

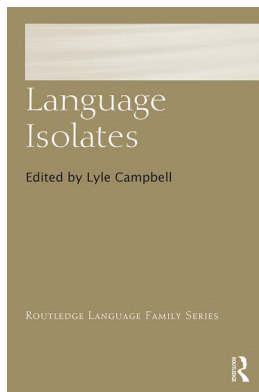
This article was downloaded by: 10.3.98.93

On: 20 Sep 2019

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



Language Isolates

Lyle Campbell

Basque and the Reconstruction of Isolated Languages

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9781315750026.ch3>

Joseba A. Lakarra

Published online on: 25 Sep 2017

How to cite :- Joseba A. Lakarra. 25 Sep 2017, *Basque and the Reconstruction of Isolated Languages* from: Language Isolates Routledge

Accessed on: 20 Sep 2019

<https://www.routledgehandbooks.com/doi/10.4324/9781315750026.ch3>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

BASQUE AND THE RECONSTRUCTION OF ISOLATED LANGUAGES

Joseba A. Lakarra

1 INTRODUCTION¹

I think it's appropriate to ask what the purpose of our genetic classification is. I believe that most historical linguists value the classifications because they help us find out about the histories of the languages in a family. We reconstruct parts of their common protolanguage and then use those reconstructions to study and compare the changes that have occurred in the various daughter languages. In other words, *to be useful to a historical linguist, a hypothesis of genetic relationship must be fruitful: a valid genetic grouping will permit reconstruction and thus lead to be a better understanding of the member languages and their histories. If a genetic hypothesis does not lead to new insights of these kinds, therein it is sterile and, within linguistics, useless.*

(Thomason 1993, p. 494; emphasis added [JAL])

The classification and comparison of languages is *not* the ultimate goal of diachronic linguists. Their main task is to describe and explain the development of languages or families studied in their different phases (whether documented or not). Comparison² is of no scientific interest except when, undertaken in strict conditions, its objective is to illuminate the structure of the languages under study and the changes produced within them, especially irregularities, exceptions, *fases sparitas* (stages with little or no attestation, impossible or difficult to investigate in the language (or dialect) itself).

In language families with a long developed tradition of diachronic research, such as Indo-European (IE), Uralic, or Semitic, demonstration of genetic relationships has not signified the culmination of comparativists' work³ but rather the start of their true vocation as historical linguists. Their work must be based on the regularity of phonetic change and on homologies, not on analogies, similarities, and superficial or casual resemblances, such as the spurious "similarities", that, as Trask (1996: 220) shows, one could find between Ancient Greek and Hawaiian or between Hungarian and Basque (1997: 412–415), which only make for an amateur entertainment.

That the Basque language is genetically isolated has been an obvious statement for a long time now: its structural differences with respect to other languages, whether geographically close or not (Romance, Germanic, Semitic, and so on), are clear for all to see. In the past, when Bascophiles or Basque apologists have sought to transform historical, political, or religious issues into linguistic questions they often claimed that in ancient times Basque (B) was spoken in the whole of the Iberian Peninsula (more so than in Gaul

or elsewhere in Europe), generally without attempting to establish genetic relationships (see Tovar 1981, Madariaga 2008).

After the development of comparative-historical linguistics, only very unorthodox linguistics, like Schuchardt, Uhlenbeck, Trombetti, or Tovar,⁴ have attempted to relate B to other languages; those more aware of the limits of this method have avoided fantasizing impossible historical scenarios and chronologies as well as improper and ridiculous comparative “tricks”.

Since having no proven genetic relationships implied imbuing the language with an ancient aura, non-Basque amateurs⁵ have searched for such relationships with all kinds of other languages, from Iberian to Na-Dene, by way of all kinds of “bright” ideas.⁶ In general, specialists in the relevant languages or families have not felt too concerned about such attempts or have refuted them as implausible, unfounded, ignorant of the data, based on erroneous analysis, and so on. Similar problems are typically found when these comparisons are judged from the perspective of Basque linguistics or even from a consideration of the basic principles of linguistic comparison (see Section 3).

