absolutive case marking. The ergative case system is not a feature of Indo-European languages, although it can be found in a handful of other languages such as Georgian, Inuktitut, and various indigenous languages of Canada and Australia. Ergative-absolutive languages distinguish between transitive and intransitive verbs with regards to the case marking of the subject. The subject of a transitive or an unergative verb is marked with ergative case (suffix –k in Basque), whilst the subject of an unaccusative verb and the direct object of a transitive verb are in absolutive case (in Basque this is unmarked). In other words, the agent receives ergative case whilst absolutive case is assigned to the theme. This is illustrated in (4)

4. Edurnek atea irekitzen du
   Edurne-ERG door-ABS open AUX
   ‘Edurne opens the door’          (King, 1994:136)

The agent (and grammatical subject) ‘Edurne’ has case marked as ergative whilst the theme (and grammatical object) ‘atea’ is in absolutive case, which has null phonological realisation. When ‘open’ is used as an unaccusative verb, the sole argument receives absolutive case, as in (5)

5. Atea irekitzen da
   door-ABS open AUX
   ‘The door opens’                   (King, 1994:136)

Similarly, the sole argument of an unergative verb is marked ergative. See (6)

6. Piratak igeri egiten du
   Pirate-ERG swim make has
   ‘The pirate swims’                (Santesteban, Pickering, Branigan, 2010)

Levin (1983, discussed by Cheng and Demirdache, 1993:75) proposes that unaccusative verbs in Basque are unlike unaccusative verbs in other languages, in that they are able to assign (absolutive) case to their object. Following this, one could assume that finite T in Basque assigns
ergative case, although this assumption would rely on an analysis of case on T being interpretable (since ergative case is not assigned in unaccusative constructions).

2.3 Sentence structure

Basque is generally considered a head final language. Word order is typically SOV (as in (4)), although it may vary depending on different information structure. Basque sentences have topic-focus word order, known as *galdegai*, which stipulates that the constituent in focus must immediately precede the verb.

7. Jonek Kepa jo zuen  
   John-ERG Peter-ABS hit AUX  
   ‘John hit Peter’

Kepa Jonek jo zuen  
   Peter-ABS John-ERG hit AUX  
   ‘It was John who hit Peter’

Kepa jo zuen Jonek  
   Peter-ABS hit AUX John-ERG  
   ‘It was Peter who John hit’

Jo zuen Jonek Kepa  
   hit AUX John-ERG Peter-ABS  
   ‘John *hit* Peter’  
   (Trask, 1997: 109)

2.4 The verb form

Basque verbal morphology is mostly periphrastic (Trask, 1997: 103); i.e. verb forms consist of a non-finite lexical verb and a finite auxiliary verb which carries tense and agreement markers. Only a handful of verbs have synthetic forms.
Two different auxiliary verbs are used in Basque: unaccusative verbs are conjugated with izan (‘to be’), whilst transitive and unergative verbs are conjugated with *edun ‘to have’. The auxiliary must agree in number, person and gender with the subject as well as in number any object present. Consequently *edun has both a singular and a plural form:

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have</td>
<td>(nik) dut</td>
<td>(nik) ditut</td>
</tr>
<tr>
<td>You have (close male friend)</td>
<td>(hik) duk</td>
<td>(hik) dituk</td>
</tr>
<tr>
<td>You have (close female friend)</td>
<td>(hik) dun</td>
<td>(hik) ditun</td>
</tr>
<tr>
<td>You have (formal)</td>
<td>(zuk) duzu</td>
<td>(zuk) dituzu</td>
</tr>
<tr>
<td>he/she has</td>
<td>(hark) du</td>
<td>(hark) ditu</td>
</tr>
<tr>
<td>we have</td>
<td>(guk) dugu</td>
<td>(guk) ditugu</td>
</tr>
<tr>
<td>you have (plural)</td>
<td>(zuek) duzue</td>
<td>(zuek) dituzue</td>
</tr>
<tr>
<td>they have</td>
<td>(haiek) dute</td>
<td>(haiek) dituzte</td>
</tr>
</tbody>
</table>

(Trask, buber.net)

Basque is pro-drop; personal pronouns are usually omitted except for emphasis. Furthermore, there are two forms of the informal second person singular depending on whether the addressee of the sentence is male or female. The masculine form ends with –k whilst the feminine form ends with –n. This is another property which is not found in Indo-European languages.

