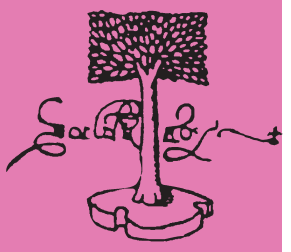
Año LII. urtea
130 - 2020
Uztaila-abendua
Julio-diciembre



FONTES LINGVÆ VASCONVM

STVDIA ET DOCVMENTA

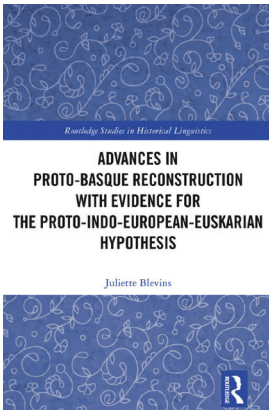
SEPARATA

Review.

Advances in Proto-Basque Reconstruction with Evidence for the Proto-Indo-European-Euskarian Hypothesis

Peter BAKKER

Nafarroako Gobernua
Gobierno de Navarra



Advances in Proto-Basque Reconstruction with Evidence for the Proto-Indo-European-Euskarian Hypothesis
Blevins, Juliette. 2018.
Routledge. 426 pages | 50 B/W Illus.
ISBN: 978-0-4295-0591-1

Erreseina/Reseña/Review:
Peter Bakker
Aarhus University
https://orcid.org/0000-0002-2072-0430

DOI: https://doi.org/10.35462/flv.130.8

1. INTRODUCTION[1]

This is a brilliant book. Yet, the author may be right or wrong. Even if the author would eventually appear to be wrong about her main thesis, it remains an ingenious attempt at demonstration of a long-distance relation between two language families/isolates, probably the best ever. In this book, grammarian, phonologist and historical linguist Juliette Blevins presents evidence for a new reconstruction of Proto-Basque, the mother of modern Basque, and argues that Proto-Basque is distantly related to the Indo-European language family.

Basque has always been taken to be an isolate, and despite serious attempts, no one has been able to convince the scholarly community of an historical connection with other languages of the world. Trask (1995) summarizes some, and he rejects all of them. Trask (1997, p. 358-429) mentions suggested connections with languages from at least 20 distinct language families, and eloquently rejects their amateuristic reasonings, providing counterevidence in more detail for the most serious ones. Indo-European is also included (p. 368-377), focusing on loans and grammatical elements. In any case, no proposal appeared to be even close to be accepted, except, in most cases, by the individual who proposed it.

Note that, according to Blevins, Basque is NOT an Indo-European language, but Proto-Indo-European (hence PIE) and Basque have a common ancestor, according to her hypothesis, hence an early form of Basque and PIE would be sister languages. Most of the proposed relatives of Basque have been found far away from the Basque homeland: the Caucasus, North America, North and West Africa, the Himalaya. Logically, the neighboring language family, Indo-European, should be the first to be taken into consideration, but that has rarely happened (but see below). My intuition had been that a connection with Indo-European and with the Caucasus, would be the most promising.

1 I thank Ross D. Kristensen MacLachlan for his help in calculating the word distances through his codes, and Joost Robbe and Jeroen Willemsen for their judgements of semantic connections.

The ideal reviewer for this book must be a superlinguist. The number of fully qualified individual reviewers of this book on our planet is probably smaller than five. A good assessor has to have solid knowledge of Basque and its dialect ramifications, about Aquitanian epigraphy of Southwestern France, about early Basque toponymy, about the sound changes that have taken place in Basque, about historical phonology in general (including long-distance geneaological relationships), about phonotactics and phoneme frequency, about the reconstruction of Proto-Indo-European, and about typology. This reviewer is not one of those very few people with knowledge of all these fields. I may have some competence in typology, historical linguistics and modern Basque. This should be kept in mind throughout this review. The ideal review would probably have to be written by a team of linguists, including an expert in the typology of sound change, an expert Bascologist and an expert in Proto-Indo-European.

Blevins worked for five years investigating her hypothesis. Her published work on sound change includes broad typological studies, as well as focused studies in a wide range of language families, including Pama-Nyungan, Austronesian (especially Oceanic), Algic, and other native Indigenous languages of the Americas. She has previously done ground-breaking work in the areas of descriptive linguistics (a.o. a grammar of an Australian Aboriginal language, 2001), the typology of phonological change (her 2004 book on evolutionary phonology) and a number of articles on sound changes in the history of California Amerindian languages. In addition, she has published a proposal about a long-distance relationship elsewhere, linking the Austronesian languages (Taiwan, Oceania, Indonesia, Madagascar) with a group of languages, the scarcely documented Ongan languages, of the Andaman islands south of India (see below).

Blevins interest in Basque was triggered in 2013 when she was invited to be an external advisor for the thesis of Ander Egurtzegi on the diachronic history of Basque, whose main supervisor was Joseba Lakarra, eminent scholar of Basque. Blevins immersed herself into the relevant bascological literature, and concluded that Proto-Basque phonology could have been very different from orthodox views. She spent years studying the materials, and began to explore possible links with Indo-European.

Blevins' thesis about the connection, which she calls a hypothesis, between Basque and Indo-European, is courageous. If this proposal is successful, she will be the one who solved the enigma of the origin of Basque. On the other hand, she may also run the risk of being ridiculed, or she may simply be ignored by colleagues. The latter seems to be the case; I am not aware of any academic reviews of the book yet. It did attract some attention in the form of conference presentations in the Basque Country, the UK, Germany and the USA, and the proposal has been discussed on several internet fora. But the most knowledgeable and qualified Bascologists have not engaged in a scholarly discussion. There is also a political dimension: the reputation of Basque as the only language having withstood the Indo-European invasion and Basque as the indigenous language of Europe is at stake.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

2. BASQUE STRUCTURAL PROPERTIES: DIFFERENCES WITH OTHER EUROPEAN LANGUAGES

Structurally, Basque is very different from its European neighbors. Here are some major grammatical differences: (1) It is the only language in Europe with ergative alignment, i.e. case marking links objects of transitive sentences with subjects of intransitive sentences, rather than subjects of both against objects. In Eurasia, such languages are found in the Caucasus and South Asia, and they are almost completely absent from Africa. Extinct Hittite (Indo-European, Anatolian branch) of present-day Turkey was the geographically closest ergative language. (2) Basque is the only European language in which verbs agree with subject, object and indirect object, whereas other European languages show agreement with the subject only. Such languages are also found in the Caucasus, Australia, Africa, the Americas and elsewhere. (3) Basque does not have grammatical gender or gender in personal pronouns, but almost all European languages, with the exception of the Finno-Ugric languages, make such a gender distinction. (4) Basque has postpositions rather than prepositions, shared otherwise only with the Finno-Ugric languages in Europe (Germanic languages have a few marginal postpositions), but many languages elsewhere have them. This property is associated with verb-final word order and other connected syntactic properties. (5) As the only language in Europe, Basque can change consonants to express a diminutive meaning. Along the westcoast area of North America it is common, but otherwise only found in Georgian, Chukchi in Eurasia (Nichols, 1971) and very few other places (West Flemish; Willemsen and Robbe, ms.; Reta in Indonesia, Willemsen & Hjorth Miltersen, 2020). (6) Basque is the only language in Europe with different constituent orders in negative and positive sentences. (7) Basque is the only language in Europe with light verbs, which take much of the inflection. This is common in Australian languages. (8) Basque has many phonaesthemes and ideophones, at least compared with other European languages. (9) Basque is the only European language, and only one of a hundred or so among the world's 7000 languages, with a genderlect, or, in the Basque case, an almost unique system of gender-distinct allocutive verb forms, in which the gender of the interlocutor is marked with a verbal suffix (De Rijk, 2008, chapter 29; Antonov, 2015). (10) Embedded sentences are created through nominalization of inflected verbs, to which nominal case markers are added, to indicate the semantic relationships between the two clauses. This is unique in Europe, but not uncommon and found in many languages with complex verbal morphology elsewhere in the world.

A language that is so different in structure is not likely to be related to other European or Eurasian languages. Not only the grammar, but also the lexicon is very different, as can be seen in the numerals 1-10: *bat, bi, hiru, lau, bost, sei, zazpi, zortzi, bederatzi, hamar,* or any other part of the basic vocabulary. The numerals *bi* 'two' (as in *bi-lingual*), *sei* 'six' and *zazpi* 'seven', however, may be reminiscent of Indo-European numbers.

Not surprisingly, many people have attempted to find sisters or ancestors of the Basque language elsewhere. The Basque language has been claimed to be related to Berber in Africa, Caucasian languages near the Black Sea and the Caucasian mountain range, as well as Burushaski in the Himalaya, Amerindian languages—to mention just

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

the most serious proposals. Strangely enough, however, Indo-European was hardly ever mentioned, even though Basque has always been surrounded by Indo-European languages: today Romance languages Gascon, French and Castilian, earlier also Celtic and Germanic languages, beside vulgar Latin in the early common era. Obviously, having been surrounded by Romance languages for almost two millennia and having been in contact with Germanic, Celtic and Romance, there are layers of loanwords from those languages.

3. PROPOSED RELATIONSHIPS WITH BASQUE

The enigma of the origin of the Basque language has fascinated many people. It is the only one in Europe that does not belong to the Finno-Ugric family (Finnish, Saami, Hungarian), the Turkic family (Turkish, Gagauz, Karaim), Afro-Asiatic family (Maltese) or the Indo-European family (the rest).

Many people who have attempted to prove that Basque is related to other languages, were amateurs in all senses of the word, but luckily not all. Also professional linguists have made more serious attempts—many of them appear to be somewhat unorthodox linguists, however. Lakarra (2017, p. 60) mentions Schuchardt, Uhlenbeck, Trombetti, and Tovar as examples. Schuchardt went against the Zeitgeist of the Young Grammarians, who defended the dogma of the exceptionlessness of phonological changes (sound laws), but he was otherwise an ingenious scholar of languages, and he tried to connect Basque with the extinct Iberian language (1907), partly documented in epigraphic and other short texts of the Iberian peninsula. Elsewhere he explored possible connections of Basque with African languages (1913, 1914). Uhlenbeck's work on Basque is of more uneven quality (see Bakker, 2011; Bakker & Hinrichs, 2011). Uhlenbeck went on a personal quest to understand the diversity of the world's languages. He discussed possible links with Caucasian languages and Berber, which were the most promising for him (e.g. Uhlenbeck, 1946). Trombetti's book on Basque (1925), connecting the roots of Basque to a great number of languages from the whole world, is characterized by creative fantasy and should be repudiated. Tovar's work is extensive and difficult to summarize. His 1980 book discusses many of the proposed connections between Basque and other languages.

