Berkeley Linguistics Society. 2010. 16-31. http://dx.doi.org/10.3765/bls.v36i1.3900
Published by the Linguistic Society of America

Language Isolates and Their History, or, What's Weird, Anyway?

LYLE CAMPBELL
University of Hawai'i at Mānoa

1 Introduction

How many language isolates are there in the world? (How many language families are there?) Most linguistics do not know, and opinions vary greatly. The answers to these questions are complicated because they depend on different views about fundamental issues in historical linguistics. The goal of this paper is to attempt to answer the questions: How many language isolates are there? How can we advance knowledge of the history of language isolates? What methodological lessons does the study of specific isolates offer to understand better the history of language isolates in general and that of other specific isolates? What are the prospects for finding relatives for some language isolates, that is, for showing that they belong to larger genetic groupings than those known at present?[1]

To begin, we need to ask, what is a language isolate? In the most common view, an isolate is a language which has no relatives, that is, that has no demonstrable genetic relationship with any other language. It is a language which has not been shown to be the descendent of any ancestral language which has other descendants (daughters). Thus, language isolates are in effect language families with only one member. The best known and most cited language isolates are Basque, Burushaski, and Ainu, though there are many others not so generally known.

Since language isolates are often contrasted with families of related languages, we also need to ask, what is a language family? As is generally known, a language family is a set of languages for which there is sufficient evidence to show that they descend from a single ancestral language and are therefore genetically related. The total number of language families in the world is the set of independent families for which no relationship can be demonstrated with any

[1] Portions of this paper are based on Campbell (2011).

other language family. And, as mentioned, a "family" can be composed of but a single language in the case of language isolates.

So, how many independent language families (including isolates) are there in the world? There are approximately 420 (Hammarström 2007, Campbell 2013).

2 Why Language Isolates and Language Families are not so Very Different

There are two senses in which language isolates are not very different from language families (of multiple members). First, some language isolates may have had relatives in the past which have disappeared without coming to be known, leaving these languages isolated.

For example, Ket in Siberia is the only surviving language of the Yeniseian family. Nevertheless, there were other Yeneseian languages, now extinct: Arin, Asan, Kott, Pumpokol, and Yugh (Vajda 2001). If these languages had disappeared without a trace, Ket would be considered an isolate. However, since data from these extinct languages was registered before they disappeared, Ket was not left an isolate, rather a member of a family of languages, albeit the only one surviving. Examples such as this show that language isolates could well have once been members of languages families whose other relatives disappeared before they could come to be known, illustrating why in this way language isolates are not so very different from languages families.

2.1 Did Basque Have Now Extinct Relatives?

This raises an interesting question: is it possible that Basque once had relatives and is thus not really a language isolate? This brings up a question about the relationship between Basque and Aquitanian. As Trask (1997:411) reports, "probably all Basque scholars now accept that Basque descended more or less directly from Aquitanian" (see also Trask 1995:87, 1997:35). However, is it possible that Aquitanian and Basque are related languages, two members of a language family, rather than Aquitanian being a direct ancestor of Basque?

It is possible that the claim of Basque as a direct descendant of Aquitanian is incorrect. The attestations of Aquitanian are from c. 2000 years ago (see below); however, most languages known from 2,000 years ago have diversified and become families of languages: Romance, Finnic, Slavic, Germanic, Turkic, Mongolian, Polynesian, etc. It is likely that Aquitanian of so long ago would also have diversified into different languages, distinct from one another. However, if no other except Basque survived, Aquitanian could be the direct ancestor of Basque. Another possibility is that Aquitanian had a sister or sisters of its own, diversified from an earlier common ancestor and that Basque descends from a sister of Aquitanian rather than directly from Aquitanian itself. The attestations of Aquitanian allow for this second possibility.

Although these attestations are sufficiently detailed to confirm that modern Basque and Aquitanian are related, they also show sufficient differences from Basque to suggest the possibility not of a direct ancestor, but as a relative, that possibly Aquitanian and Basque are sister languages representing two branches of the original proto-language. The corpus of Aquitanian data is limited to about 400 personal names and 70 names of deities, found in texts written in Latin. There are no Aquitanian texts larger than these names. (See Michelena 1988, Gorrochategui 1984, 1995, Trask 1997:398-403.) Compare in Table 1 the Aquitanian words (from c.100 CE) with those of Basque on the one hand, and the Gothic words (from 382 CE) with those of English on the other.