2 SOME BASIC DATA FOR THE HISTORY OF THE LANGUAGE⁷

2.1 Context of Basque

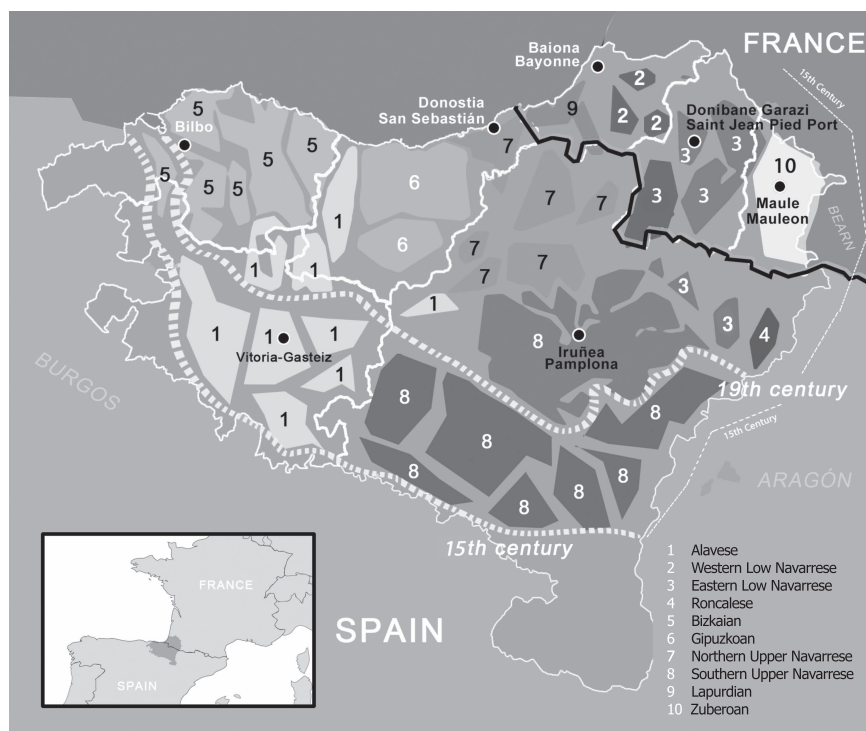
There are just over 800,000 speakers of Basque (*euskara*, variants *heuskara*, *üskara*, *euskera*, *eskuera*, *eskuara*, etc.) on both sides of the Pyrenees, bordering the Bay of Biscay, in the territory – almost 27,000 km² – known in B as *Euskal Herria* (Country of the B Language). This area coincides in general terms with the historical provinces of Lapurdi (Labourd), Behenafarroa (Basse Navarre), and Zuberoa (Soule) in the North and the Kingdom of Navarre (Nafarroa/Navarra) and the provinces of Bizkaia (Vizcaya), Gipuzkoa (Guipúzcoa), and Araba (Álava) in the South.

In the latter three provinces (which together make up the Basque Autonomous Community, BAC) the language has been co-official since 1982 with Spanish, and its introduction into the education and administrative systems⁸ has led to more people knowing it, especially among the younger generations – although not so much to a significant increase in its use – and a clear improvement in expectations for its future. In Navarre it only enjoys co-official status (and in a restricted way) in the B-speaking zone – north of Pamplona (B *Iruña* or *Iruñea*) – following a major decline in B-speakers between the 19th and 20th centuries. In the continental territories to the north, Basque is clearly declining toward a likely near extinction, led by the fairly typical encouragement of linguistic genocide of postrevolutionary France (see Calvet 1981).

Historically, the language was spoken in those territories (although there are no testimonies from parts of the Encartaciones area of Bizkaia or southern Navarre) and their surroundings during the Middle Ages (Rioja, Burgos) and even farther afield (from the Pyrenees to the Mediterranean according to Corominas, though not so extensively according to most; cf. Salaberri 2011a, Manterola 2015) in Late Antiquity (see Map 3.1).

2.2 Basque prehistory

The prehistory of the language – a period lacking any direct or indirect information – ends at the beginning of the Common Era, thanks to the Aquitanian inscriptions (between the 1st and 3rd centuries AD), which contain some 300 to 400 anthroponyms and theonyms.⁹



MAP 3.1 BASQUE MAP

Here we enter into protohistory, a period without any texts written in Basque but with abundant information, especially in the medieval era, thanks to the large number of names and toponyms included in Latin and Romance (Navarrese, Gascon, Castillian) texts; here, Luchaire was a pioneer, as Mitxelena (1964) points out.