2.5 Negation

To negate sentences, the negative quantifier ez is placed directly before the auxiliary verb and immediately after the topic. The remaining elements of the sentence then follow. See (8).

8. Euskalduna naiz * Ez naiz euskalduna
    basque (I) am not (I) am Basque
    “I am Basque” “I am not Basque” (King, 1994: 14-15)

2.6 Postpositions

1 *edun has lost its non-finite form.
As is typically the case in SOV languages, Basque has postpositions as opposed to prepositions. These postpositions often take one or more case suffixes, as shown in (9) and (10).

(9) haitzaren gainean
    rock-GEN on top-LOC
    “On top of the rock”

(10) harriaren aurrean
    stone-GEN in front of-LOC
    “In front of the stone”

2.7 Lexicon

Basque has thousands of words which are borrowed from Romance (especially French and Spanish), since it is in intense contact with these languages. There are also numerous early loanwords from Latin, for example ‘mila’ (one thousand), ‘liburu’ (book) and ‘diru’ (money) (Trask, buber.net). Linguists have also identified a small number of words believed to be early Celtic borrowings; these include ‘maite’ (beloved) and ‘adar’ (horn) (Trask, buber.net). Despite many foreign loanwords, however, the majority of the Basque lexicon consists of indigenous words. These are typically words for family members, animals, parts of the body, toponyms and words for natural materials such as water (‘ur’), iron (‘burdina’) and earth (‘lur’) (Trask, buber.net). It is also noteworthy that words for tools such as ‘axe’ (‘aizkora’ or ‘haizkora’) have the same root as the words for ‘rock’ (‘aitz’ or ‘haitz’), suggesting that they may have originated in the Stone Age, when stone was used to make tools. Some other words, which do not have direct translations, also have connotations of a pre-historic lifestyle. Such an example is the noun ‘ilini’ which refers to ‘burning wood’ or ‘a half-burned stick’. It is likely that this word was also first coined many thousands of years ago, when firebrand was essential for human survival, suggesting that the Basque language can be traced back several millennia.

2.8 Counting system

Basque uses a vigesimal counting system, i.e. base twenty, rather than base ten, is used with respect to the linguistic structure of almost all ordinal numbers, with the exception of numbers below twenty and multiples of one hundred.
Basque multiples of ten:

Hamar - ten
Hogei - twenty
Hogeita hamar – thirty (literally ‘twenty and ten’)
Berrogei – forty (literally ‘two times twenty’)
Berrogeita hamar – fifty (literally ‘two times twenty and ten’)
Hirurogei – sixty (literally ‘three times twenty’)
Hirurogeita hamar – seventy (literally ‘three times twenty and ten’)
Laurogei – eighty (literally ‘four times twenty’)
Laurogeita hamar – ninety (literally ‘four times twenty and ten’)
Ehun – one hundred

(Trask, buber.net)

3. Possible Origins of Basque and connections with other languages

It is thought that an earlier form of Basque was present in Western Europe prior to the arrival of the Indo-Europeans. It seems fairly certain that Basque is descended from, or at least closely related to Aquitainian, the language spoken in ancient Aquitania in South-western France. Aquitanian is attested in the Roman period in the form of about 400 personal names and 70 divine names. A small selection of Aquitanian words is shown in the table in (11).