Table 1: Comparison of the relation between Aquitanian and Basque with that between Gothic and English

Aquitanian	**Basque**	**Gothic**	**English**
arixo	haritz	(faírguni)[2]	oak
atta	aita	fadar	father
belex	beltz	swarts	black
bon-	on	goÞs	good
sembe-	seme	sunus	son
hanna	anaia	broÞar	brother
seni-	sehi/seiñ	magus	boy
oxson	otso	wulfs	wolf
siri(co)	suri	hweits	white
ausci	euska(ra)		'Basque'

It is easy to see that in spite of clear similarities between Aquitanian and Basque, there are also marked differences. It is also clear that the similarities and differences between the cognates from Gothic and English are quite similar to those between Aquitanian and Basque. However, Gothic cannot be considered a direct ancestor of English – the two belong to distinct branches of Germanic. Their differences and similarities, when compared with those between Aquitanian and Basque, turn out to be very similar in nature. Given this similarity, it should be asked, could the relationship between Basque and Aquitanian be that of related languages as exists between Gothic and English, and not like that between Latin and Spanish, where an ancestral language and its descendant are involved?

2.2 Could Basque Have Modern Relatives?

In the other sense in which language isolates and language families are not so different from one another, some languages which were thought to be isolates

2 Gothic has no attested form for 'oak'; *faírguni* 'mountain' is generally believed to have in it the root for 'oak', from Proto-Indo-European **perkw*-.

have in reality proven to be members of small families of related languages. For example, Japanese would be a language isolate if Ryukyuan languages (of Okinawa) had not been shown to be distinct languages, related to Japanese. Thus Japanese belongs to a family of languages (often called Japonic) and is no longer an isolate. Similarly, Etruscan, long cconsidered an isolate, was shown to be related to Lemnian. Lemnian is known from a stella and ceramic fragments from the Greek island of Lemnos, dating from c.400 BCE (cf. Steinbauer 1999:357-66).

Some other cases of small families no longer considered language isolates because related languages have come to be known are:

Atakapan Texas and Louisiana (two languages: Atakapa, Akokisa)
Jicaquean (Tol) Honduras (two languages: Tol, Jicaque of El Palmar)
Lencan El Salvador, Honduras (two languages: Chilanga, Honduran Lenca)
Xinkan Guatemala (four languages: Chiquimulilla, Guazacapán, Jumaytepeque, Yupiltepeque)
Hurrian (Hurro-Urartean) Northeast Anatolia, from the state of Mitanni, known from the second and first millennium BCE.

Cases such as these lead us to ask, is it possible that Basque constitutes a small family of languages? Basque would not be considered an isolate if Aquitanian proves to be a separate language, not an ancestor of Basque but genetically related to it. This is quite possibly the case. Also, when two or more 'dialects' are not mutually intelligible, by standard criteria they are considered separate languages. Basque would become a small family if its dialects have diversified so much that some are no longer mutually intelligible. This appears to be the case. Since Proto-Basque (c.600 CE, Lakarra 1995:193) enough time has transpired for Basque to have diversified. As Trask (1997:5) affirms, "the differentiation [between the dialects] is sufficiently distant that speakers of different areas can have significant difficulty understanding one another when using the vernacular form of Basque."

The classification of Basque dialects varies, with disagreements about how many dialects there are and how to classify them; but the following are recognized:

Bizkaiera (Vizcaíno, Viscayan)	Gipuzkera (Guipuzcoano)
Nafarrera garaia (High Navarro)	Aezkera (Aezcoan)
Salazarera (Salacenco)	Erronkariera (Roncalese)
Lapurtera (Laburdino, Labourdin)	Nafarrera behera (Low Navarro)
Zuberoa (Suletino, Souletin).	

It is generally conceded that at least Zuberoa (Suletino, Souletin) is not mutually intelligible with the others. Given this, some consider Basque no longer

an isolate, but a small language family. The 2005 edition of *Ethnologue* (Ethnologue.com) listed three Basque languages, though the 2014 edition has one, noting "some inherent intelligibility among regional varieties except Souletin."

The point is not to insist that Basque formerly had relatives (questioning the status of Aquitanian) nor that it has relatives now (assuming lack of mutual intelligibility), but rather to show that Basque easily could cease to be a language isolate, and therefore that language families and isolates are not so very different.

3 Further Clarification of the Concept 'Language Isolate'

It is necessary to distinguish language isolates from unclassified languages, languages so poorly known that they cannot be classified, though sometimes listed as isolates. An unclassified language is one for which there is not enough data (documentation/attestation) to know whether it has relatives – these languages lack sufficient data for them to be compared meaningfully with other languages and therefore their possible kinship remains unknown. Isolated languages are not grouped in larger genetic classifications because for them there do exist data and comparisons of these data with other languages do not reveal linguistic kinship.