2.3 Basque historical periods

The historical period of the Basque language is usually said to begin with the publication of *Linguae Vasconum Primitiae* (Bordeaux 1545) by Bernard Etxepare, even though there are a couple of 11th-century glosses¹⁰ and the odd poetic fragment or letters dating from before 1500. It would seem preferable (cf. Mounole and Lakarra 2017) to move the beginning of this period back to 1400, given that recent philology has demonstrated that the language of texts collected from the oral tradition (sayings, ballads, and elegies) in the late 16th century and first half of the 17th century correspond to much older periods.

The typical periodization, in eight stages (Lakarra 1997a), combines internal linguistic criteria with other references to sources or phenomena in the literary language:

- 1 Up to the 1st century of the Common Era: Proto-Basque (PB)
- 2 1st–3rd centuries: Basque in Antiquity

- 3 10th–14th centuries: Medieval Basque
- 4 1400–1600: Archaic Basque
- 5 1600–1745: Old and Classical Basque
- 6 1745–1876: First Modern Basque
- 7 1876–1968: Second Modern Basque
- 8 1968–: Contemporary Basque

Thus, for example, some phonetic changes in composition and derivation and the presence of the article in the Middle Ages separate (2) and (3); the nature of the (secondary/primary) corpus distinguish (3) and (4); archaic verb forms (aorist and other extinct tenses and moods, more numerous synthetic forms) serve to separate (4) and (5); the consequences of Larramendi's work – 1729, *Grammar*, 1745 *Sp-B-Latin Dictionary* – mark the boundary between (5) and (6); and the unification or standardization of the language is the landmark between (7) and (8).

Dialectal differences,¹¹ which to uneducated enthusiasts and speakers may appear great, are few for comparativists (cf. Mitxelena 1964, 1981), so that the origins of the initial dialectal divergence must be dated close to its early documentation. Mitxelena dated Old Common Basque (OCB), toward the 5th–6th centuries (see 9th section), a thousand years after Late Proto-Basque (LPB), which is a stage of language defined as “the language that the Romans encountered”.

3 CRITIQUE OF COMPARISONS

The Basque language has been the object of many attempts to link it genetically to languages nearby and distant in both space and time. However, 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 as it has been developed in truly established language families (see, *inter alia*, Campbell 2013, Trask 1996, Watkins 1990 and Meillet 1925 and Mitxelena 1963); see Campbell 2011 and *supra* chapter 1 about Basque-Aquitanian relation.

No standard evidence of genetic relationship has ever been provided (nor attempted) discovering phonetic rules (sound correspondences) relating – for example – Iberian and Basque or elaborating the historical grammar, and the few “promising” cognates (homophones) have dwindled to such an extent that they have all but disappeared from the literature (cf. de Hoz 2010–2011), either as a result of changes in reading and/or interpretation within Iberian or because of advances in Basque linguistics and philology that make them impossible; See also chapter 2, section 3.4.

The different hypotheses regarding the genetic relationships of Basque – classic ones such as Basque-Iberian, Basque-Caucasian, more recent ones as Basque-Uraloaltaic, Vasconic, Basque-Indo-European, etc. (see Mitxelena 1964, Trask 1997 and Lakarra 2017: Section 2) – share multiple characteristics that discredit them immediately:

- 1) They start from the simple yet false idea according to which, given that neither Basque nor other isolate language belongs to IE, Semitic, Uralic, or other well-established families – with known histories and acceptably reconstructed protolanguages, i.e., impossible to be manipulated on the whim of the amateur of the moment – all of them, and particularly B and one or another of the remaining languages, must belong to the

same family. Since the demonstration of this a priori assumption is a something good, a goal in itself, the objectives, methods, and criteria of comparative-historical linguistics are not sufficient to cause such “discoverers” to desist.