(11) Aquitainian vocabulary

<table>
<thead>
<tr>
<th>Aquitainian</th>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>arixo</td>
<td>haritz</td>
<td>oak</td>
</tr>
<tr>
<td>atta</td>
<td>aita</td>
<td>father</td>
</tr>
<tr>
<td>andere</td>
<td>and(e)re</td>
<td>woman</td>
</tr>
<tr>
<td>cisson</td>
<td>gizon</td>
<td>man</td>
</tr>
<tr>
<td>gori, cor(r)i</td>
<td>gorri</td>
<td>red</td>
</tr>
<tr>
<td>hars</td>
<td>hartz</td>
<td>bear</td>
</tr>
<tr>
<td>heraus</td>
<td>herauts</td>
<td>boar</td>
</tr>
<tr>
<td>ilun(n)</td>
<td>ilun</td>
<td>dark</td>
</tr>
<tr>
<td>osson, oxon</td>
<td>otso</td>
<td>wolf</td>
</tr>
<tr>
<td>sahar</td>
<td>zahar</td>
<td>old</td>
</tr>
</tbody>
</table>

(Trask, buber.net)
Aquitainian is accepted by most historical linguists as a precursor of modern Basque. Over the years, linguists have made many attempts to link Basque with other languages, although as yet there is no firm evidence to suggest that it is related to any language besides Aquitanian. Some of these suggested connections are briefly discussed in sections (3.1-3.3)

### 3.1 Iberian
The ancient Iberian language was spoken in the Iberian Peninsula from approximately the 7th – 1st centuries BC. Various lexical items in Iberian show similarities with Aquitanian and Basque; however most linguists accept that these similarities are most likely a result of lexical borrowing and conclude that there is not sufficient evidence for a genetic relationship between Iberian and Basque.

### 3.2 Pictish
Spoken in Scotland by the Picts until the early Middle Ages, the Pictish language is something of a mystery. Not a single sentence of Pictish has survived in a written document, and thus the only sources for the language are inscriptions on carved stones (dating from approximately the 8th century CE) and a small number of place names. Nothing is known about Pictish grammar. The language is labelled ‘Pictish’ from the Latin ‘picti’ (meaning painted) which the Romans applied to the unidentifiable inhabitants of present-day Scotland. Some linguists believe that Pictish was a non-Indo-European language spoken in northern Britain prior to the arrival of the Celts, and that this language may have been related to Basque, being non-Indo-European also. However, these claims are purely speculative and there is not a shred of linguistic evidence to suggest that Pictish and Basque were related. Most linguists now regard Pictish as Celtic, but belonging to the P-Celtic branch along with Welsh and Breton as opposed to the Q-Celtic branch to which Irish and Scots Gaelic belong.

### 3.3 Caucasian
There have been attempts to link Basque to Georgian/South Caucasian. The main similarity between Georgian and Basque is the common use of ergative-absolutive case marking. However, this evidence alone is not enough to suggest that the languages are genetically related. Moreover, there are a significant number of other world languages which use this case system (see section (2.2)), suggesting that this similarity between Basque and Georgian could simply be chance resemblance. The second striking similarity is the vigesimal counting system of both Basque and...
Georgian. The word for ‘twenty’ (Basque hoge; Georgian otsi) is used in counting as a base for all cardinal numbers between twenty and one hundred.

4. The Old European Substrate theory

Europe was inhabited for at least 30,000 years before the arrival of the Indo-Europeans from the east, and the Indo-European languages must therefore have obliterated many other languages spoken on the continent. Whilst there is no direct evidence of such languages, some linguists propose that traces of their vocabulary and structure can be found in Indo-European languages. Since Basque is the sole survivor of these pre-Indo-European tongues, several linguists have searched for possible traces of Basque syntax and vocabulary in modern Indo-European languages. Theo Vennemann proposes the ‘Vasconic substratum hypothesis’ based largely on studies of European toponyms. Vennemann considers Basque to be the only survivor of a larger language family to which Aquitanian also belongs, termed the ‘Vasconic’ family. He believes that this family once extended throughout most of Europe:

“When the continent was becoming warmer about ten thousand years ago and the ice sheet was beginning to withdraw from large parts of Europe, both in a northerly direction towards the pole and a southerly direction towards the Alps, the Vasconic Old Europeans moved forward in Western, Central and Eastern Europe, so that nearly the entire continent became Vasconic.” (Vennemann, Languages in Prehistoric Europe north of the Alps).

According to Vennemann, the Vasconic peoples gave names to rivers and geographical landscape features, some of which still remained after Vasconic languages were replaced by Indo-European tongues.