There are two sorts of unclassified languages. The first are extinct languages which are too poorly attested to be grouped with any other language or language group. Some examples include:

Adai, Louisiana
Aranama-Tamique, Texas
Camunico, Northeast Italy (survived to 2nd half of 1st millennium BCE)
Eteocretan, Crete, 4-3 centuries BCE.
Ibrerian, Iberian Penninsula (2nd half of 1st millennium BCE to 1st half of 1st millennium CE).
Indus Valley, India and Pakistan, 2500-1900 BCE. (Based on undeciphered inscriptions, whose status as a real writing system is disputed, cf. Farmer et al. 2004).
Kara, possible language of Korea, only from 13 toponyms.
Kaskean, Northeast Anatolia 2nd millennium BCE.
Koguryo possible language, NE China, Manchuria, Korea, 1-8 centuries CE, known only from toponyms and a few words.
Ligurian, Northeast Italy, few words, 300 BCE-100 CE.
Maratino, Northeast Mexico
Minoan Linear A, undeciphered, 1800-1450 BCE.
Mysian, Western Anatolia before the 1st century BCE.
Naolan, Tamaulipas, Mexico.
Northern Picene, Adriatic coast of Italy, 1st millennium BCE.
Pictish, Scotland 7-10 centuries CE, few inscriptions.
Puyo, Manchuria (with Koguryo?), few attested words.

Quinigua, Northeast Mexico.
Raetic, Northern Italy, Switzerland, Austria, 1st millennium BCE.
Sicanian, Central Sicily, pre-Roman epoch.
Solano, Texas, Northeast Mexico.
Sorothaptic, Iberian Peninsula, pre-Celtic, Bronze Age.
Tartessian, Spain, 1st millennium BCE.

The second kind are the extant languages not classified for lack of data, languages not described sufficiently to compare them with other languages in order to determine whether they may be related. Examples include:

In Africa: Bung, Lufu, Kujargé, perhaps Mpre, Oropom, Rer Bare, Weyto.
In Asia and the Pacific: Sentinelese (Andaman Islands), Bhatola (India), Waxianghua (China), Doso (Papua New Guinea), Kehu (Indonesia Papua), Kembra (Indonesia Papua), Lepki (Indonesia Papua).
In South America: Amikoana, Arara, Cagua, Carabayo, Chipiajes, Coxima, Ewarhuyana, Himarimã, Iapama, Kaimbé, Kamba, Kambiwá, Kapinawá, Karahawyana, Kohoroxitari, Korubo, Miarrã, Natagaimas, Pankararé, Papavô, Pataxó-Hãhaãi, Tapeba, Tingui-Boto, Truká, Tremembé, Uru-Pa-In, Wakoná, Wasu, Yarí, etc.

It should be noted that some of these unclassified languages could also be language isolates, but without evidence we cannot know.

4 How Many Language Isolates are There, Really?

With the clarification that the unclassified languages are not language isolates, we return to the question, how many isolates are there in the world. The list, compiled from consensus reports from specialists in each region, is:

Africa: [10]

Bangi-me	Centúúm
Hadza	Jalaa
Kwadi	Laal
Mekejir (Shabo)	Meroitic Extinct
Ongota (Birale)	Sandawe

Asia: [9]

Ainu	Burushaski
Korean	Elamite Extinct
Hattian Extinct, poorly know	Nihali
Kusunda	Nivkh (Gilyak) (two languages?)
Sumerian Extinct	

Australia [7]

Enindhilyagwa (Andilyaugwa)	Kakadju (Gaagudu)

Laragiya Extinct?
Ngurmbur Extinct?
Umbugarla
Minkin Extinct
Tiwi

Oceania [14]

Abinomn, Indonesia Papua)
Anem, New Britain
Elseng, Indonesia Papua
Isirawa, New Guinea
Kuot (Panaras), New Ireland
Mpur, Indonesia Papua
Pele-Ata (Wasi), New Britain
Sulka, New Britain
Yalë (Nagatman), New Guinea
Yele, Papua New Guinea
Abun, Indonesia Papua
Busa (Odiai), New Guinea
Hatam, Indonesia Papua
Kol, New Britain
Massep, Indonesia Papua
Odiai, Papua New Guinea
Pyu, New Guinea
Taiap (Gapun), Papua New Guinea
Yawa, New Guinea (two languages?)
Yuri (Karkar), New Guinea

Europa: [1+]

Basque (Some would include Tartessian and Iberian, extinct languages of Spain, probably better considered unclassified due to insufficient information.)