Among long-rangers, a connection of Basque with Caucasian languages is the most popular, and the best argued proposals (e.g. Bengtson, 2016; see also 2017, 2020) endeavor to reveal deep genetic links between Basque and one or more of the three families of Caucasian languages. The connection of Basque with the Caucasian languages (divided by specialists into several families) is based on both grammatical and lexical similarities, and regular sound correspondences have been proposed. But as there are (at least) three families of languages in the Caucasus, with different levels of lexical and grammatical diversity, the number of potential cognates to choose from is obviously enormous, with 30+ languages. Bengtson proposes some regular sound correspondences, one of the criteria required for the proof of a genetic relationship. Bengtson's work is certainly intriguing.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Some people have proposed that many European and Asian languages are related in a supergroup called Nostratic (for discussion, see e.g. Salmons & Joseph, 1998; Renfrew & Nettle, 1999; Bomhard, 2008). Sometimes the Nostratic proposals are presented as part of a broader proposal including also languages of the Himalaya and even North America. Strikingly, Basque is not usually part of Nostratic or similar families in those proposals, which is mostly based on presumed or observed lexical similarities between the languages. The lack of connection with Nostratic indicates that long-rangers normally do not perceive similarities between Indo-European and Basque. Also Greenberg, who conjectured that Indo-European, Uralic, Yukaghir, Mongolian, Japanese and more were all part of one Eurasiatic superfamily (2000, 2002), did not include Basque is this group. And that makes Blevins' proposal all the more remarkable, as she does link Basque with Indo-European.

Blevins' hypothesis is internally consistent, and, as far as I can judge, consistent with the available evidence. If her proposal will be accepted, it will be a very unusual development in the study of language relationships. It is in fact quite rare that isolates, or small language families, are successfully connected with other language families. In the past century or so it may have happened only a handful of times, and it may take a long time for the connection to be accepted (I will give some examples below).

In an evaluation of long-distance proposals for Basque, Lakarra (2017) pointed out, «none of them has achieved the standards demanded by the comparative method, and above all, they have not achieved the objectives of diachronic comparison; namely, such attempts have been of no use when it comes to illuminating aspects of the structure and evolution of the language, and therefore, they are inadmissible by the comparative method» (Lakarra, 2017, p. 62). In other words, in Lakarra's view, following the common standards in historical linguistics, none of the proposals are acceptable as proof using the accepted burden of proof for language relationships. Does Blevins' proposal fair better?

4. BLEVINS USE OF DATA

Blevins uses a very broad range of data in her assessment, from inscriptions, to historical developments, dialectology and typological knowledge. She has studied data from inscriptions in the extinct Aquitanian language, the only language that has been proven to be related to Basque. There are some funerary and votive texts from the centuries around the beginning of the Common Era, all in all some 300 to 400 identified roots, mostly names of people and deities. They are found in the current Basque speaking areas, and up to several hundreds of kilometers from there.

Blevins also uses data from modern Basque dialects, and she uses dialect variants and dialect-specific phenomena to argue for her analysis. In her comparisons, she also includes the more deviant dialects such as the extinct varieties of the Roncal valley, still spoken until a few decades ago.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

The current standard view on the historical phonology of Basque builds upon the pioneering work of Koldo Mitxelena (Luis Michelena), whose ground-breaking book *Fonética Histórica Vasca* appeared in 1961 (second edition 1977). Subsequent work, for instance the work by Basque specialists like Joaquín Gorrochategui, Joseba Lakarra and R. L. Trask, followed in his footsteps and largely accepted his reconstructions of the phonological system of pre-Basque, based on careful study of Basque dialects and phonological typology. A recent overview of the reconstruction of Proto-Basque can be found in Martinez Areta (2013). Blevins, however, deviates from Mitxelena and the current state-of-affairs at several crucial points (see below).

In terms of comparisons, Blevins is extremely conservative, using Proto-Indo-European (PIE) reconstructions that are widely accepted, and adhering to the comparative method. It is clear that she wants to utilize the earliest reconstructable forms for both the Basque (Proto-Basque, or PB, in her own reconstruction rather than the one of Mitxelena and his successors) and Proto-Indo-European parts (based on orthodox sources) of the comparison—as is desirable in this type of research.

5. EVOLUTION AND CHANGE IN LANGUAGE

There are several areas in which languages evolve, at different levels of analysis: phonemes evolve leading to sound changes, assumed to be regular; the roots may survive but often they will have changed their forms, sometimes even radically, to a point of unrecognizability; the roots of the language can fall into disuse and thus disappear; another root may take over (part of) the semantic domain; the meanings of the roots that are still in use will often change, sometimes substantially; finally, the grammatical system of the language may evolve, in some cases radically. Thus, evolution in the form of small incremental changes take place at the phonological, lexical, semantic and grammatical levels.

Blevins' proposal relates to phonological changes and thus continuation of roots in different forms, as well as semantic changes, where meanings of roots have evolved, sometimes radically. Her proposal does not deal with grammatical changes, or only marginally. It seems that specialists agree that Basque grammar, so different from the neighboring languages, is relatively young, perhaps a few millennia old (Trask, 1997).

The choice to focus on phonological and semantic changes is of course a legitimate choice, and more work on the reconstruction of the grammar by Blevins can be expected. Indeed, in new publication, Blevins (2020a) presents further support for her view of derivational morphology in the Proto-Basque noun.

6. ACCEPTED LONG-DISTANCE PROPOSALS

It has rarely happened that long-distance proposals were accepted. In the Americas, for instance, some 184 geneaological linguistic entities are found, of which 90 are language families and 94 are isolates (https://glottolog.org/glottolog/family). In North

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

America, 76 genealogical units are identified, and 31 of them are isolates. All of them have been known, to some extent, since the early 1900s at the latest. Only three or four isolates or small language families were joined with other families in the past century or so in that language-rich continent. Wiyot and Yurok, the two languages of the small Ritwan family of California (identified as related in 1913), were first proposed to be related to Algonquian languages, spoken mostly in the Northeastern and Northern parts of North America, by Edward Sapir in 1913. His proposal, however, was met with skepticism at the time. Almost all rejected the connection (see Bakker, 2009b). Yet, it was ultimately accepted, but first more than 60 years later, after Algonquianist Ives Goddard vetted it, in his seminal (1975) article on the connection (Poser, 2003; Bakker, 2009b)—even though Goddard rejected a large part of Sapir's proposed lexical cognates. Wiyot and Yurok are now universally accepted as sisters of Algonquian in the Algic superfamily.

Similarly, in South America, a rare demonstration of a small family being part of a wider family was achieved when Ribeiro and Van Der Voort (2010) made it very likely that the languages of the small Jabuti family were part of the large Macro-Je (super) family. In this case as well, it had taken many decades before its proof was accepted, 80 years after it had first been suggested by Kurt Nimuendaju in the 1930s.

More recently, Brown et al. (2011) proposed a geneaological connection between two Central American language families, Totonacan and Mixe-Zoque. In 2014 a language of Louisiana in North America, Chitimacha, until that time considered a language isolate, was added to this Totozoquean group, which seems to be accepted now (Brown et al., 2014).

These are the only three or four proposals that have been accepted by the scholarly community. Campbell (1997, p. 6-329) devotes two long chapters of his book about American Indian languages on the discussion of proposals for «distant genetic relationships» and evaluates them, taking into account their probability and the level of confidence, but hardly any of these ca. 50 proposals were acceptable in his eyes, let alone proven.

In short, distant geneaological proposals are rarely accepted when first proposed, and if they are, the recognition often happens first many decades later. These four successful proposals mentioned above, are the only ones from the Americas in the past century, but many more are not accepted. Note that in all four cases, the established relations cover non-adjacent areas: the related languages are spoken relatively far from one another. As an illustration of unsuccessful proposals, we discuss briefly three well-motivated proposals that have not led to acceptance.

The proposed genealogical connection between surviving Ket of the small Yeniseian family in Western Siberia and the widespread Athapaskan (Na-Dene) languages of North America, propagated mostly by Edward Vajda, is well-argued and it is accepted by some (see the papers in Kari and Potter 2010), but rejected by others (see e.g. Campbell 2011). Evidence is presented both at the lexical and grammatical levels (e.g. verbal affix order).

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Bakker has argued for a genealogical connection between Salishan languages, the isolate Kutenai and Algonquian languages (2006, 2007, 2009a, 2013), first suggested by Sapir. The total lack of reactions (with the sole exception of Van Eijk, 2007, who is unconvinced) indicates an outright rejection, or skepticism, or, in the best case, the lack of cogent counterarguments against the proposal from the part of the adversaries, leading to silence.

Outside of the Americas, Blevins herself (2007) has proposed a daring genealogical connection between languages of the Andaman Islands (South of India), united in an Ongan grouping, and Proto-Austronesian, which has not been received positively (see below).

In short, serious proposals on deep genealogical connections are uncommon, and the acceptance rate is low, and if a proposal is ultimately accepted, it is often only decades later that others are convinced. Proposals to connect Basque with other language families and isolates have thus far all failed to convince the scholarly community. I will discuss a few recent ones.

7. SOME OTHER LONG-DISTANCE PROPOSALS FOR BASQUE

Five years before Blevins' book appeared, a proposal linking Basque and Proto-Indo-European was published in the *Journal of Indo-European Studies* (Forni, 2013), with peer commentaries. The comments were rather devastating for the proposal. Gorrochategui and Lakarra (2013), for instance, commented his ideas in an article called «Why Basque cannot be, unfortunately, an Indo-European language», as the work of an amateur, not mastering the basic principles of comparative linguistics, and call their own judgment «harsh». Kassian (2013, p. 197) likewise classified the proposal among those of «amateur linguists» with «negative results» based on traditional, lexicostatistical and probabilistic arguments. Bengtson (2013) called Forni's assumptions «highly questionable», full of misunderstandings about the methods of genetic linguistics and his verdict, as a scholar devoting his research time to long-range connections of Basque, is also decidedly negative. Blevins does not discuss it in detail but she agrees with the others that Forni does not make use of international reconstruction and the comparative method (2018, p. 82, 147-148, 150). No trained linguist endorsed it.