- 2) In cases in which a minimum attempt has been made to prove an argument, analyzing alleged cognates in B and one or more other language, the Basque part of the argument (as well as often the other part, of course) is strewn with errors: erroneous, dialectal, or later meanings and forms; nonexistent loanwords, words, or variants; flawed and arbitrary morphemic analyses; and so on (cf. Campbell 1988 and 2013).¹²
- 3) No attention is paid to the body of work on Basque historical linguistics or the existing literature on the language(s) that are compared with it.
- 4) As a consequence, after many decades of such comparative efforts, no light has been shed on any aspect of the historical phonology or grammar of B (nor of the other languages).
- 5) Frequently, the false illusions or statements derived from those essays are perpetuated in later works: thus, for example, a claim was still recently made about the supposed abundance of initial vowel in B (cf. Odriozola 2016) that, explained arbitrarily as old articles, Schuchardt used to underpin his B-Hamitic-Semitic edifice, which had already been demolished 90 years before.¹³
- 6) In the search for “explanations” for languages compared to Basque, numerous things are overlooked such as significant variants, underused archaic testimonies, irregularities, generalizations and clarifications about phonetics or grammar, and, in general, the most important data for reconstructing B: if the relationship of *hiri*₁ ‘city’ and *hiri*₂ ‘close, near’ (cf. **her*; *her(t)si* ‘to close’ and Sp *cerca*₁ ‘closure’ and *cerca*₂ ‘near’; see Corominas and Pascual 1980–1991, s.u.) is not examined, this can only be due to the age-old and false belief that *hiri*₁ derives from the family of Iberian ILTIR.¹⁴

Altogether, not only is there bad comparative practice in much of this work, but the true nature of comparison is either misinterpreted or ignored. Comparison cannot be a goal in itself and even less so when the universal standards its practice requires are not fulfilled (see Section 1).

4 THE CLASSIC RECONSTRUCTIVE PARADIGM OF PROTO-BASQUE

4.1 Martinet and the plosives

Following diverse prestructuralist essays (by Campi3n, Azkue, Schuchardt, Gavel, Uhlenbeck, and others), Martinet (1950) approached the study of B in order to test diachronic phonology with a specific case – the well-known voicing of initial plosives in B – just as he had done with other language – Semitic, IE, Slavic, Celtic, and Romance languages. In calling for the need for a structural view of the problem – not the change of isolated sounds but of the system as a whole – he transformed the bases of research in this field. Likewise, he also realized that the situation of prolonged bilingualism and uninterrupted contact between dialects was an obstacle for distinguishing variants derived from regular sound change from other types of variants.¹⁵

Latin-Romance loanwords offered the surest support, given that we know both their origin and their Romance evolution, in contrast to what happens with the inherited lexicon. The use of loanwords, allows us to determine the chronology of many internal

changes in the language that have taking place during the last thousand years based on what is known about the evolution of the Romance languages (Mitzelena 1974, Echenique 1984).¹⁶

The changes in Basque plosives had been addressed previously: according to Uhlenbeck, word-initial voiceless consonants were voiced through dissimilation from other intervocalic consonants. The problem with this hypothesis was that, in addition to the uncommon character of this phenomenon, there is also voicing in words without an intervocalic voiceless consonant (*gerezi* ‘cherry’, *gela* ‘room’ < Lat *ceresia(m)*, *cella(m)*, etc.). Gavel (1920: 314ff) suggested that all word-initial plosives voiced regularly in a period subsequent to the adoption of the oldest loanwords, because in B there are no word-initial voiceless consonants, except in recent loanwords, phonosymbolism, after the loss of vowels, regressive assimilations, and so on.

Martinet¹⁷ accepted the important part of Gavel’s explanation: Ancient B had only voiced plosive phonemes in word-initial position, whereas word-finally only voiceless plosives were allowed. On the other hand, both voiced and voiceless plosives were found in intervocalic environment. Nevertheless, the treatment of loanwords demonstrates that something more than voicing is necessary to characterize the old system. For Martinet, the starting point would not be the voiceless/voiced opposition, but another very different one, a *fortis/lenis* contrast as in Danish, which he had previously studied. There would, thus, be two series of plosives, strong /P, T, K/ and weak /p, t, k/. The strong plosives would be realized as aspirated [p^h, t^h, k^h] word-initially and as plain [p, t, k] intervocally. The weak plosives would be produced as soft voiceless [p_o, t_o, k_o] in word-initial position, and as fricative [β, δ, γ] between vowels.

Thus, the initial plosives in Latin loanwords, both voiced and voiceless, were adapted to PB as *lenis* phonemes, given that the aspirated allophones of the *fortis* ones were very different from the Latin sounds. In intervocalic position, each Latin plosive would have its corresponding PB phoneme: voiceless ones strong and voiced ones weak. Word-initially the strong phonemes would not be used in the adaptation of any Latin loanwords, and later they would disappear through the influence of surrounding languages.¹⁸

4.2 The reconstruction of the phonological system by Mitxelena

Mitzelena (1951) accepted Martinet’s hypothesis, adding important observations in its support regarding (1) the development of geminates and clusters of voiced and voiceless consonants, (2) the sound [f], and (3) Aquitanian and medieval graphic testimonies.