North America: [20]

Adai, extinct, Texas, Louisiana
Chimariko, extinct, California
Esselen, extinct, California
Karankawa, extinct, Texas
Kootenai, Idaho, Montana, BC
Salinan, extinct, California
Takelma, extinct
Tonkawa, extinct, Texas
Washo, California, Nevada
Yana, extinct, California
Cayuse, extinct, Oregon
Chitimacha, extinct, Louisiana
Haida, Alaska, British Columbia
Karok, California
Natchez, ext., Mississippi, Louisiana
Siuslaw, extinct, Oregon
Timucua, extinct, Florida
Tunica, ext., Mississippi, Louisiana
Yuchi, Georgia, Oklahoma
Zuni, New Mexico

(See Golla et al. 2008.)

Mexico [6]

Coahuilteco, ext., Texas, NE México
Cuitlatec Extinct, Guerrero
Purhépecha (Tarascan)
Cotoname, extinct, NE México
Huave, Oaxaca
Seri, Sonora

(See Campbell 1997)

South America [55]:

Aikaná, Brazil
Awaké, Venezuela, Brazil
Betoi, Colombia
Candoshi, Peru
Cayuvava, extinct, Bolivia
Chono, Chile
Culle, extinct, Peru
Guató, Brazil
Andoque (Andoke), Brazil, Peru
Baenan, Brazil
Camsá (Sibundoy), Colombia
Canichana, Bolivia
Chiquitano, Bolivia
Cofán (A'ingaé), Colombia, Ecuador
Gamela, extinct, Brazil
Irantxe (Iranche, Münkü), Brazil

Itonama, Bolivia
Jotí (Yuwana), Venezuela
Kapixaná (Kanoé), Brazil
Koayá (Kwaza, Koaiá), Brazil
Máko (Máku), Brazil
Matanauí, Brazil
Movima, Bolivia
Natú, extinct, Brazil
Omurano, Peru
Pankararú, extinct, Brazil
Rikbaktsá, Brazil
Tarairiú, Brazil
Taushiro (Pinche), Peru
Ticuna, Colombia, Brazil, Peru
Tuxá, extinct, Brazil
Wamoe, Brazil
Xokó, Brazil extinct
Yagan, Chile
Yuracaré, Bolivia
Yurumangui, extinct, Colombia
Jeikó, Brazil
Kaliana (Sapé),Venezuela
Kirirí, Brazil
Kukurá (Cucura, Kokura), ext., Brazil
Mapudungu, Chile
Mochica (Yunga, Chimu), Peru
Munichi (Otanabe), extinct, Peru
Ofayé (Opayé), Brazil
Otí, extinct, Brazil
Puquina, extinct, Bolivia
Sabela (Auca), Ecuador
Taruma, Brazil, Guyana
Tequiraca (Auishiri), Peru
Trumai, Brazil
Urarina, Peru
Warao, Guyana, Surinam, Venezuela
Xukurú, extinct, Brazil
Yaté (Fulniô), Brazil
Yuri (Jurí), extinct, Colombia, Brazil
(See Campbell 2012)

Thus, the total number of isolates in the world is **136**. There are c.420 independent language families (including isolates), for which it is not possible to demonstrate a genetic relationship with any other language family. Isolates make up 32% of all "language families," about one-third of the world's linguistic diversity. Seen from this perspective, isolates are not at all weird; they have as their "cohorts" over one-third of the "language families" of the world.

How do we explain the general attitude that language isolates are weird, so unusual that they are suspicious, and the frequent feeling that languages with no relatives should not be tolerated? I suspect these feelings stem from lack of understanding about how many isolates there are and of how little isolates differ from other languages families, as seen above.

5 How Can We Advance Our Knowledge of the History of Language Isolates?

How can we learn about the history of a language without relatives? One attitude about Basque has been that if it has no relatives then it has no history. De Saussure (1916:298, see Michelena 1995:101) said that "we cannot derive anything from Basque because, being an isolate, it does not allow any comparison." Meillet (1925:11-2, see Michelena 1995:101) said that "if a language is an isolate, it lacks history…so if we cannot find a way to demonstrate a relationship between Basque and some other language, there will never be any

hope of finding out anything about its history." These attitudes make us ask, how can we learn about the history of languages without relatives? Must we accept the claim that an isolate is a language without history? What lessons does the study of Basque and other isolates offer us to understand better how to investigate the history of language isolates in general? As we will see, these attitudes are mistaken –isolates indeed have history and there are means of studying their history. The means that can be employed to learn about the history of isolates include:

Internal reconstruction	Philological study of attestations
Toponyms	Personal Names, names of deities
Historical reports	Comparative reconstruction based on dialects
Evidence from loanwords	Language contact and areal linguistics
Wörter und Sachen	

These are taken up in turn.