Bengtson (2020) attempts to find phonemic regularities with Caucasian languages, and he concludes that lexical similarities between the North Caucasian languages are promising, but a connection with Kartvelian (South Caucasian) is to be excluded. The detailed data on the lateral consonants presented by him are certainly intriguing and suggestive. A more detailed discussion of the Basque-Caucasian hypothesis can be found in Bengtson (2017).

Jaime Martín, specialist in Romance languages, proposed in the past decade (e.g. 2014) that Basque is related to Dogon, a language spoken in Mali, Africa. His proposal received quite a lot of attention in the Spanish press (see e.g. http://www.semevadelalengua.es/?p=154). The Dogon varieties are so diverse, that it would be better to speak

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

of perhaps five distinct languages. Dogon has been suggested earlier to be part of the vast Niger-Congo phylum, but more recent expert classifications do not accept this connection (glottolog.org), and consider Dogon an isolate or a small family. In unpublished work, I compared 40 supposedly conservative words from six different Dogon varieties with both modern Basque and modern Spanish. The results showed that there were forms resembling Spanish in 25 of the 40 words, and forms resembling Basque, using the same criteria, in only 9 of 40 words. Thus, lexically, Dogon is more similar to Spanish than to Basque. That should be enough to reject the connection. The grammatical similarities are not convincing either: the shared SOV order is the most common constituent order in the world, and other orders correlate strongly with verb-final languages (genitive-noun, noun-adjective, noun-demonstrative, postpositions). The connection has to be rejected. Specialists have not responded to the preposterous claims in academic work, as the proponent is also obviously ill-informed about the principles of historical linguistics and typology.

8. ISOLATES

Basque is not the only isolate in the world. According to Glottolog.org, there are currently 188 identified isolates on planet earth. Basque is undoubtedly the best known of them, and the only one spoken in Europe. There are more than 400 language families and isolates in the world. Dialects diverge and become more different, so they evolve into distinct languages, within the same family. Languages or language families can diversify to such an extent, that it is no longer detectable after, say, ten millennia, that languages descend from the same original language. That is one obvious way that isolates emerge. And that is likely to have happened to Basque.

If one would assume that language emerged once in the history of the planet, then it can only mean that all 429 identified groupings are ultimately related, and that many splits took place, starting many millennia ago. So many changes have taken place, that one cannot perceive anymore that these 429 groupings actually go back to a single point of genesis, or at most a few. In that case, it just means that our methods are limited, knowing that languages evolve so rapidly, that all traces of a proto-stage have disappeared after, say, 10 000 years, implies that our horizons are not wide enough and our time depth too shallow to enable us to discover that two or more of these groups could be related. As all humans living on earth descend from one man who lived some 250 000 years ago and one and the same woman who lived some 150 000 years ago, there are *at least* 15 to 25 total cycles of related languages that have become so different that common descend is no longer discernible—assuming that these mitochondrial Eve and Y-chromosomal Adam spoke a language, and that is highly probable.

The alternative would be to accept that language emerged independently at least 429 times (the number of currently known language families and isolates) on earth, and that is utterly unlikely. Especially taking into consideration that all now-living humans likely descend from a small group of a few dozen individuals who lived some 55 000 years ago in Africa. There is thus no doubt, that many former language families and iso-

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

lates have disappeared from the globe before they even could be documented, including dozens of name-given languages that we know of and that have gone extinct in the past millennia. And there is no doubt that isolates are more likely than not to have relatives, but the relatives are not detectable because of an early split.

The truth of the origin of languages, is probably closer to one or very few independent geneses of languages, rather than hundreds of them. The surviving languages that are spoken today, have all changed so dramatically within six to ten millennia of development from their ancestor language, that the remaining similarities between descendant languages are very few. Thus, the shared words and constructions between language families, do not allow identification, since the numbers of surviving roots are below the level of chance. There will always be 3-10 % accidental similarities between unrelated languages (numbers depending on how far you want to stretch perceived semantic similarities). Any connection on the basis of shared words is thus unprovable after 6000 to 10 000 years.

As for Basque, this would mean that Basque split off from the language it is most closely related to, whatever that may be, at least six to ten millennia ago—the limit that historical linguists have set for being able to detect a signal of shared history in the lexicon, based on extrapolations on dates for known splits. There are simply no discernible signals that can be discovered with the tools and standards that are available for lexical comparisons. Is Blevins able to convince the scholarly community that there is a genealogical shared history between Basque and PIE?

9. THE BOOK

The book consists of two parts, and a long appendix. The first part is a reconstruction of Proto-Basque, at the phonological level (p. 1-122). Part two is a lexical comparison between Proto-Basque and Proto-Indo-European (p. 123-212). Both parts start with lucid introductions (p. 3-12) and (p. 125-132) on the topic, that make the material accessible to general readers. The appendix that links Proto-Basque and Proto-Indo-European roots fills almost as much space (p. 213-386). Note that the study is limited to phonology and word roots, and it does not cover much grammar, except occasionally (the shape of) some affixes, Ablaut/grade forms and some derivational processes. The bibliography (p. 387-402) covers some 400 titles. There is a subject index (p. 403-406), a word index (p. 407-425), with separate lists for Proto-Basque, Aquitanian, medieval Basque and modern Basque, and for the different documented Indo-European languages (notably Albanian, Latin, Hittite, Sanskrit) as well as for reconstructed Proto-Indo-European roots.

Chapter 1 (Basque and Proto-Basque, p. 13-25) provides background knowledge about sources, dialects and orthography. It also outlines the reconstruction of Proto-Basque phonology. Chapter 2 (The Proto-Basque Vowel System, p. 26-36) deals with reconstructed vowels, with different ideas about the diphthongs. Chapter 3 (A Revised Proto-Basque Consonant System, p. 37-82) is more radically different from previous reconstructions, and presents a less typologically unusual system than hitherto proposed

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

(see below). Also in Chapter 4 (Proto-Basque Phonotactics, p. 83-99), a revision of orthodoxy is invoked with more complex syllable structure and clusters than hitherto proposed. Chapter 5 (Proto-Basque Stress and Accent, p. 100-116) in the same way pleads for innovative ideas. Basque and Proto-Basque accent and stress have been subject of discussion for a long time. Blevins' reconstruction has both quantitative stress assignment and lexical accent. Chapter 6 (Advances in Proto-Basque Reconstruction, p. 117-122) ends Part I and it summarizes the findings and lists the reconstructed phonemes, segmental sound changes and the new proposal is compared with previous systems.

Part II (Comparison of Proto-Basque and Proto-Indo-European) starts with an introduction (p. 123-132) and Chapter 7 (Results of the Comparative Method, p. 133-150) then provides regular sound correspondences, illustrated with basic vocabulary from Proto-Basque (PB) and Proto-Indo-European (PIE). In Chapter 8 (Statistical Evidence for Relatedness, p. 151-183), Blevins uses statistics to argue that it cannot be chance that there are similarities between PIE and PB. She also compares the well-known vowel-grading of PIE and finds reflexes in PB. Chapter 9 (Proto-Basque and Proto-Indo-European Historical Phonology, p. 184-200) she describes implications for Hittite, and several phenomena know from Indo-European with traces in Basque, including continuity of laryngeals. In chapter 10 (Potential Implications for Indo-European Linguistics, p. 201-212), the new reconstructions and the repercussions these may have for studies of Indo-European are discussed. The rest of the book lists PB reconstructions and the parallels in Indo-European in detail, including the modern Basque forms that are derived from the protoforms. It is impossible to do justice to the detailed reasonings in the book in this discussion of the book, especially for one who does not have the same level of knowledge of the two language groups. Interested readers, and that should be anyone working in the field of Basque historical linguistics or Indo-European, must read the book.

Blevins' reconstructions of Proto-Basque phonology and root structure differ from other proposals, and her novel reconstruction of the phoneme inventory, she claims, «yields roots and grammatical morphemes that show regular correspondences with Proto-Indo-European constructions» (p. 3). Hence, she hypothesizes that «Euskarian», the reconstructed ancestor of Basque (including consideration of its only known relative, extinct Aquitanian) and Proto-Indo-European were sister languages. She presents it as a hypothesis, not as an established proven fact. Such reservations and modesty distinguish the true scholar from the cranks and amateurs. Yet, reading the book, one gets the impression that the author considers her hypothesis confirmed.

The research is well-informed and full of references to earlier studies of the Basque language and its history, including: the history of Basque, including medieval Basque, and Basque dialectology, loanwords in Basque, Aquitanian epigraphy, and Indo-European studies. This reviewer does not have the necessary expertise to spot gaps and errors in the detailed data. Blevins indicates explicitly that she follows the well-established principles of the comparative method and internal reconstruction, but she makes clear that «some aspects of the sound system and certain sound changes» (p. 12) are novel. Indeed, she does not accept Mitxelena's widely acclaimed reconstructed phoneme inventory of Proto-Basque.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Where does she differ from current consensus about reconstructed forms for Basque? First and foremost, she accepts a phoneme *m for Proto-Basque, which Mitxelena/ Michelena (1977) suggested had not been present, and many others since have accepted that. Second, in contrast to more traditional views, she reconstructs *p(h), a voiceless bilabial stop. Third, she argues that initial voiceless laminal sibilant <*z*-> in Basque derives from a sibilant followed by an aspirated stop (*stH, *spH, *skH). Thus, she reconstructs a single sibilant rather than two or four, and three aspirated stops. Fourth, she only reconstructs one rhotic, not two. This makes Proto-Basque look less exotic than the current reconstruction by Mitxelena and others. Further she has different views on Basque phonotactics. Whereas others take CVC as the canonical root in Basque, Blevins has both monosyllabic and disyllabic roots, and CC combinations.

The sound changes from Proto-Basque to modern Basque identified by Blevins are summarized on pages 120-121. Blevins proposal is typologically unremarkable, whereas the traditional system is typologically remarkable, for instance in lacking the cross-linguistically extremely common consonant /m/.

10. BLEVINS' RECONSTRUCTIONS

The reconstructions are limited to stems and words, mostly nominal and verbal, but inflection is not studied. Basque has a rich case system and a complex verbal morphology. Some derivation is discussed, and elaborated in Blevins (2020a). Only a few grammatical elements are dealt with at this stage. The detailed appendix (p. 213-387) presents the reconstructed Proto-Basque roots, enumerating its descendants in modern Basque and its proposed cognates in Proto-Indo-European, and sometimes also information about the processes of change. The appendix contains, if I counted right, 445 starred lemmas with Proto-Basque forms and suggested PIE (sometimes archaic Indo-European languages instead) roots. That is, it should be said, an impressive proportion of the lexicon in this kind of work.