The development of geminates would be the clearest sign that Basque speakers paid attention to the strength of sounds (geminates = long, i.e. strong) rather than sonority, because voiced geminates were adapted as (voiceless) *fortis* plosives just like voiceless geminates: Lat *-bb-* > B *p*, cf. Lat *abbas*, *ad valle(m)*, *sabbatu(m)* > B *apaiz* ‘priest’, *apal* ‘humble’, *zapatu* ‘Saturday’ versus Sp *abad*, *sábado*. Romance voiced groups have also given simple voiceless ones: *cobdiçia* > B *gutizia* ‘whim’ versus Sp *codicia*.

The testimony of the sound *f* – nonexistent in the phoneme inventory of PB – would corroborate such a phenomenon, because it was at certain point a variant of *b* (*nafar/nabar* ‘multi-colored’, *afari/abari* ‘dinner’), then *p* (*alfer/alper* ‘idle’, *ifini/ipini* ‘to place, put’). In ancient times, sonority was not the main distinction between *b* and *p*, but rather its strong or weak pronunciation, so that because the fricative *f* was weak, it was replaced by *b*. If in PB the voiced/voiceless distinction had been pertinent, [f] would have been

replaced by one or the other, not by both, for that reason we should assume that this opposition was not the basic one.¹⁹

Lastly, Mitxelena points out that the data on Aquitanian inscriptions coincide with this analysis: the written symbols and <CC> represent strong plosives (vs. <TH>, aspirated). Likewise, the distribution of the plosives shows that *p* was infrequent and a variant of *b* in strong position (i.e. on initial in the second member of the compound (Aquit *Seniponnis* vs. usual *-bon(n)*) and after the sibilant: cf. Aquit *Andoxponni*).

Mitxelena (1957) extended Martinet’s argument to the whole system, believing that the strong/weak opposition affected all the consonants, except /h/. The complete reconstruction would be as follows (where, as in modern Basque orthography, *z* represents a voiceless pre-dorso-alveolar or dental fricative [s̺] and *s* is a voiceless apico-alveolar fricative [s̺̟]):

Strong	—	T	K	TZ	TS	N	L	R	
Weak	b	d	g	z	s	n	l	r	h

The fact that more phonemes do not appear does not mean that there were no more sounds but rather that it is impossible or unnecessary to reconstruct them with the available data for the nuclear phonological inventory. Mitxelena dispensed with (subsequently phonologized) sounds such as the strong bilabial, palatals, historically expressive or automatically generated (secondary) sounds, the /m/, nonexistent outside loanwords and phonosymbolisms, except as an allophone in /b/ in nasal contexts. Vowels are not mentioned either in the original argumentation (1957) or in the summary of the *Fonética histórica vasca (FHV)*, since they had hardly changed since Aquitanian: cf. *Nescato, Cison, Sembe, Ummesahar*. . .). Thus the five vowels of modern B are reconstructed as such, with three levels of height and with no distinction based on quantity; the additional vowel [y] of Zuberoan comes from **u*. The approximants [j] and [w] are later, emerging in different places: [j] from *e-* in old verbs (*joan* ‘to go’ < **e-oan*; cf. *ebili* ‘to walk’ < **e-bil-i*) and [w] in loanwords when it is not a case of a final diphthong.

The plosive and sibilant subsystems seem to coincide in neutralization points (in initial and final position, after a sonorant and before a consonant), reflecting the same opposition: we would have two series of sibilants opposed to one another by point and means of articulation: apical and laminal, fricative and affricate.