5.1 Internal Reconstruction

Internal reconstruction is the best known and most used tool to investigate the history of isolates. Excellent examples of its deployment are Michelena (1988, 1995), Lakarra (1995, 2006), and Trask (1997). For example, Lakarra (1995) on the reconstruction of the roots of Pre-Proto-Basque is a valuable application of internal reconstruction to obtain significant historical understanding of a language isolate. Here I offer a single example of a single word, to illustrate what can be gained. Basque 'wine' is reconstructed as **ardano*. It has regional variants *ardo, ardao, arno,* and *ardu* (Lakarra 1995:195), but even with standard Basque *ardo* in isolation and *ardan-* in compounds, internal reconstruction takes us close to the **ardano* reconstruction – evidence internal to Basque reveals the change of *-n-* > *Ø* (loss of intervocalic *n*), see below.

5.2. Philological Investigation of Attestations

Michelena (1988) and Gorrochategui (1984, 1993, 1995) have made very valuable studies of the older attestations of Basque. These include older citations of Basque forms, toponyms, personal names, deity names, and historical reports.

5.3 Comparative Reconstruction

A less well known but extremely valuable tool is the comparative method applied not to separate related languages, but to regional dialects. Successful and instructive cases include: Basque (Gorrochategui and Lakarra 1996, 2001, Michelena 1988, 1995, Trask 1997); Ainu (Vovin 1993); Huave (Suárez 1975);

and Tarascan (Friedrich 1971). These studies show that obviously Meillet, de Saussure, and others were mistaken in insisting that nothing can be known of the history of a language isolate if it has no relatives, and that isolated languages do not lend themselves to any comparison. (See above for Basque dialects.)

5.4. Loanwords

Another source of evidence on the history of isolates is loanwords. For example, from the semantic content of the more than 300 ancient loanwords from Latin into Basque it is clear that the Romans had much influence in the areas of laws, administration, technology, religion, and refined culture. Moreover, the relative age of many of these loanwords in Basque is known from phonological traits. Many were borrowed before the changes in Romance of the 5 long and short vowels to a system of just 7 vowels, as in Basque *gertu* 'certain, ready' [cf. Spanish *cierto*] < CERTU; *joko* 'game' [cf. Spanish [*juego*] < JOCU), and before the palatalization of velar consonants before front vowels, as seen in Basque *gertu* 'certain, ready' < CERTU; *gisu* 'lime' < GYPSU) [cf. Spanish *gis* [xis] (formerly [šis] from [žis]), and before the voicing of intervocalic stops, as in Basque *bake* 'peace' < PACE) [cf. Spanish *paz*]. (Michelena 1988, 1995, Trask 1997).

Loanwords can indeed provide considerable historical information about isolates, as in Basque. In another case, from Mesoamerica, we know something of the history of Huave (isolate) and its speakers from words borrowed from Mixe-Zoquean (MZ). Some examples are:

Huave *pom* 'copal (incense)' < PMZ (Proto-Mixe-Zoquean) **poma* [necessary in Mesoamerican ritual]
Huave *koy* 'rabbit' < PMZ **koya* [calendric name]
Huave *patsi* 'lizzard' < PMZ **patsi* [calendric name]
Huave *pïkɪ* 'feather' < PMZ **pïk* [important in pre-Columbian trade] (*ï* = barred "i")
Huave *kawak* 'chicozapote, mamey' < PMZ **ka'wak* 'chicozapote fruit' (the *'* = glottal stop)
Huave *yati* 'anona, chirimoya [soursop]' < Zoque *yati, ati*.

Several of these loans show cultural influence from Mixe-Zoquean on Huave, loans that reflect cultural concepts in ancient Mesoamerica. They support the hypothesis that the ancient Olmecs – the first highly successful agricultural civilization in Mesoamerica – spoke a Mixe-Zoquean language. Mixe-Zoquean influenced many other languages in the area (Campbell and Kaufman 1976).