11. BLEVINS' PROPOSAL

We mentioned above that dozens of genealogical connections have been proposed for Basque, all of them unsuccessful. In what respect does Blevins' proposal differ from other long-range proposals for Basque?

First, she does not compare Basque with a range of languages (e.g languages from several Caucasian families, all in all covering 30+ fairly diverse languages), but with just one, a reconstructed language PIE. Occasionally she also uses extinct IE languages Hittite (Turkey) or Vedic Sanskrit (India), old written IE languages, but that is always because those languages provide information otherwise not recoverable from PIE. Occasionally she spots potential cognates not in PIE, but in more 'marginal' branches of the family, such as Albanian, Armenian or Baltic.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Note that it is much easier to find potential cognates of Basque in a *group* of languages, like Caucasian languages, than in one *specific* language, PIE, from a statistical perspective: with many languages (in the Caucasian case, around 35) to choose from, the chance of finding a form that resembles a Basque form, is high. Thus, it is more challenging to prove a one-to-one connection, and that makes the enterprise much more challenging than in proposals in which only two languages are compared. We have to stress that almost all of Blevins' forms do not show just accidental resemblances, but forms connected through regular sound correspondences.

Second, for each form that she cites, she provides a clear source, referenced with page numbers or item numbers, and thus checkable. That is not always the case among the long-rangers. In my experience, it is often difficult or impossible to check much of the data provided by long-rangers. Sometimes material seems to be invented, but Blevins sticks to reality. This is also indicative of the quality of the proposal.

Third, she uses reconstructed forms, rather than modern forms, and hence the oldest possible identifiable forms, and these are obviously closer to the original forms. Here we can add that she uses PIE reconstructions made by the specialists, rather than her own. For Basque, however, she makes her own reconstructed forms, and that is necessary because her ideas about the phoneme inventory and root structure differ from established scholarship. This could be a weak point since a researcher may want to reconstruct that form that fits any one of the suspected cognates. These seem to be consistent though radical.

In short, Blevins' proposal is more restricted in its choice of languages of comparison, using the oldest possible (documented or reconstructed) forms, the sources are more controllable and hence more scientific, and the reconstructed forms do not seem to be ad hoc.

As I lack the detailed knowledge to assess the specific reconstructions, I will discuss a handful of ways to evaluate long-distance relationships.

12. SIX QUALITY TESTS FOR LONG-DISTANCE GENEALOGICAL RELATIONSHIPS

Long distance relationships are often rejected for one or more of the following reasons: undetected loanwords have been included in the comparison. Unidentified neologisms are used. The semantic connections between the roots are too broad to make sense (semantic lumping). Ghost words have been included. Words have been segmented wrongly, and parts have been chosen that do not have the meaning that is connected with the other language. Essential parts of words (e.g. glottal stops, tones) have not been included in the comparison. The number of proposed cognates is too low. There are no or too few cases of regular sound correspondences. And, beyond linguistics, the geographical distance between the languages is too great.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Several of these have been raised as objections against Blevins' proposal. A group of scholars mentioned that they found problems with around 80 % of the proposals in their systematic sample of words pairs proposed by Blevins (Ariztimuño, Zuloaga & Krajewska, 2019).

12.1. Test one: calculating formal distances between proposed cognates and non-cognates

One test to establish the solidity of proposals of language relationship was proposed by Oswalt (1970). It was proposed when computerized methods were uncommon, but today it is much easier to do. The idea is roughly this: if you have a set of proposed cognates between two languages, with one (more or less shared) meaning and two forms (one in each of the languages), one can statistically establish the likelihood of this being right, by shifting one of the forms one row, and recalculate the level of similarity, and do this as many times as one has forms in the set. If the proposed set of cognates is correct, then the similarities between the forms with the same meanings in L1 and L2 can be calculated. The proposed cognates can be compared with the other forms, with which they are not cognates. One should thus expect the root forms of the proposed cognate sets be more similar in form than the root forms of the non-cognate sets. Dunn & Terrill (2012) used a variant they called the Oswalt Monte Carlo Test, where the word pairs were randomized. Dunn (2015) also discusses it briefly. It is said not to work very well for long distance relations because of the disturbing effects of loanwords. As they tested it with contemporary and adjacent languages, we would run less risk with two languages so distant in time and apace as PB and PIE.

Blevins (2018, p. 173) argued that this method was unsuited for Basque, but encouraged the readers to use it on her material. We took up the challenge. (In the meantime, however, Blevins has done it herself. See the afterword. This was written without knowledge of her on-going work with Richard Sproat; see Blevins and Sproat, to appear).

We did the test for the proposed cognate sets between Basque and PIE (I thank Ross D. Kristensen MacLachlan for his help). We chose the first 100 appropriate proposed cognate pairs in Blevins list for scrutiny. We used Levenshtein distances in (Levenshtein, 1966) order to calculate the distances in the forms of the words. Simplified, Levenshtein distance measures the number of steps to get from a word to another word with the same meaning, by measuring deletions, additions and substitutions and thus the number of steps to realize the change. Thus, it would take just one step to get from Dutch *hond* to German *Hund* (o > u), both meaning «dog», but at least three steps to get from *hond* to *dog*: replacement of h by d (several steps in reality, e.g. h > x > k > p > b > d), the deletion of <d> and the change of <n> to <g> (e.g. n > ŋ > ŋk > g). The lower the number of steps, the more similar the forms are, and thus more likely to be possible cognates—if the meanings match.

We choose only words for which both a PB and a PIE form were available (thus excluding for example those where Sanskrit or Hittite forms represented Indo-European). We also excluded pairs that had question marks. Our calculation should emphatically only be taken as a very rough indication, in that we had some double forms and some

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

diacritics that were counted as signs by the computer, and similar confusing factors for the machine. The first word included in the comparison was PB *ágo* (p. 217), the last word PB *hankha* (p. 276). As the words are arranged alphabetically, one can expect a somewhat lower diversity of roots, because the list will contain e.g. all roots starting with *d-* or *g-*, but none starting with e.g. *s-* or *t-*. Thus, the roots can be expected to be more similar than they would have been if all roots had been taken into consideration.

We first found the pairs of words with the shortest Levenshtein distances. One would expect those pairs to be the proposed cognates to have the shortest distance, and thus more similar. That was the case in only 17 % of the cases, which is quite discouraging for the proposal. However, double forms, diacritics and other interfering factors indicate that this number is not to be taken too seriously.

In a second step, we used a variant of the Oswalt method. The Levenshtein distance of all the proposed cognate pairs was calculated, and also the Levenshtein distance between each form from one language, with the 99 non-cognates, and the latter were averaged. When more than one form variant was mentioned, we arbitrarily chose the first one. One should expect the Levenshtein distance between the cognate pairs to be lower than between the non-cognate pairs. The average Levenshtein distance for proposed cognate pairs is 37.8, and for the other pairs 50.0. This indeed appeared to be the case for an overwhelming number of forms. In 87 % of the pairs, it was the proposed cognate pair that was closest or equal. Even though it is probably inflated, this number is actually surprisingly high. It may be interpreted that Blevins is on the right track. Her proposed cognate word pairs are overwhelmingly more similar in form than non-cognates, even though not the full range of cognates was included. But what about the range meanings of the proposed pairs, and the non-pairs?

12.2. Test two: semantic broadness

It is much more difficult to calculate differences between meanings than differences between forms.

Whereas forms of morphemes can be compared through automatic measuring, this is actually not possible with meaning. The broadness of meanings and the differences of the meanings is always a bone of contention in proposed long-distance relationships. It was also a criticism raised against Blevins' proposals for Proto-Ongan-Austronesian (POA) and for PB and PIE. Criticisms can be divided into two types: the semantic link between proposed pairs is far-fetched, and/or the semantic range of one of the forms (with several meanings) is so broad that several of the meanings can be linked to other forms that perhaps also have several meanings. An example of the first is the proposed cognate set with the meanings 'small round object that makes up a multitude' (PB) and 'bee (insect)' (PIE). Two examples of the latter are PB 'to give away, deprive oneself of' linked with PIE 'to share out, to apportion', and PB 'block, obstruct, bar, enclose, fence in; stop' with PIE 'stop, ward off, defend'.

In order to measure somehow the semantic distance between the proposed cognates, the meaning sets of the same 100 cognate pairs, without showing the forms,

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

were independently assessed by three linguists. They were asked to match the two columns and find the meanings that were closest for the two languages, after I had alphabetized the meanings for both columns. This matching is a somewhat time-consuming task, especially when several meanings or meaning sets are given, sometimes up to five. By having three raters, one can estimate the interrater validity by measuring how deviant the scores could be. The number of pairs for which the same meanings were matched by all three was 28 %. Pairwise rates were 48, 42 and 38 %. These are somewhat disappointing numbers. However, these are lower than the real numbers, as a few had identical meanings, and then the choice is arbitrary, and we did not take that into account. These percentages should only be taken as impressionistic. The resulting figures are rather low, with less than a third meaning matches.

12.3. Test three: automatic classification on the basis of the most stable words

Wichmann and his colleagues hypothesized that one could create a tree of the world's languages based on stable words, or actually form-meaning combinations. They tried to create a tree of all the documented languages of the world on the basis of a small word list (see Wichmann et al., 2020). In the first global and automated comparison of the world's languages based on the 40 most conservative words of the Swadesh list, however, no link between Basque and Indo-European was detected. In the global tree of 2010, 4350 of the world's languages and dialects were compared automatically. In that tree, Basque appeared to be most similar to languages of Peru, the Amazon, Papua New Guinea, Kainji languages of West Africa and two isolates of Africa, Fur and Shabo (both Nilo-Saharan, according to some) (see Müller et al., 2010; Basque can be found on p. 40). As lexical signals of relatedness disappear quite quickly, this should not be taken too seriously, and relegated to chance similarities. Clearly, this has not led to the discovery of a language that Basque is potentially related to IE, and Basque and Indo-European languages appear very distant.

12.4. Test four: finding additional cognates on the basis of proposed sound changes

One way of testing a proposal is to apply sound changes in order to find potential other cognates, for instance by taking forms, and see whether similar forms or meanings appear after applying observed sound laws. We will try that by taking modern Basque roots as our point of departure.