However, this theory is not widely accepted. Other linguists, such as Trask, claim that pre-Indo-European Europe was very diverse, linguistically speaking:

“The idea the invading Indo-Europeans would have encountered as linguistically homogeneous Europe seems implausible in the extreme. Far more probably, what the Indo-Europeans found was a patchwork of languages, large and small, some related, some not, resulting from previous millennia of settlement, displacement and shift, just like anywhere else. Consider, for example, the linguistic position of the pre-Roman Iberian Peninsula.” (Trask, 1997: 364)
Donald Range supports Trask’s view, drawing his conclusion based on the findings of Johanna Nichols (1990) in an article discussing patterns of pre-state linguistic diversity. According to Range, the linguistic layout of pre-Indo-European Europe would have comprised approximately some sixty different languages, belonging to at least forty different language families. However, due to the nature of pre-state communal life, coastal areas, in particular the Mediterranean coastal zone would have exhibited greater linguistic diversity than the interior of the continent:

“In the most general terms, aboriginal Europe should have exhibited a degree of linguistic diversity comparable to that of western North America, with the Mediterranean region comparable to aboriginal California, the Atlantic coast comparable to the northwest coast of North America, and the hinterlands very roughly comparable.” (Range, The Linguistic Diversity of Aboriginal Europe)

Thus, regarding the linguistic pre-history of Europe prior to the arrival of the Indo-Europeans, there appears to be two main conflicting hypotheses:

1. An earlier form of Basque was one just one of many languages spoken in western Europe, and covered a geographical area similar in size to the area in which it is spoken today; or;
2. Modern Basque is the sole survivor of a family of languages which were once spoken throughout Europe.

Note that Vennemann does not dispute the existence of other, non-Vasconic languages in pre-Indo-European Europe, but does not consider them evidence of language survival, and argues that “the existence of such languages has to be demonstrated and cannot be assumed as a matter of course.” (Vennemann, Languages in Prehistoric Europe north of the Alps)

Let us focus on Range’s view that linguistic diversity in pre-state Europe would have been similar to the pattern of indigenous American languages in present-day North America. Range considers North America a suitable model for comparison with prehistoric Europe, as it illustrates patterns
of linguistic diversity without influence of state or government. The map in (12) shows the geographical areas occupied by different Amerindian languages families:

(12) Location of North American (Amerindian) language families

(http://www.absoluteastronomy.com/topics/Indigenous_languages_of_the_Americas)

The western coast of the United States exhibits the greatest degree of linguistic diversity. This is the type of diversity, which, according to Range, would have been found on the Mediterranean coast in aboriginal Europe. Other Amerindian languages families, however, are spoken over much larger areas. The Eskimo-Aleut, Na-Dene and Algic families cover areas which are not vastly smaller than Europe. Although Range compares the linguistic diversity of pre-Indo-European Europe to that of western North America, the distribution of such families suggests that it is possible for an indigenous language family (such as the proposed Vasconic family) to occupy a
vast geographical area. The area in which the Na-Dene language family is spoken, for instance, is similar in size to a substantial part of Europe.

From this it is possible to conclude that the idea of a (Vasconic) language family once being spoken over a sizeable part of Europe need not be ruled out. At the very least, Vasconic could have occupied a larger area than that in which modern Basque is spoken. The following section looks at the spread of pre-historic populations following the Ice-Age.

4.1 Geographical and Genetic evidence
During the last ice-age, when the air was about 10-12 degrees cooler than it is today, it was necessary for our ancestors to move south of the tree line to hunt game (Fleming, 2002). Most of northern Europe was completely inhospitable. The three main human populations in Europe at the peak of the ice-age are shown on the map in (13). It is significant that one of these three main regions inhabited by humans lies in the Iberian Peninsula, very close to modern Basque country.

(13) Human populations during the last ice-age

(13) Human populations during the last ice-age

(Fleming, 2002)

When the Earth became warmer around 12,000 years ago and the ice sheets began to retreat, our ancestors were able to move north and re-populate central and northern Europe. Geneticists have
identified three distinct DNA haplogroups which developed from genetic mutations whilst our ancestors waited the ice-age out in the locations shown in (13). These are R1b, I, and R1a. The map in (14) shows how these groups spread north as the ice-age ended.