5.5 Areal Linguistic Traits

Another source of information about the history of isolates is areal linguistics. A linguistic area (*Sprachbund*) is a geographical region in which, due to language

contact, languages of the area share structural traits, not through inheritance, but due to borrowing/diffusion. Areal linguistic traits reveal historical contacts and help to explain certain changes in the languages involved, including in isolates, as exemplified by the following areal traits in Basque owed to contact with neighboring languages:

(1) *s* is apico-alveolar in most varieties, but is apico-post-alveolar for most French Basque speakers (Trask 1997:84), due to French influence.
(2) The Basque phoneme written <j> is [ž] in Zuberoa, presumably due to influence from French. It is [x] or [X] in Gipuzkoa and East of Biscaya, from influence from Spanish. (Trask 1997.)
(3) Loss of intervocalic (lenis) *-n-* and *-l-* is apparently an areal trait, shared also with Portuguese, Galician, and Asturian. (Trask 1997.)
(4) Basque *u* has become *ü* in Zuberoa, probably due to influence from French.
(5) Basque initial *h-* is lost in most dialects (not in Zuberoa). This loss is probably due to influence from Spanish and French. (Trask 1997.)
(6) Basque epenthesized a vowel before initial *r*, e.g. *errege* 'king' (borrowed from Latin *rege*) and *erloju* 'clock' (borrowed from Spanish *reloj*). This feature is shared also with Aragonese and Gascon (Lakarra 1995:198.)
(7) An older trait possibly due to areal influence, older Basque, Aquitanian, and Iberian all basically lack *p*. (Michelana 1995:112, Trask 1995:78, 87.)

These facts also provide information about the history of Basque.

5.5 Wörter und Sachen

Wörter und Sachen strategies also provide information on the history of isolates. These are strategies for detecting past language-and-culture relations.

One strategy involves the **analyzability of words** (their morphological complexity) – words that can be analyzed into transparent parts are believed to be more recent than words which have no internal analysis. It is believed that words which can be analyzed into parts were created more recently than words which have no such internal composition, thought potentially to be older forms (Campbell 2013:434-6). For example, Basque *garagardo* 'beer' is analyzable morphologically: *garagar* 'barley' + *ardo* 'wine'; however, *ardo* 'wine' has no evident morphological analysis; therefore, it is inferred that the word for 'wine' is probably older than the word for 'beer'. Similarly, Basque *gari* 'wheat' is inferred to be older than *garagar* 'barley', since *garagar* is a reduplicated from of the word for 'wheat' and thus morphologically analyzable. And, the word for 'wheat' too must be older than that for 'beer', since the 'barley' component of 'beer' is morphologically complex, with 'wheat' in it. Basque *janarbi* 'radish' is analyzable as *jan* 'eat + *arbi* 'turnip'; however, *arbi* 'turnip' has no such internal structure; it is inferred that the 'turnip' word is older than the 'radish' word.

Another *Wörter und Sachen* strategy involves the **analyzability of toponyms**. It is also inferred that place names that can be analyzed into component parts probably came to be known more recently than those which have no such internal analysis. Thus, for example, it is inferred that *York* is older than *New York*, since the latter is composed of identifiable pieces, but not the former. In Basque, since the names of several rivers in the French Basque area have no clear etymology (not analyzable into parts), it is inferred that they are old names, for example *Atturri* (Adour), *Bidasoa, Biduze, Errobi*. The names of several rivers of Biscaya, on the other hand, are analyzable, for example *Ibaizabal* from *ibai* 'river' + *zabal* 'wide', and *Artibai* from *arte* 'between(?') + *ibai* 'river'. It is inferred that these latter names are not as old in the language as the former.

A third *Wörter und Sachen* strategy involves words which bear **non-productive (irregular) morphemes**; these are assumed to be possibly older than words composed only of productive morphemes. In Basque, for example, the morph *-di* is frozen, not productive, and its presence in the animal names *ardi* 'sheep', *zaldi* 'horse', *idi* 'ox', and *ahardi* 'sow' suggests that these animals have been known for a long time. For example *zaldi* 'horse', with non-productive *-di*, appears older than *zamari* 'horse', which is confirmed as a loanword (from Latin SAGMARIU 'pack-horse'). In general, though, it is possible only to conclude that words containing the non-productive morphology are old, but nothing can be inferred about the age of words lacking such forms. For example, for *otso* 'wolf' and *ahuntz* 'goat', lacking the irregular morphology, it is not possible to say anything of their relative age in the language.

In sum, based on these resources just seen, much is known of the history of Basque. This demonstrates that we can learn about the history of isolates.