Blevins proposed that Basque words starting with *z-* can be derived from PIE words starting with a sibilant or a cluster of a stop consonant (aspirated or not) preceded by a sibilant. Blevins proposes Basque initial <z> could come from /s/ or from *sp^{h}, *st^{h} or *sk^{h}, and medial <z> from these clusters or from *rs.,*ls.,*ns. Thus, the prediction would be, that Basque words starting with a sibilant, are most likely to have a PIE equivalent starting in a sibilant, or sC clusters.

Taking Trask's manuscript etymological dictionary (2008) as a point of departure, one could select all Basque words starting with <z>. Then one can look for meaning equivalents of these words, and if the hypothesis is correct, one should expect to

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

find Indo-European roots with initial sibilant stops to have statistically often with sibilant-stop clusters, having a comparable meaning. For this, we looked only at forms that are not borrowed, limiting us to forms marked OUO «of unknown origin» in Trask (2008), and look at possibly emerging patterns. The list is given below. The words on the left side are modern Basque words as found in Trask. On the right side, I first give Blevins' equivalent in PIE, or in other IE languages (OCS = Old Church Slavonic). The information found by me follows, preceded in all cases by «Cf.» It will appear that indeed quite a few start with the predicted segments.

It should only be considered as what it is: an impressionistic view based on a mixture of sources. I start with Blevins, and then add other potential sources, without attempting to adhere to sound changes beyond the first consonant. If no meanings are indicated, they are the same as in the first column.

zahar 'old'	Blevins 159 OCS *starŭ* 'old'.
zabal 'broad'	Blevins 332: PIE **pleh2-* 'be wide, flat' [LIV3]. Cf. Proto-Germanic **braida-*, possibly linked with Proto-Germanic **spraidijan-* 'to expand, disperse'.
zail 'tough'	Blevins 353: PIE **(s)kel3-* 'crooked; bent, curved' [CW:80]. Cf. PIE **(s)ter-g-* 'to stiffen'.
zain 'guard, watchman'	Blevins 367: PIE **(-)stah2-n/o-* [NIL:641], [CW:86] 'place' < 'where one stands'.
zakil 'penis'	Not in Blevins. Cf. German *Stiel*, *Stengel*, *Schaft* 'handle, stem, shaft', PIE *steg/stengh* 'be stiff', Old High German *stachilla* 'pointed growth'.
zakur 'dog'	Blevins 305: PIE **kʲuō(n)* 'dog'. Cf. also words for 'dog' like Portuguese *cachorro*, Old Indic 3219 **kuccura* 'dog', Sanskrit *jukuṭa-*, Romani *džukel*.
zaldi 'horse'	Blevins 364 **sthahal-di*. Blevins 323: connected with *nar2* 'vital, strong; virile, manly'. Cf from PIE **sta-* 'to stand, make or be firm', Old High German Old High German *stuot* 'herd of horses', English stud. Cf. *stallion*, derived from a word for 'stable'. Cf. Dutch *staldier* 'animal kept in stable'.
zalke 'vetch'	Blevins 353: PIE **(s)kel3-* 'crooked; bent, curved'. Cf. also PIE **weyk-* 'to curve, bend', the source of Germanic *wikke* 'wetch'.
zata (G) n. 'nightjar, goatsucker'	Not in Blevins. Cf. Proto-Germanic *swarta-* 'black, dark', PIE 'pie. **suord-*, **surd-* 'black, dirty color' (IEW 1052).
zazpi 'seven'	Blevins 350 PIE **sep/tm*. Cf. *hatz* 'finger', *bi* 'two'.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

ze- 'interrogative stem'	
zehar 'across, through'	Blevins 353 PIE **(s)ker3-* 'to turn, bend' [CW:81].
zelai 'plain'	Not in Blevins. Cf. Proto-Germanic *felþ-* < PIE **pelt-*, **p(e)lh2-* (IEW 805) 'flat'.
zerri 'pig'	Blevins 369: *sther(h)* 'stiff, solid, massive'. Cf. PIE **pórḱo-* (IEW 841), PIE **sū̆s, *su̯os* (IEW 1038) 'pig'
zezen 'bull'	Blevins 166: Connected with PB stehen. Reduplicated form. Cf. also Proto-Germanic **steura-* 'bull', perhaps PIE **th$_2$euro-*
zi 'acorn'	Blevins 360: PIE **spei-* 'sharp point'. Cf. Lettish *zīle* 'acorn'.
zigor 'staff, rod, stick'	Not in Blevins. Cf. Proto-Germanic **stukka-* 'stick'.
zikin 'dirty'	Not in Blevins. Cf. Proto-Germanic **stankwa-* 'stench, stink'
zilar 'silver'	Not in Blevins. Cf. Proto-Germanic **silubara-* (borrowing from Basque according to Cowan, 1971).
zilegi, *zillegi*, *zilhegi* (L) 'licit, permitted'	Not in Blevins.
zin 'oath, truth'	Not in Blevins. Cf. Old Indic *śapátha* m. 'curse' RV., 'vow' [12290] cf. Proto-Germanic **swarjan-*, cf. OCS *svariti* 'to scold', PIE **suer-*
zintzur 'throat, neck, summit'	Not in Blevins. Cf. Proto-Germanic *spitja-* 'pointed object'. Cf. PIE **trud-* ('to swell, become stiff') > English 'throat'.
zizari 'worm'	Blevins 95: **sis* 'moth, moth larva' + **thari* 'thread'. Cf. Baltic *žalìti-*, *zalti* 'snake'. Old Indic **kīṭarūpa* 'worm'; *kīṭá*, *kīṭá1* 'insect, worm'.
zoko 'corner'	Not in Blevins. Cf. Proto- Germanic **hōk-*. No PIE.
zor 'debt'	Blevins 269: **gohor* 'lacking, deprived; barren'. Cf. Lithuanian *skolos* 'debt'; PIE *-rn-* 'debt'.
zori 'omen'	Blevins 361: PIE **speh1-* 'get on well, succeed' [LIV3]; PIE **sph1ro-* 'thrive, prosper' [CW:84].
zorne 'pus'	Not in Blevins. Cf. PIE **puH-* (LIV480), Albanian *qelb*, Lettish *strutas* 'pus'.
zorri 'louse'	Blevins 360: PIE **sper-* 'propagate, spread' [LIV3]. Cf. Albanian *morri*, PIE **lūs-*, *lus-*.
zorro 'bag, sack, pouch'	Not in Blevins.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

zortzi 'eight'	Not in Blevins. Cf. Old Indic *cáta·sr-* and *catvāri* 'four' + *bi* 'two'.
zotz 'stick'	Not in Blevins. Cf. Proto-Germanic *stukka-* 'stick', Proto-Germanic *strunka-* 'stump of tree'.
zozo 'blackbird'	Not in Blevins. Cf. Old Indic 12405 *śā́ri* 'bird species, blackbird'
zu 'you'	Blevins 373. PB *sTu* 'you'. Cf. *tu* 'you', widespread in Indo-European.
zulo 'hole'	Not in Blevins. Cf. PIE *halu* Cf. *tu* 'you' in many branches of IE. 'hole'.
zumar, zunar 'elm'	Not in Blevins. Cf. *zu(r)* 'wood'; Proto-Germanic **alm-*.
zuntz 'thread, strand, fibre'	PIE *ĥand-* 'thread'.
zur 'wood'	Blevins 372. pre-PIE **sth2-ur-* > **stuh2r-* 'big, strong, thick' [NIL:639].
zuri 'white'	Blevins 361. PIE **pewH-* 'purify, cleanse' [LIV3]. Cf. Old Indic *śvetá-* 'white, clear', *śvitrá* 'whitish'.
zurtz 'orphan'	Blevins 372. PIE *?*sterh1-* 'lack, deprived; be robbed' [LIV3]. Cf. Proto-Germanic **sterbanan* 'be stiff, starve', from. PIE **sterh1-* 'to be stiff, solid' [LIV3]
zuzen 'straight'	Blevins 371. PIE **sth2-* 'to stand (somewhere), to stand up' [LIV3]. PIE **sth2-ú-* 'accurate, real, correct' [NIL:637]. Cf. PIE **strenk-* 'tight, narrow; pull tight, twist' (source of English straight).
zuzi 'torch'	Blevins 95. from *su* 'fire' + *zi* 'pinecone; tip' (v. [PB] **su*, **sphi*). Cf. torch from late Latin *torqua/torquere*, PIE root **terkw-* 'to twist'
zuzun 'poplar'	Not in Blevins.

Thus, I have added a number of potential cognates to those presented by Blevins, without, however, working out the exact details of the sound laws, and without being able to judge the quality of my sources. I will leave that to others.

These rough comparisons lead us to a dilemma. One the one hand, we get surprisingly many forms where at least the first consonant fits the sound laws as proposed by Blevins for the development form potential PIE cognates towards the forms of the Basque roots. The meanings are also close. This success could probably interpreted in two ways. Either Blevins is on the right track, and the number of potential cognates is expandable, and that is a good sign for her hypothesis. Alternatively, it could be a consequence of the sound change patterns being so general, that it would be too easy to find forms in any language.

However, there is another observation one can make. Many of the new potential cognate sets in this list that were identified by me, without the rigorous application

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

of sound laws for the whole word as done by Blevins, are almost all from Germanic. Some of the forms are limited to the Germanic branch, others have a known root in PIE. Could it be that some of these Germanic words give a clue towards Indo-European roots only preserved in Germanic? Or, alternatively, could some of them be unidentified loans from Germanic into Basque? It could be worth exploring these possibilities. At any rate, whether correct or not, it is intriguing that one can find sC with meanings that are so close to the Basque forms. If Blevins is on the right track, it could also lead to revisions of reconstructions for PIE.

12.5. Test five: abstract grammatical similarities

Blevins does not deal with grammatical features (e.g. forms of affixes), or with abstract grammatical features (e.g. ergativity, presence/absence of gender, presence/absence of case) in which the specific forms are not taken into consideration. Nichols (1990, 1992, 1998) was a pioneer in the study of the stability of grammatical features, and she concluded on the basis of her ingenious typological survey, that abstract grammatical properties like alignment and gender were more stable than lexical roots, and that such features could detect a genealogical signal that was much older than with any other method. She estimated that such features could persist for 30 000 years.

In the meantime, large databases of grammatical features have been compiled, notably the *World Atlas of Language Structures* (WALS) (Haspelmath et al., 2005; Dryer & Haspelmath, 2013).