The weak/strong distinction persists historically in rhotics (trill and tap), but in contrast to what happens with all other consonants, these do not appear word-initially. The existence of /L/ and /N/ is guaranteed by their behavior in the intervocalic context and by Aquitanian and medieval written symbols. Intervocalically not all the *n*’s and *l*’s behave the same way as one can see clearly in loanwords, and traces remain in inherited words; thus, if from Lat *angelu, gula* and *caelu*, we get *aingeru* ‘angel’, *gura* ‘desire’, and *zeru* ‘sky’ (*Z zeliü*); we could have a weak *-l-* in the causative (*-ra-*) but not in *ilhe/ule* ‘hair’ (> **irhe/*ure*) from **enon-le* > **e.ole* > **eule*, etc. With *ahate* ‘duck’ and *lehoi* (< **leohe*) ‘lion’, the result of Lat *anate(m)* and *leone(m)*, a strong /N/ would be needed by both *arrano* ‘eagle’ and *baino* ‘than . . . (comparative)’, but not **ardano* ‘wine’ (> *ardao, arno, ardan-*) or **bini* ‘grain’ (> *bihi*), or **seni* ‘child’ (> *sehi, sein*) or *bai(n) a* ‘but’.²⁰ There is prothesis before initial *r* in late Lat *ropam* > *arropa*, late Lat *ratonem* > *arratoi*, etc. (as in Aragonese and Gascon via Basque substract); the *muta cum liquida* consonant clusters are broken up, with a vowel before /r/ – *garau* ‘grain’, *daraturu* ‘drill’,

boronte ‘front’ < Lat *granu(m)*, *taratru(m)*, *fronte(m)*, etc., and with the plosive deleted before the lateral: cf. *laket* ‘pleasant’, *loria*, *lau* ‘flat’ < Lat *placet*, *gloria(m)*, *planu(m)*, etc.²¹

Since PB, there has been a tendency to neutralize the fortes/lenes opposition in sibilants and sonorants except intervocalically, in favor of lenis in initial position and fortis in final position:²² cf. *gorputz*, *bake* < Lat *corpus*, *pacem*, and *zeru* < Rom *tselū*; in inherited words cf. *gazi* ‘salty, savory’/ *gatz* ‘salt’.

Mitxelena recognized a phonemic character to aspiration in PB, /h/ being the only consonant outside the strong/weak opposition. Gavel (1920) had a different opinion, arguing in favor of an adventitious late character of /h/, which would not have existed in southern dialects. Analyzing Aquitanian and other medieval peninsular testimonies – and the then recently discovered inscription in Lerga (Navarre) – Mitxelena observed that the old /h/ (a clear difference with respect to its neighbor languages) appeared throughout the historical territory and that, although already lost in Navarre by the 10th century through Romance influence,²³ it was attested until the 13th–14th centuries in Alava and Rioja, and even in Bizkaia and Gipuzkoa.

Therefore, *zahar* ‘old’, *ahuntz* ‘goat’, *zuhur* ‘wise, prudent’, etc. are more archaic than *za(a)r*, *a.untz* (*awntz*, *ajntz*), *zur*, etc., since they retain the structure and number of syllables of a previous stage, in the same way as *ahate* ‘duck’, *ohore* ‘honor’, etc. are older than the corresponding contracted forms (*a(a)te*, *ôre*, etc.).²⁴ Mitxelena established four etymological origins for the historical /h/: (1) PB plosives in absolute initial position, (2) Latin-Romance *f*-, (3) intervocalic lenis **n*, and (4) PB **h*.²⁵ Many etymological *h*’s disappear in historical periods of the language (cf. *FHV* 525 and 219–220), for example, those situated after the accent, or the first of two *h*’s within a root. Historically, there is no *h* beyond the second syllable, but there is in Aquitanian and even a thousand years later in Medieval Basque.

The accent is a difficult point in the reconstruction of LPB. The most common pattern in modern B is phrasal, rather than word-level and only weakly contrastive. This system cannot be very old, given that the evolution of consonants depends on whether they appear in initial or in medial position. There have been two main proposals regarding the old accentual system:

- 1 Martinet: demarcative accent on the initial syllable, which would explain the distribution of plosives.
- 2 Mitxelena: accent on the 2nd syllable, to explain the modern distribution of *h* (never after the 2nd, nor two *h*’s in the same word).²⁶

Later (1995 and ff.) Hualde has placed the old accent on the last syllable of the phrase (lexically contrastive accent arising later in borrowings and morphologically complex words), but his argument seems to correspond more to OCB than to LPB.²⁷

Mitxelena (1979, *FHV*²) proposed for PB the syllabic structure (C)V(W)(R)(S)(T), similar to that observed in Iberian. He established two restrictions: (1) in word-initial position only one of the following consonants could appear: *b*-, *g*-, *s*-, *z*-, *n*-, *l*-; (2) most likely, not all the segmental slots in a syllable were ever filled (*geurtz* ‘next year’ would be the closest).²⁸

As regards morphemes, Mitxelena assumed that the Canonical Root Structure (CRS) of Iberian onomastic compounds and derivatives was “[2 + 2]” and “[2 + 1]” and that the maximum size of roots was two syllables and proposed a similar structure for Old Basque.