6 What Can We Predict About the Possible Distant Genetic Relationships for Some of These Language Isolates?

What prospects are there for coming to reliable classifications that would include some of the language isolates in larger genetic groupings than those currently known? In answer to this question, we can cite relatively recent successful demonstrations where it has been possible to show a relationship for some languages previously considered isolates, for example:

Harakbmut-Katukinan ? (Adelaar 2000)
Lule-Vilela (Lule and Vilela) (Viegas Barros 2001)
Western Torres Island and Pama-Nyungan (Alpher, O'grady, and Bowern forthcoming)
Tikuna-Yurí (Tikuna and Yurí) (see Campbell 2012)

Judging from these successful instances, it can be expected that with more data and dedication, following adequate methods (see Campbell and Poser 2008), more cases of genetic relationship involving some language isolates will be discovered.

Nevertheless, it is not to be expected that there will be many of these. In particular in the case of Basque it has already been shown that the proposals that have been made do not support the assumption of a genetic relationship between Basque and any other language or language family (Lakarra 1996, 2006, Trask 1995).

7 Conclusions

From the above considerations, the conclusions that follow are:
(1) There is nothing unusual about isolates; there are 136 isolates in the world.
(2) Language isolates make up about one third of the language families in the world's total of c.420 independent families (including isolates).
(3) Language isolates are not very different from languages which have relatives. Isolates could easily have had relatives now lost or could diversify into small families of related languages.
(4) Language isolates, which have data, should not be confused with unclassified languages, which are not classified for lack of data.
(5) We have made progress in the search for relatives of Basque and other language isolates in that we have proven that many hypotheses of distant genetic relationship are not supported by the evidence, and much more is known now of the methods necessary to demonstrate a genetic relationship among languages (see Campbell and Poser 2008).
(6) In spite of doubts about discovering anything about the history of isolates, there are several resources (tools, techniques) which can help to recover considerable historical information about these languages; these include: internal reconstruction, philological investigation of earlier attestations, comparative reconstruction based on the dialects, evidence from loanwords, language contact and areal linguistics, and *Wörter und Sachen* strategies.
(7) It can be expected that with more data and dedication, employing adequate methods, new genetic relationships will be discovered for some language isolates. However, it is not to be expected that there will be many such cases, and this is highly unlikely in the case of Basque.

References

Adelaar, Willem F. H. 2000. Propuesta de un Nuevo Vínculo Genético entre dos Grupos Lingüísticos Indígenas de la Amazonia Occidental: Harakmbut y Katukina. *Actas I Congreso de Lenguas Indígenas de Sudamérica*, ed. by Luis Miranda Esquirre, 2:219-236. Lima: Editorial Universitaria.

Alpher, Barry, Geoffrey O'grady, and Claire Bowern. Forthcoming. Western Torres Strait Language Classification and Development.

Bowern, Claire. 2007. Australian Models of Language Spread. Paper presented at the International Conference on Historical Linguistics, Aug 6-12, 2007, University of Quebec, Montreal.

Campbell , Lyle. 1997. *American Indian Languages: The Historical Linguistics of Native America*. Oxford: Oxford University Press.

_____. 2011. La investigación histórica de las lenguas aisladas, o ¿es raro el vasco? *II. Congreso de la Cátedra Luis Michelana*, ed. by Joseba A. Lakarra, Joaquín Gorrochategui, and Blanca Urgell, 23-40. Vitoria Gasteiz: Editorial de la Universidad del País Vasco.

_____. 2012. The classification of South American indigenous languages. *The Indigenous Languages of South America: A Comprehensive Guide,* ed. by Lyle Campbell and Verónica Grondona, 59-166. (The World of Linguistics, vol. 2.) Berlin: Mouton de Gruyter.

_____. 2013. *Historical Linguistics: An Introduction.* (3rd edition). Edinburgh: Edinburgh University Press, and Cambridge, MA: MIT Press.

Campbell, Lyle and Terrence Kaufman. 1976. A Linguistic Look at the Olmecs. *American Antiquity* 41:80-89.

Campbell, Lyle & William J. Poser. 2008. *Language Classification: History and Method*. Cambridge: Cambridge University Press.

Farmer, Steve, Richard Sproat, and Michael Witzel. 2004. The Collapse of the Indus-Script Thesis: The Myth of a Literate Harappan Civilization. *Electronic Journal of Vedic Studies* 11:19-57.

Friedrich, Paul. 1971. Dialectal Variation in Tarascan Phonology. *International Journal of American Linguistics* 37:164-187.