Greenhill et al. (2010) took all the features of WALS, and used a computer program to produce a network or tree of the languages of the world documented in that database. The languages of the world ended up in a neighbornet showing limited geneaological signals, and a clear Eurasian cluster. In Greenhill et al. (2010, p. 2445), Basque appeared as part of the Eurasian cluster. The language appeared to have most in common (typologically) with the Amerindian language Awa Pit of Northern Ecuador and Southern Columbia, and the Nakh-Daghestanian languages Ingush, Lezgian and Hunzib of the Caucasus. Their study was based on the whole set of 138 typological characters in WALS.

Bakker et al. (2011) used a subset of 43 features from WALS, and found Basque in a cluster with Hindi (Indo-European, India/South Asia), Burushaski (isolate, North Pakistan) and Hunzib (East-Caucasian, Caucasus) (boxed on the right side of the network in Figure 10 on p. 34 in the original publication). In subsequent unpublished and highly speculative work by Bakker, he selected the most stable grammatical features, and found Basque to have most in common with languages associated with the first and second of four waves of global migrations out of Africa as discussed in Nichols' model, notably associated with Australia, Papua New Guinea, the Southeast and the east of North America, the Northeast of South America. Those migrations (Nichols, 1998) took place during glaciation (70 000-10 000 years ago) and late glaciation (ca. 13 000 years BP), which would indicate a split-off date of at least ten millennia ago. Nichols' study focused on the New World, and Indo-European did not play a role.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

These three studies based on abstract grammatical features find North Caucasian and Himalayan languages to be most similar to Basque, and a more speculative approach based on stable features led to languages of the Americas and Australasia as the most similar ones. Hindi was the IE language closest to Basque grammatically.

12.6. Test Six: comparison with another Proto-language

One can also test long distance proposals by checking whether the proposed reconstructed forms and meanings are also linkable to other proposed protolanguages. If these protolanguages are beyond reasonable geographical range, similarities are more likely due to chance. The forms of the reconstructed roots tend to be short, and the meanings broad, and that is also the case in Blevins' book. If it is just as easy, or even easier, to link her PB/PIE reconstructions with other protolanguages, the results could be attributable to chance.

Blevins has previously attempted to prove a genetic relationship between languages of the Andaman islands and Proto-Austronesian. We can use that reconstruction as a test ground. Proto-Ongan (PON) is the name given to the reconstructed language with ancestral forms for the two documented languages of the Andaman islands, south of India, and Proto-Austronesian is abbreviated PAN. Austronesian languages are spoken a.o. in Taiwan, the Philippines, Indonesia, Micronesia and Madagascar. PAO is the abbreviation for the reconstructed Proto-Austronesian-Ongan. There are different proposals for dating Proto-Austronesian, and one of these is Greenhill & Gray (2009), who calculated an age of 5310 years for Proto-Austronesian and 4240 for Proto-Malayo-Polynesian, the only branch that moved out of Taiwan. The time depth for PAO would be close to, or beyond, the limits of the comparative method, said to be 6000 years.

As an exercise, one can try to add Proto-Basque equivalents to reconstructions for PON-PAN, and one should expect a much lower number of lookalikes between PON and Basque than between PON and PAN, if the latter two are related and if Basque and PON are assumed to be related (which is of course highly unlikely). First, because of the unlikely historical connection. Second, the PON-PAN connections were based on a selection of the total lexicon, rather than the 25 selected meaning-form pairs that were already identified for PON-PAN. For possible Basque lookalikes, the choice is therefore much more restricted.

I chose the 25 first PON-PAN reconstructions. The similarities with Proto-Basque are striking. The PAO forms are from Blust's (2014) list extracted from Blevins (2007), and the PB (Proto-Basque) forms are quoted from Blevins (2018). The quotes end with pages numbers in Blevins (2018) between brackets, and some additions are by me when Blevins does include roots with these meanings in her (2018) book.

(1) PAO **aCa* 'high up'.
PB **ga* 'top, high', cf. Sanskrit *gagana* 'sky…' (259)

(2) PAO **aCay* 'liver'.
PB **gi* 'meat, flesh', *bel* 'dark', B *gibel* 'liver', no PIE etymology (266)

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

(3) PAO *-*aku* 'self, ego; third person pronoun'.
Not treated; cf. Basque *hau*, third person < DEM. Cf. Roncalese *kaur* and Salazarese *kur-*.

(4) PAO *-*ala* 'fetch, get, take'.
PB **har* 'take, receive', PIE h_2r 'to take, acquire' (277)

(5) PAO **aNak* 'child, offspring'.
Not treated, but cf. B *haur* 'child', *anai* 'brother'

(6) PAO **aNiC* 'skin'.
PB **dahar* 'attached, fixed', cf. B *larru* 'skin' (245), PIE *dher/egh* 'bushy plants';
PB **(s)khal* 'shell, husk, skin', B *azal* 'skin, crust…' PIE *(s)kelh* 'wither' (352)

(7) PAO *-*an* 'locative suffix'.
Not treated, but cf. B *-n* 'locative suffix'

(8) PAO **apa* 'carry on back', PON *-*apa* 'carry on shoulder'.
PB **lepho* 'flat, flat part, shoulder blade', PIE *lep* (2) 'to be flat, palm, sole, shoulder blade' (309)

(9) PAO **aqajaw* 'sun, sunlight, day'.
PB **gur(h)* 'warm', PIE $g^{w}er$-, 'to warm, heat' (273), cf. Basque *egun* 'day', *eguzki* 'sun'.

(10) PAO **aRi* 'come, go (movement toward speaker)'.
PB **thor* 'come, come through, bore through', PIE ter/h_1 'to bore, drill', ter/h_2 'to come through, cross, traverse' (381)

(11) PAO **ati-a* 'our, us; 1PL possessive pronoun'.
Not treated. Cf B. *gu* 'we, 1PL', or *-da* 1SG

(12) PAO *-*aya* 'mother'.
PB **ama*, Vedic *ambi, amba* 'sweet mother!', Albanian *ama* (219).

(13) PAO **baqeRuh* 'new'.
Not treated but cf. B *berri* 'new'.

(14) PAO **bel* 'smoke', *PON* *bel 'smoke'.
PB * *kheni* 'vapor, steam, smoke', PIE **keni-*, *koni-* 'dust, ashes'. (298)

(15) PAO **beRay*, PON **bele* 'give'.
PB **ma-n* 'give, hand over', PIE **man* 'hand' (315)

(16) PAO **biaC* 'bow; draw a bow', PON **iya.*
Loanword in B; no equivalent.

(17) PAO **biraŋ* 'anger; angry', PON **biraŋ* 'angry'.
Not treated, cf. B *haserre* 'anger, angry', B. *beldur* 'fear, afraid'.

(18) PAO **buaq*, PON **wa* 'fruit'.
PB **ma* 'fruit' (311).

(19) PAO **bubu* 'conical bamboo basket trap for fish'.
No equivalent.

(20) PAO **buhet* 'squirrel', PON **uhe*.
Not treated, B *urtxintxa* 'squirrel', B *untxia* 'rabbit'.

(21) PAO **bukeS* 'hair', PON **ukec/ele* 'tail of animal'.
PB *hulhe* 'wool, hair' (291), PB *bursthan* 'extremity, tip, tail, penis', Albanian *bisht* 'tail' (241).

(22) PAO **bukij* 'forested area, with many trees', PON **ukiy* 'area with many trees'.
PB **baso* 'wild, uncultivated', PIE *bhos-o* 'naked' (227); cf also B *zuhaitz* 'wood'.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

(23) PAO **buluq* 'type of plant with slender stem for use as small poles, arrow shafts, and general construction', PON **ulukw*, **ulu*.
No equivalent.

(24) PAO **Cekelu* 'beckon to come back', PON **e/jegulu*.
No equivalent. But cf. B. *itzuli* 'return'.

(25) PAO **Cenek* 'hard exterior; sharp exterior', PON **cenek* 'shell; hard exterior', PAO **Cana* 'hard exterior, sharp exterior'.
No equivalent, but cf. B. dialect *kaskolin* 'nutshell'.

With some creativity, one can observe similarities between Basque and Proto-Ongan in forms 1-5, 7-13, 18, 20, 21, that is, in 15 out of 21 word pairs (four out of 25 PONPAN forms have no meaning equivalent in PB, or no native root, and these are therefore not included). That is an impressive 71 %. On the basis of these words alone, it is even possible to propose regular sound correspondences (e.g. b/b in (13), (17), (22)). With an increase of the number of lookalikes, more sound correspondences will obviously suggest themselves, once more forms are used, or when the semantic range is stretched to the maximum admissible. In these (Proto)Basque-Proto-Ongan comparisons, the meaning ranges are considerably smaller than in most long-range comparisons. PB looks surprisingly close to PAO, perhaps closer to PAO than to PIE. It shows that the method is not easily constrained, and it is probably easy to find a sufficient set of formal lookalikes with similar meanings, have them promoted to cognates, and formulate a number of sound correspondences covering a limited number of word pairs.

13. DISCUSSION AND CONCLUSION

If ultimately all languages of the world go back to a single point of genesis, all languages would be related, but attempts of reconstructions of «Proto-World» rarely draw attention from mainstream linguists, and are usually only taken seriously by other «lumpers». The comparative method has a time depth limit of 6000 to 8000 years, according to optimists perhaps 10 000 years. In practice, studies on genealogical connections are based on lexicon only. More recent proposals that suggest that specific conservative grammatical similarities may extend the ceiling to perhaps 30 000 years (Nichols, 1992), have not been widely accepted. Language has been around, we can assume, for at least several hundreds of thousands of years, if not a few millions of years, and that simply means that the currently accepted methods for establish deep genealogical relationship on the basis of lexical cognates is limited to a time depth of less than six to ten millennia. That ceiling is much closer in time than the emergence of the earliest human languages even in the most conservative estimates, implying that lexicons of the world's languages have been completely recycled at the very least ten times.

The six testing scenarios we used above, pointed in two different directions. In some cases, Blevins hypothesis was strengthened, in other cases it was weakened by the results.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

The hypothesis that PIE and PB are related is intriguing, well argued and it would fall within the 10 000 year window in its current form, but there are a few weak points.