(14) How humans populated Europe at the end of the last ice-age

![Map showing the spread of DNA haplogroups R1b, I, and R1a across Europe](image)

(Fleming, 2002)

The group which is of most interest to scientists studying Basque origins is the R1b group, which is associated with the population living in the Iberian Peninsula in (13). The R1b group gradually spread across western Europe and the British Isles as the population from the Iberian Peninsula moved north after the ice-age. Populations in Western Europe today carry a high percentage of the R1b haplotype. It is found mainly in Spain, France and the British Isles, and to a lesser extent, Germany and southern Scandinavia. See (15).

(15) Haplogroup R1b in Western Europe today
The inhabitants of the Basque country in south-western France and north eastern Spain carry a very high percentage of R1b DNA, as do the Celtic speaking regions of Britain and Ireland. The southern and eastern parts of the British Isles have a slightly lower percentage, as does France, the Netherlands and parts of Germany and southern Scandinavia. Considering this evidence, it seems that, after the ice-age, people living around modern Basque country moved north and populated regions of Europe from the British Isles, to southern Scandinavia and central Germany. The issue to consider now is whether these people spoke Vasconic, and spread their language(s) over the regions they populated (i.e. those with high R1b percentage). The distribution of Amerindian languages in (12) shows that it is indeed possible for an indigenous language to occupy a region of this size. Section (5) explores possible linguistic evidence which may support this idea.

5. Traces of Vasconic in Western Europe

5.1 Place names
When prehistoric peoples migrate, it is necessary for them to name important feature of their environment, and as they begin to settle in certain places, communal living areas or townships
which are established. Often, some of these names are retained by successive intruding populations, whilst others are re-named in the language(s) of the new settlers. If Vasconic speaking peoples did indeed re-populate Western Europe after the Ice Age, they would have given Vasconic names to features of their landscape. The vast majority of these would have disappeared without a trace, as the various towns and topographical features were re-named by Indo-European speaking populations. However, it is argued by some linguists (most notably Vennemann) that some Vasconic toponyms in Western Europe have persisted to today. Vennemann discusses many European toponyms which he believes to be of Basque origin. He considers this evidence that Vasconic speaking peoples once inhabited areas in which these names are found. Perhaps the most significant of these is the Basque word for valley ‘aran’, which is an element in various toponyms across Western Europe. These places, all valleys, are: Val d’Aran in the French Pyrenees, Arundel in Southern Britian, Arntal in Germany, Ahrental in Austria and Arendal in Southern Norway (Vennemann, 2003). The ‘aran’ component has been adapted in each instance to fit the language of the new population in each case by the addition of a grammatical head. The ‘aran’ element adopts the role of specifier to the new head and must either precede or follow it depending on whether the language of the new population is pre-specifying or post-specifying. In each case, the heads val, del, tal and dal all mean ‘valley’ in their respective languages. This process, involving the addition of a head which has the same meaning as the as the original toponym, is not uncommon in the formation of place names. Several place names in England, for instance, contain combinations of Celtic, Anglo-Saxon and Old Norse elements which have the same meaning in their respective languages.

Other possible Basque-related toponyms are the hydronomical roots is- and ur-, which mean ‘(body of) water’ (Venneamnn, 2003). The ur- element can be found in river and settlement names in Germany such as Ur-ach and Aur-ach (Vennemann, 2003): (the Germanic head –ach also means ‘river’). Vennemann considers the name of the river Thames in southern England to be Vasconic, consisting of the elements tam-is (Vennemann, 2003). This contains the Basque root for ‘river’, preceded by the Celtic word for ‘dark’, resulting in a Celtic-Basque world meaning of ‘dark river’. However, other etymologists consider ‘Thames’ to be of Latin origin, the river having been given its name by the Romans during their occupation of Britain. Further to this, Vennemann draws attention to the etymon bard-/part- in European place names:
“The etymon bard-/part- of Partenkirchen and its river Partnach is seen in many other toponyms all over Europe, e.g. Partenheim, Perticus saltus (La Perche), Partney, the river Parthe, Partington, Bardemara, and Bardenbach, and is connected to a weakly attested Basque word barta/parta meaning ‘swamp’ (Vennemann, 2003)