Golla, Victor, Ives Goddard, Lyle Campbell, Marianne Mithun, and Mauricio Mixco. 2008. North America. *Atlas of the World's Languages*, ed. by Chris Moseley and Ron Asher, 7-41. London: Routledge.

Gorrochategui, Joaquín. 1984. *Onomástica indígena de Aquitania*. Bilbao: Universidad del País Vasco.

_____. 1993. Onomástica Indígena de Aquitania: Adiciones y Correcciones I. *Sprachen und Schriften des antiken Mittelmeerraums: Festschrift für Jürgen Untermann*, ed. by Frank Heidermanns, Helmut Rix, and Elmar Seebold, 145-155. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

_____. 1995. The Basque Language and its Neighbors in Antiquity. *Towards a History of the Basque Language*, ed. by José Ignacio Hualde, Joseba A. Lakarra, and R.L. Trask, 31-63. Amsterdam: Benjamins.

Gorrochategui, Joaquín and Joseba Lakkara. 1996. Nuevas aportaciones a la reconstruccíon del protovasco. La Hispania prerromana: Actas del VI Coloquio sobre lenguas y culturas de la Penísula Ibérica, ed. Por F. Villar y J. d'Encarnação, 101-145. Salamanca.

_____. 2001. Comparación Lingüística, Filología y Reconstrucción del Protovasco. *Religión, Lengua y Cultura Prerromanas de Hispania: Actas del VIII Coloquio sobre Lenguas y Culturas de la Penísula Ibérica*, ed. Por F. Villar y J. d'Encarnação, 407-438. Salamanca.

Hammarström, Harald. 2007. *The Languages Families of the World: A Critical Synopsis*. www.cs.chalmers.se/~harald2/.

Lakarra, Joseba A. 1995. Reconstructing the Pre-Proto-Basque Root. *Towards a History of the Basque Language*, ed. by José Ignacio Hualde, Joseba A. Lakarra, and R.L. Trask, 189-206. Amsterdam: Benjamins.

_____. 1996. Sobre el Europeo Antiguo y la Reconstrucción del Protovasco. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 30:1-70.

____. 2006. Protovasco, Munda y Otros: Reconstrucción Interna y Tipología Holística Diacrónica. *Oihenart: Cuadernos de Lengua y Literatura* 21:229-322.

Meillet, André. 1925. *La Méthode Comparative en Linguistique Historique.* Olso: H. Aschehoug & Co.

Michelena [Mitxelena], Luis [Koldo]. 1988. Sobre Historia de la Lengua Vasca, ed. by Joseba A. Lakarra. (*Suplementos del Anuario de Filología Vasca 'Julio de Urquijo'*, 10.) Donastia/San Sebastian: Diputación Foral de Guipúzcoa.

_____. 1995. The Ancient Basque Consonants. *Towards a History of the Basque Language*, ed. by José Ignacio Hualde, Joseba A. Lakarra, and R.L. Trask, 101-135. Amsterdam: Benjamins.

Saussure, Ferdinand de. 1916. *Cours de Linguistique Générale*. Paris: Payot.

Steinbauer, Dieter H. 1999. *Neues Handbuch des Etruskischen*. St. Katharinen: Scripta Mercaturae.

Suárez, Jorge. 1975. *Estudios Huaves*. (Colección Científica, Lingüística, 22.) México: Departamento de Lingüística, INAH.

Tovar, Antonio. 1957. *The Basque language*. Philadelphia: University of Pennsylvania Press.

Trask, L. R. 1995. Origin and Relatives of the Basque Language: Review of the Evidence. *Towards a History of the Basque Language*, ed. by José Ignacio Hualde, Joseba A. Lakarra, and R.L. Trask, 65-99. Amsterdam: Benjamins.

_____. 1997. *History of Basque*. London: Routledge.

Vajda, Edward. 2001. *Yeniseian peoples and languages: a history of Yeniseian studies*. London: Curzon.

Viegas Barros, J. Pedro. 2001. *Evidencias del Parentesco de las Lenguas Lule y Vilela*. (Colección Folklore y Antropología 4.) Santa Fe: Subsecretaría de la Provincia de Santa Fe.

Vovin, Alexander. 1993. *A Reconstruction of Proto-Ainu*. Leiden: Brill.

Lyle Campbell
University of Hawai‘i at Mānoa
Department of Linguistics
Moore Hall 569, 1890 East-West Road
Honolulu, HI 96822

lylecamp@hawaii.edu