One is that Pre-Proto-Basque and PIE were spoken at quite different points in time. Thanks to the presence of hundreds of IE languages, which have been documented for almost 4000 years in time, the reconstructed PIE language can be dated back 6000, according to some perhaps even as far back as 10 000 years B.P. The time-depth for Basque-Aquitanian, however, is much shallower. There are only some very limited 2000 year old inscriptions to rely on for the time-depth, and a reconstruction based on differentiation between Basque dialects that are not hugely dissimilar. Thus, one can go back perhaps 2500 years in time in establishing the earliest reconstructable form of the ancestor of Basque. This means that the languages that are being compared (Proto-Indo-European and Pre-Proto-Basque) were spoken in quite different time periods. The difference is at least three and a half millennia, and perhaps as much as 7500 years. Blevins is of course aware of this, and she motivates this with an assumption that Basque (PB) must have been a very conservative language, from the viewpoint of its phonology, and that it had not undergone much change during several millennia, and therefore it can be compared with PIE. This observation, by the way, has also been made by others, who have pointed out the conservative nature of Basque phonology during the past two millennia at least (Trask, 1997). The whole enterprise thus builds on the assumption that IE diversified hugely in 6 to 10 millennia into 10 quite different branches, covering some 450 languages, and that Basque-Aquitanian hardly diversified during six millennia. That scenario is not very likely, especially as there is very little evidence of such conservatism in languages anywhere in the world. That is the most problematic point of the proposal.

Another major problem pointed out, is the enormous semantic range of the proposed cognates, often so broad that the semantic similarities cannot be detected by just comparing meanings of PB and PIE word glosses. This, however, has been dealt with in the meantime (see the afterword).

A third problem is that there may be a possibility to find cognates also in other reconstructed languages with brief reconstructed proto-forms. It appears possible to find accidental lookalikes that can be presented as being cognates between Basque and other proto-languages as well, as I showed in the comparisons of Proto-Basque with Proto-Ongan, with 71 % similarities in a limited set of meaning sets. The comparison between Basque and Proto-Ongan-Austronesian seems to be as good as the proposals between POA and PON, but historically much more unlikely.

Based on admittedly a fraction of the knowledge that Blevins displays, I would have to defer my judgment about the success of the proposal. It is clearly a proposal that has to be taken seriously, and it could well be that there will be a breakthrough in the near future. The proposal has to be taken seriously, and a scholarly and open debate should take place. Personally, I have had the (pre-scientific) impression that Basque had some connection with languages in the Himalayas and South Asia, including Indo-European languages, based on some lexical and grammatical similarities. I am aware that future

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

research may show a closer connection of Basque with other Eurasian languages, both Indo-European and North Caucasian.

Grammatical similarities are often not taken seriously for genealogical signals, also because most classifications are based primarily on lexical-phonological comparisons. If one considers those, Basque may be linked (but not necessarily in a geneaological manner) with North Caucasian languages and languages of the Himalayas. At least, that is what the results of global comparisons on the basis of WALS data have led to. Languages of the Himalaya include Hindi (Indo-European) and non-Indo-European (Burushaski isolate) languages, suggesting a very deep time depth.

Blevins is without doubt one of the best argued cases for a link of Basque with other languages. Her impressive scholarship, the knowledge of the facts necessary to make statements of Basque-PIE connections, the sober argumentation, the strict application of established methods in historical linguistics, the lack of fear to go against mainstream bascologists, and the openness to consider counterarguments and revise parts of her thesis (see the revisions and errata, Blevins, 2020b), all point to the credibility of the proposal.

A historical linguist of the caliber of Juliette Blevins, is not likely to make elementary errors. As far as I can see (but I am not an expert in all of the areas, let that be clear), there are no cases of invented roots and forms. Proposed sound correspondences cover several words rather than single pairs of words. Sound changes (of which there are many) are listed as well.

A work of this enormity cannot be flawless. Based on criticism of specific points, after the book came out, she corrected some errors (e.g. unidentified loanwords), and dropped some forms (see https://julietteblevins.ws.gc.cuny.edu/files/2020/09/PBerrataAll2020.pdf). It is of course a strong point to be able to admit errors, and adjust the details of the hypothesis accordingly. Her list of errata shows how important the details are for her, and they hardly affect her cognacy pairs. Her general claims still stand. The general hypothesis has at least not been falsified on the basis of those objections, and thus it still stands.

But is it convincing? I started this review stating that the book is a brilliant endeavor, but unfortunately I think there are two main objections that lead me to a judgement that the case is not proven.

One is the problem of the time depth. Blevins uses, and rightly so, the oldest reconstructable forms for both PB and PIE. There is, however, a huge time lapse between the two. Proto-Basque goes back perhaps two or three millennia, whereas the estimates of PIE range from six to ten millennia. The assumption is that Basque hardly changed between the period that PB was spoken and the period PIE were spoken, but languages and are known to evolve, and they can change considerably in 3000 to 8000 years. That is a weak point. Blevins is aware of that, and she assumes that Basque lexical roots were stable, much more stable than we are used to, on the basis of our knowledge of language change in the world.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

The second problem is that the semantic leeway taken in the comparisons is very wide, but that has been dealt with in Blevins and Sproat (in press). In quite a few cases, the PIE forms were probably located by applying identified sound laws, and then the meanings were matched, but many of these meaning mappings were not convincing. The raters agreed in no more than a third of the proposed meaning pairs. But Blevins and Sproat have found a way to set limits to such semantic ranges.

Still, there are intriguing observations by Blevins and some striking similarities. How many of these should one have in order to say that two languages are related? In the end, by lack of a fixed criterion about the quantity or quality of the proposed cognates, it is a question of taste when one accepts a proposal. Some linguists deny that Hungarian and Finnish are related and that the Uralic family does not exist (Marcantonio, 2002), even though the existence of this family may be undeniable for most linguists.

Blevins' hypothesis deserves to be taken seriously, and I look forward to a detailed discussion of the book, beyond pointing to a few selected weaknesses, by the most knowledgeable experts, bascologists and Indo-Europeanists. Unless they are able to provide counterarguments, not just picking a few points, but at the same level of detail as Blevins in her book, it may well be that the hypothesis will be proven and accepted in half a century or so from now.

14. AFTERWORD

After having written the above review, another article by Blevins, co-authored with Richard Sproat (Blevins & Sproat, to appear), came to my attention. In this article, Blevins & Sproat perform a carefully designed Oswalt Monte Carlo test on the PB and PIE materials. For one thing, they only accept meaning ranges that are attested in existing languages, basing themselves on a recently published database of meaning combinations. For example, some languages have the same word from «arm/hand» and «shoulder», or «to lie» and «bed», and only such documented combinations that are attested in at least two language families, are acceptable. This limits the meaning ranges considerably, and in this way, they are no longer unconstrained. The forms of the words were also put to the test, and the conclusion was that the likelihood of the PB/PIE cognate pairs being just arbitrary, is very small. They conclude that PIE and PB «are related languages». This newly developed method, or this innovative application of the Oswalt Monte Carlo test, sounds promising, and some of the objections raised in the review are no longer valid.

I also like to mention that Lakarra et al. (2019) was not at my disposal when writing this article.

15. REFERENCES

Antonov, A. (2015). Verbal allocutivity in a crosslinguistic perspective. *Linguistic Typology*, *19*(1), 55-85. DOI: https://doi.org/10.1515/lingty-2015-0002

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Ariztimuño, B., Zuloaga, E. & Krajewska, D. (2019). Against the Proto-Indo-European-Euskarian hypothesis, or why Basque continues to be a language isolate. Paper presented at the 52nd Annual Meeting of the Societas Linguistica Europaea, Leipzig, Germany, August 21-24.

Bakker, P. (2006). Algonquian-Ritwan, (Kutenai) and Salish: proving a distant genetic relationship. *University of British Columbia Working Papers in Linguistics*, *18*, 1-32.

Bakker, P. (2007). Proto-Algonquian *çk and Salish. *Algonquian and Iroquoian Linguistics*, *32*(4), 26-27.

Bakker, P. (2009a). Response to Jan van Eijk: More arguments for an Algonquian-Salish connection. In K. S. Hele & R. Darnell (eds.), *Papers of the 39th Algonquian Conference* (p. 24-51). London: The University of Western Ontario.

Bakker, P. (2009b). C.C. Uhlenbeck and the Ritwan languages of California. In: I. Genee & J. P. Hinrichs (eds.), *C.C. Uhlenbeck (1866-1951): A linguist revisited*, *Canadian Journal of Netherlandic Studies*, *29*(2)/*30*(1), 151-163.

Bakker, P. (2011). C. C. Uhlenbeck's work on the Basque language. *Anuario del Seminario de Filología Vasca Julio de Urquijo*, *42*(2), 237-264.

Bakker, P. (2013). Diachrony and typology in the history of Cree (Algonquian, Algic). In F. Josephson & I. Söhrman (eds.), *Diachronic and typological perspectives on verbs* (p. 223-259). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/slcs.134.09bak

Bakker, P. & Hinrichs, J. P. (2011). An annotated bibliography of C.C. Uhlenbeck's writings on Basque (1888-1981). *Anuario del Seminario de Filología Vasca Julio de Urquijo*, *42*(2), 265-282.

Bengtson, J. D. (2016). A Study in Basque etymology and phonology from a Euskaro-Caucasian perspective. Discussion Draft | February 25, 2016.

Bengtson, J. D. (2017). *Basque and its closest relatives: A new paradigm*. Cambridge, Mass.: Mother Tongue Press.

Bengtson, J. D. (2020). ms. A non-trivial pursuit? Some notes on the etymologies and putative phonological correspondences between Basque and North Caucasian languages. Discussion draft 09/22/2020 JDB.

Blevins, J. (2001). *Nhanda: An aboriginal language of Western Australia* (*Oceanic Linguistics Special Publications*). Honolulu: University of Hawaii Press.

Blevins, J. (2004). *Evolutionary phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, J. (2007). A long lost sister of Proto-Austronesian? Proto-Ongan, mother of Jarawa and Onge of the Andaman Islands. *Oceanic Linguistics*, *46*, 154-198.

Blevins, J. (2018). *Advances in Proto-Basque Reconstruction with Evidence for the Proto-Indo-European-Euskarian Hypothesis*. London & New York: Routledge.