To this list the town of Parthenay in western France can also be added (which indeed lies on a river bank, hence the connection to ‘swamp’). Another possible Vasconic toponym is ‘Mendi’ in ‘Mendip Hills’ in south west England. Since ‘Mendi’ in Basque means ‘hill’, this toponym could have been formed in a similar way to those involving ‘aran’, in the sense that it has been integrated into English by the addition of a new head (‘hills’).

5.3 Loanwords

Vennemann cites several possible loanwords in Germanic which he believes to have been borrowed from Vasconic. These are: silver (Bq. zilar), iron, hook, callow, adze, shank, to stink and Gm. halde ‘sloping ground’, garbe ‘sheaf’, harn ‘urine’, latte/laden ‘to board’ (Vennemann, 2003). In addition to this, Venneman notes various similarities in Indo-European words for the concepts of ‘evening’ and ‘west’: Gk. hésperos, Lat. vesper ‘evening’, OIr. fescor ‘evening’, MCym. Gosper, gosber ‘evening’, Lith. vākaras ‘evening’, Pl. vakaraĩ ‘west’, Latv. vakars ‘evening’, OCS večerŭ ‘evening’, Arm. gišer ‘evening’, ON vestr ‘west’, OHG/OE/OS westan ‘(from the) west’. According to Vennemann, a single source for these loanwords would be something resembling wesk(w)er or wesk(w)ar. This is strikingly similar to euskera/euskara, i.e the Basques’ name for themselves, their land and their language. The term may well have been used by Indo-Europeans to refer to the people, the land and the language of the west. Over time, it would have been semantically broadened to mean ‘the west’ in general.

In addition, Trask identifies two probable loanwords which are attested in Basque and Old Irish:

1. Basque  ‘adar’ (horn/branch)  
   Old Irish  ‘adarc’ (horn)  

2. Basque  ‘andere’ (lady)  
   Old Irish  ‘ander’ (young woman)  
   (Trask, 1997: 369)
There is clearly a common origin for the Basque and Celtic words. The Basque word is the more ancient of the two, because it is attested in Aquitainian; however this does not necessarily prove that the borrowing is from Basque into Celtic. Trask states that ‘…the loan, if it is one, could equally have been in either direction, or into both from an unidentified source’ (Trask, 1997: 369).

5.4 Structural similarities

Vennemann identifies two key similarities between Basque and Western Indo-European. These are:

1) The first syllable accent in Germanic, Romance and Celtic. This is probably the result of a substrate language, which may well be Vasconic (stress in Basque always falls on the first syllable). This is not the case in all Indo-European languages and was probably not the case in Proto-Indo-European. (Vennemann, Languages in Prehistoric Europe north of the Alps)

2) Post-Specifying adjective placement in Romance (especially if compared to Germanic). This could also be a result of contact with Vasconic (post-specifying adjective placement is the only exception to the otherwise consistently pre-specifying Basque syntax). It is likely that this noun-adjective ordering is a substratal influence of some non Indo-European tongue, since proto-Indo-European had SOV syntax and therefore most probably had adjective-noun ordering. (Vennemann, Languages in Prehistoric Europe north of the Alps)

A further structural similarity between Celtic and Basque is the use of periphrastic verbal form. Celtic verb phrases are formed with an inflected form of the auxiliary ‘to be’, followed by the infinitive or lexical verb:

**Welsh**  
Dw i yn chwarae  

am I play  

“I play/am playing”

**Scots Gaelic**  
Tha mi a dol  

am I go  

“I go/am going”

However, as is the case in Basque, Celtic languages also have some synthetic verb forms.
5.5 Vigesimal Counting