Blevins, J. (2020a). Derivational patterns in Proto-Basque word structure. In L. Körtvélyessy & P. Štekauer (eds.), *Complex words: Advances in Morphology* (p. 222-243). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781108780643.011

Blevins, J. (2020b). Errata. Updated August 18, 2020. https://julietteblevins.ws.gc.cuny.edu/files/2020/09/PBerrataAll2020.pdf

Blevins, J. & Sproat, R. (To appear). Statistical evidence for the Proto-Indo-European-Euskarian Hypothesis: A word-list approach integrating phonotactics. *Diachronica*.

Blust, R. (2014). Some recent proposals concerning the classification of the Austronesian languages. *Oceanic Linguistics*, *53*(2), 300-391. DOI: https://doi.org/10.1353/ol.2014.0025

Bomhard, A. R. (2008). *Reconstructing Proto-Nostratic: Comparative phonology, morphology, and vocabulary*. Leiden: Brill.

Brown, C. H., Beck, D, Kondrak, G., Watters, J. K. & Wichmann, S. (2011). Totozoquean. *International Journal of American Linguistics*, *77*, 323-372. DOI: https://doi.org/10.1086/660972

Brown, C. H., Wichmann, S. & Beck, D. (2014). Chitimacha: a Mesoamerican language in the Lower Mississippi Valley. *International Journal of American Linguistics*, *80*(4), 425-474. DOI: https://doi.org/10.1086/677911

Campbell, L. (1997). *American Indian languages: The historical linguistics of native America*. New York & Oxford: Oxford University Press.

Campbell, L. (2011). Review of The Dene-Yeniseian Connection, ed. by James Kari and Ben A. Potter. *International Journal of American Linguistics*, *77*(3), 445-451. DOI: 10.1086/660977.

Cowan, H. K. J. (1971). Prae-Indo-Europese relicten in de Nederlanden? *Leuvense Bijdragen*, *60*, 159-198.

Dryer, M. S. & Haspelmath, M. (eds.). (2013). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info

Dunn, M. (2015). Language phylogenies. In C. Bowern & E. Bethwyn (eds.), *The Routledge handbook of historical linguistics* (p. 190-211). London: Routledge.

Dunn, M. & Terrill, A. (2012). Assessing the lexical evidence for a Central Solomons Papuan family using the Oswalt Monte Carlo Test. *Diachronica*, *29*(1), 1-27. DOI: 10.1075/dia.29.1.01dun

Forni, G. (2013). Evidence for Basque as a IE Language. *Journal of Indo-European Studies*, *41*(1-2), 39-180.

Goddard, I. (1975). Algonquian, Wiyot and Yurok: Proving a distant genetic relationship. In D. Kinkade, K. L. Hale & O. Werner (eds.), *Linguistics and anthropology: In Honor of C.F. Voegelin* (p. 242-262). Lisse: Peter de Ridder.

Gorrochategui, J. (1995). The Basque language and its neighbors in antiquity. In J. I. Hualde, J. A. Lakarra & R. L. Trask (eds.), *Towards a history of the Basque language*. (p. 31-63). Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.131

Gorrochategui, J., Igartua, I. & Lakarra J. A. (eds.). (2018). *Historia de la lengua vasca*. Vitoria-Gasteiz: Eusko Jaurlaritza.

Gorrochategui, J. & Lakarra, J. A. (2013). Why Basque is not, unfortunately, an IE language? *Journal of Indo-European Studies*, *41*(1-2), 203-237.

Greenberg, J. (2000). *Indo-European and its closest relatives. The Eurasiatic language family, v. 1 Grammar*. Stanford, Ca.: Stanford University.

Greenberg, J. (2002). *Indo-European and its closest relatives. The Eurasiatic language family, v. 2 Lexicon*. Stanford, Ca.: Stanford University.

Greenhill, S. J., Atkinson, Q. D., Meade, A. & Gray, R. D. (2010). The shape and tempo of language evolution. *Proceedings of the Royal Society B, Biological Sciences*, *1693*, 2443-2450. DOI: https://doi.org/10.1098/rspb.2010.0051

Greenhill, S. J. & Gray, R. D (2009). Austronesian language phylogenies: Myths and misconceptions about Bayesian computational methods. In A. Adelaar & A. Pawley (eds.), *Austronesian historical linguistics and culture history: a festschrift for Robert Blust* (p. 375-397). Canberra: Pacific Linguistics.

Haspelmath, M., Dryer, M. S., Gil, D. & Comrie, B. (eds.). (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Jakobson, R. (1962). Why 'mama' and 'papa'? In Jakobson, R. *Selected writings, vol. I: Phonological Studies* (p. 538-545). The Hague: Mouton.

Kari, J. & Potter, B. (2010). *The Dene-Yeniseian connection* (Anthropological Papers of the University of Alaska, new series, 5). Fairbanks: University of Alaska Fairbanks, Department of Anthropology.

Kassian, A. (2013). On Forni's Basque-Indo-European Hypothesis. *Journal of Indo-European Studies*, *41*(1-2), 181-201.

Lakarra, J. A. (2017). Basque and the reconstruction of isolated languages. In L. Campbell (ed.), *Language isolates*. London: Routledge. DOI: https://doi.org/10.4324/9781315750026-3

Lakarra, J., Manterola, J., Segurola, I. 2019. *Euskal Hiztegi Historiko-Etimologikoa (EHHE-200)*. Bilbo: Euskaltzaindia.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*(8), 707–710.

Marcantonio, A. (2002). *The Uralic Language isolates* (p. 50-59). *Facts, Myths and statistics*. Oxford: Blackwell.

Martín, J. (2014). Dogón y euskera: algo más que un origen común. Euskeraren Jatorriaren 9. Biltzarra. Gernika-Lumo. 2014.05.10.

Martinez Areta, M. (ed.). (2013). *Basque and Proto-Basque*. Frankfurt: Peter Lang.

Michelena, L. (1977). *Fonética histórica vasca* (2nd edition). Donostia-San Sebastián & Vitoria-Gasteiz: Diputación Foral de Guipuzcoa & University of the Basque Country.

Müller, A., Wichmann, S., Velupillai, V., Brown, C. H., Brown, P., Sauppe, S., Holman, E. W., Bakker, D., List, J.-M., Egorov, D., Belyaev, O., Mailhammer, R., Urban, M., Geyer, H. & Grant, A. (2010). ASJP World Language Tree of Lexical Similarity: Version 3 (July 2010). Retrieved from https://asjp.clld.org/static/WorldLanguageTree-003.pdf

Nichols, J. (1971). Diminutive consonant symbolism in Western North America. *Language*, *47*(4), 826-848. DOI: https://doi.org/10.2307/412159

Nichols, J. (1990). Linguistic diversity and the first settlement of the New World. *Language*, *66*(3), 475-521.

Nichols, J. (1992). *Linguistic diversity in space and time*. Chicago: University of Chicago.

Nichols, J. (1998). The origin and dispersal of languages: Linguistic evidence. In N. Jablonski & L. C. Aiello (eds.), *The Origin and Diversification of Language*, (p. 127-170) (Memoirs of the California Academy of Sciences, 24). San Francisco: California Academy of Sciences.

ISSN: 0046-435X ISSN-e: 2530-5832 ISSN-L: 0046-435X

Oswalt, R. L. (1970). The detection of remote linguistic relationships. *Computer Studies in the Humanities and Verbal Behavior*, *3*, 117-129.

Poser, W. (2003). On the end of the Ritwan controversy. Retrevied from http://www.billposer.org/Papers/oerc.pdf

Renfrew, C. & Nettle, D. (eds.). (1999). *Nostratic: Examining a linguistic macrofamily*. Cambridge: McDonald Institute for Archaeological Research.

Ribeiro, E. R. & Van der Voort, H. (2010). Nimuendajú was right: The inclusion of the Jabutí language family in the Macro-Jê stock. *International Journal of American Linguistics*, *76*(4), 517-570. DOI: https://doi.org/10.1086/658056

Salmons, J. C. & Joseph, B. D. (eds.). (1998). *Nostratic: Sifting the evidence*. Amsterdam: John Benjamins.

Sapir, E. (1913). Wiyot and Yurok, Algonkin languages of California. *American Anthropologist*, *New Series*, *15*, 617-646. DOI: https://doi.org/10.1525/aa.1913.15.4.02a00040

Schuchardt, H. (1907). Die iberische Deklination. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-Historische Klasse*, *157*, 1-90.

Schuchardt, H. (1913). Baskisch-hamitische Wortvergleichungen. *Revista Internacional de Estudios Vascos / Revue International des Études Basques*, *7*, 289-340.

Schuchardt, H. (1914). Baskisch und Hamitisch. Zu 78 tšilbor Nabel. *Revista Internacional de Estudios Vascos / Revue International des Études Basques*, *8*, 76.

Tovar, A. (1980). *Mitología e ideología sobre la lengua vasca: historia de los estudios sobre ella*. Madrid: Alianza.

Trask, R. L. (1995). Origin and relatives of the Basque language: review of the evidence. In J. I. Hualde, J. A. Lakarra & R. L. Trask (eds.), *Towards a history of the Basque language* (p. 65-99). Amsterdam: John Benjamins.

Trask, R. L. (1997). *The history of Basque*. London: Routledge.

Trask, R. L. (2008). Etymological Dictionary of Basque. Edited for web publication by Max W. Wheeler, University of Sussex.

Trombetti, A. (1925). *Le origini della lingua basca*. Bologna: Azzoguidi.

Uhlenbeck, C. C. (1946). Gestaafde en vermeende affiniteiten van het Baskisch [Proven and presumed affiliations of Basque], *Mededeelingen der Koninklijke Akademie van Wetenschappen, Afdeeling letterkunde, N. R.*, *9*(2), 13-24. [French translation in *Eusko-Jakintza, 1* (1947), 171-182].

Van Eijk, J. (2007). Salish and Algonquian: A Possible Relationship Revisited. In H. C. Wolfart (ed.), *Papers of the 38th Algonquian Conference* (p. 403-416). Winnipeg: University of Manitoba.

Wichmann, S., Holman, E. W. & Brown, C. H. (eds.). (2020). The ASJP Database (version 19). Retrevied from https://asjp.clld.org

Willemsen, J. & Miltersen, E. H. (2020). The expression of vulgarity, force, severity and size. Phonaesthemic alternations in Reta and in other languages. *Studies in Language*, *44*(3), 659-699. DOI: https://doi.org/10.1075/sl.19073.wil

Willemsen, J. & Robbe. J. R. (n. d.). Vestigial phonaesthemic alternations in West-Flemish: consonantal play as evidence for the emergence of sound symbolism through language contact.