Vigesimal counting systems are found in several Indo-European languages in Western Europe, most notably Celtic languages. In Scots Gaelic and Irish Gaelic, all multiples of ten are expressed in terms of ten and twenty (although Irish has also recently developed a modern system which is non-vigesimal). In modern P-Celtic languages, (Welsh and Breton), vigesimal counting is used for all multiples of ten except ‘thirty’ and ‘fifty’ (although, again, Welsh also has a modern non-vigesimal system). There are also traces of this in English, which occasionally uses ‘score’ to refer to ‘twenty’ in structures such as ‘four score and ten’ (ninety). In Danish, vigesimal counting is used for all multiples except thirty and forty, and the numbers ‘eighty’ and ‘ninety’ in standard French are expressed in terms of ‘four times twenty’. Albanian also has a trace of vigesimal counting, with ‘forty’ expressed as ‘two times twenty’. The only non-Indo-European language other than Basque which has vigesimal counting is Georgian (which interestingly also uses ergative-absolutive case marking; see section 3.3). It is noticeable that the regions of western Europe in which vigesimal counting is used roughly match those with a high R1b percentage, particularly Celitic speaking regions. Although this is generally Western Europe, the area around Georgia, just east of the Black Sea also has a slightly higher proportion of R1b than other areas of Eastern Europe (note the darker shading around the Caucasus on the map in (15)). Vennemann considers the vigesimal counting systems of modern western Indo-European languages to be a substratal influence of Vasconic. It is fairly safe to say that Vigesimal counting in western Indo-European is a substratal loan from a non-Indo-European source. There seem to be no other explanation as to why Danish, for instance, should use vigesimal counting whilst other North Germanic (Scandinavian) languages, to which Danish is closely related, do not. The question that remains is whether vigesimal counting was borrowed from Vasconic into western Indo-European languages, or whether the structural loan was into both families from an unidentified source.

6. Connections between Basque and Celtic

Based on evidence so far, it seems fair to say that Basque has a stronger connection with Celtic than any other Indo-European sub-family. This is reflected in the following:

1) The genetic map of Western Europe in (15), which shows that Basque and Celtic speaking regions of Europe share a high percentage of haplotype R1b. This is significantly higher than other parts of Europe.
2) Linguistic structures. Basque and Celtic both use vigesimal counting, periphrastic verbal forms and noun-adjective ordering. Unfortunately, there are very no items of shared vocabulary other than those mentioned in (5.3).

The reason for this Basque-Celtic connection lies in the fact that the Celts were the first Indo-Europeans to populate Western Europe, including possible Vasconic speaking regions such as Britain and Ireland. In the 13th Century BC, the distribution of Indo-European language subfamilies was as shown in (16):

(16) Europe in 13th Century BC

![Europe in 13th Century BC Map](http://indoeuro.bizland.com/project/chron/eur1.html)

Notice that, whilst the Celts had migrated furthest west of all the subfamilies, they had not yet reached as far as Spain, western France or the British Isles. These regions would still be non-Indo-European speaking. By 500 BC, Celtic had spread across the British Isles and parts of France and Spain. If Vasconic was indeed spoken across this area, the two language families would have been in contact at this point. The mixing of the two cultures would have resulted in a sizeable degree of Basque-Celtic intermarriage, and consequently, R1b genes would have spread into Celtic. If this is the case, the Celts must have been economically or agriculturally superior to the Vasconic population which they encountered and thus, Celtic would have been the dominant tongue. The
Vasconic peoples would have given up their native language and learned Celtic; however they would have acquired it imperfectly, and in the process injected into it grammar and structures from their native language, in what is known as a substratum effect. This is how some structural patterns of Vasconic (e.g. periphrastic verbal form, vigesimal counting) were adopted into Celtic languages. The borrowing of linguistic structures- from the former, less superior language into the language which was imposed on the area- is typical of a substratum effect. Over generations, these structures would have spread throughout the language, being adopted by people of Celtic descent as the two populations mixed. In a substratum effect, few, if any, items of vocabulary from the substrate language (in this case Vasconic) are found in the dominant language (in this case Celtic). The only real exception to this is the adoption of toponyms (see section 5.1).

Over the following two and a half millennia, Celtic came to be (mostly) obliterated by imposing Romance and Germanic languages, in a similar way to which Vasconic was obliterated by Celtic. However, just as Vasconic had a substratal influence on Celtic, Celtic had a substratal influence on certain Romance and Germanic languages. Some Celtic structures (some of which were originally Vasconic structures) passed into English and into French. This is why relics of vigesimal counting can be found in French (see section 5.2) and, to a certain extent in English. Some scholars think that the English present continuous tense, which is periphrastic, is a substratal loan from Celtic, although this is highly debatable. Regarding genetics, the intermarriage of Celts, who carried a high percentage of R1b DNA inherited from Vasconic, and Anglo-Saxons who invaded Britain circa 400AD allowed for R1b genes to be passed to the new population. Anglo-Saxon invaders eventually came to dominate Britain, pushing Celtic towards the western extremities of the country in which languages such as Welsh and Scots Gaelic are now spoken. See (10).

(17) Europe today
In summary, Vasconic genetic types were passed on to the Celts and then to successive populations such as the Anglo-Saxons. Similarly, certain linguistic features of Vasconic were adopted by the Celts, and later passed from Celtic into other languages:

Vasconic → Celtic → Romance

However, if Vasconic languages were indeed once spoken throughout Western Europe, one would question why they have not survived in any region other than modern Basque country. It is believed that Basque survived in the western Pyrenees due to Roman neglect of the area, which enabled to culture and language to become better established. Further to this, the Basques later developed a strong seafaring culture, which not only defined them as a people but also left them less prone to being overwhelmed by intruding populations. Often when states become established, minority languages tend to dwindle and eventually become extinct, as has been the fate in Britain for many Celtic languages (E.g. Cornish, Cumbric, Manx). However, the Basque culture of seafaring has been significantly advanced and respected, such that it is not in danger of diminishing. The Basques’ reputation at sea became truly established in the seventh century when they began hunting whales, of which meat, bone and blubber were valuable commodities.
(Connelly, 2004: 201), and they subsequently became innovative shipbuilders. It may well be that the nautical expertise of the Basques has been the key to the survival of their language and culture.

7. Conclusion

To summarise, the Basque language has many grammatical features which set it apart from Indo-European languages; most notable the ergative-absolutive case system and the use of dual auxiliary marking. Whilst it is not genetically related to any known language other than Aquitanian, possible traces of Basque syntax can be found in many western Indo-European languages, especially Celtic languages. Further to this, there are a significant number of Western European toponyms which have plausible Basque roots. This evidence suggests that Basque, or an earlier form of it, may have been spoken over this area prior to the arrival of the Indo-Europeans. The distribution of indigenous languages in North America suggests that it may well have been possible, in prehistoric times, for a single family of languages to occupy an area of this size. As far as Genetic evidence is concerned, the population who inhabited the Iberian Peninsula during the last Ice-Age eventually moved north and re-populated Western Europe as the ice retreated. They spread the DNA haplogroup R1b across these regions. Today, R1b is found throughout Western Europe, though mainly in Ireland and the Basque country. It is fair to say that the regions of Europe whose populations have the highest R1b percentage are those whose languages bear some structural resemblances to Basque, particularly in the use of vigesimal counting. This begs the question of whether we can conclude on these grounds that Vasconic languages were spoken in the Iberian Peninsula during the Ice-Age, before spreading across France, the United Kingdom, Ireland and parts of Germany and Scandinavia. There is also the possibility that any linguistic similarities between Basque and western Indo-European were borrowed into both language groups from an unknown source of which no other traces can be found. This would have been an indigenous language (or family of languages) once spoken in Western Europe, which had substratal influences on both Basque and Indo-European. Nonetheless, no strong suggestions have been put forward as to what this language may have been. On balance, there is not enough evidence to be able to conclude in any certain terms that Basque, or an ancestral form of it was once spoken across most of Western Europe. In the absence of evidence to the contrary, however, it is a plausible suggestion which need by no means be ruled out in studies of pre-Indo-European Europe.
8. References