4.3 Morphosyntactic reconstruction

As in other languages and language families, morphosyntactic reconstruction of Basque is less advanced. The relationship and organization of elements, which is evident in phonological reconstruction, is difficult to observe in morphosyntax. Furthermore, if in the reconstruction of the phonology, elements belonging to different periods are sometimes mixed together, as we have seen, this applies even more to morphemes, such as derivative suffixes²⁹; plural verbal markers (cf. *gaitean* ‘let us be, we should be’, *gara* ‘we are’ versus *gatoz* ‘we are coming’, *gabiltza* ‘we are walking’); elements in first- and second-level declension; and so on.

In the study of diachronic Basque morphology, one must mention the names of Schuchardt and Lafon (see 1943 and 1999), although in the case of the former, his war against the sound laws of the Neogrammarians was damaging for Basque studies, as Mitxelena pointed out time and again. Several works by Mitxelena are indispensable, even if they are not specifically morphological, among them *Fonética Histórica Vasca* and *Apellidos Vascos*; others that address toponymy (1971), old texts (1954a, etc.), and scattered studies on the form of suffixes, postpositions, or the history of the verb (1977).

Yet, in our opinion, the work of Trask (1977) on the structure of the verb and historical syntax is the most important contribution in the last 60 years. It established that *da-* (cf. *dator* ‘is coming’, *dakar* ‘is bringing’) was not originally a marker of person or time, but was aspectual, and that its position to the left of the root does not favor SOV order of agglutinative languages, but another order altogether, VO (SVO according to him, VSO for Gómez 1994 and Gómez and Sainz 1995). In the same work, Trask clarifies the later character of the dative markers and the origins of ergativity, making it a starting point for later works on Basque diachronic morphosyntax.

5 ANALYSIS OF THE CANONICAL ROOT STRUCTURE

5.1 Antecedents in other languages

An important topic in IE diachrony is that of the CRS, particularly since Benveniste (1935). He reproached his predecessors for having reconstructed for the protolanguage, not a system, but a disjointed amalgam of polymorphous roots. Exploring the economy of the root, he established a general outline from which all those would derive. The results of this line of inquiry have been splendid (cf. Watkins 1984) in morphology, phonology, and the lexicon of the protolanguage and derived languages.

More recently, the CRS has been studied diachronically in other families such as Semitic (del Olmo Lete 2003), Uralic (Bakrò-Nagy 1992), Japanese (Janhunen 1997), Sinitic (Sagart 1999), and others. All these studies have demonstrated significant results for the phonology, morphology, and structure of the core lexicon – especially the reconstruction of old families of words (as well as those cited, see Schuh 2013 on Hausa) – and in some cases (see Elmendorf 1997 on Yuki – Wappo) may contribute evidence about more profound genetic relationships beyond the protolanguages.

5.2 First results of the new paradigm

Until recently little attention was paid to the CRS in Basque. Azkue (1923–1925) dedicated only 20 lines to it in a book of a thousand pages and did not observe that the

root was monosyllabic in ancient times. There is barely anything else until Uhlenbeck (1947 [1942]) and Lafon (1950). One can scarcely say anything positive about the latter, since he ignores internal reconstruction and proclaims that comparison with the Caucasian languages – supposed relatives – is the only existing recourse for analyzing B roots. Mitxelena, in *FHV*, does not use this term, nor does he derive any important or clear consequences derived from it, though he notices, for example, the restriction against homorganic consonants in rhotics and sibilants within the root: *erur* ‘snow’ > *elur/edur*; *berar* ‘grass’ > *belar/bedar*, and *sasoi* < Sp *sazón*, *frantses* < *francés* as *sinetsi* ‘to believe’ < *zinhetsi* (1545), but, across morpheme-boundaries, *erro-aren* ‘of the root’. Uhlenbeck deserves praise for calling for an analysis of root models, although he did so in order to buttress his theory of the polygenesis of the B language, where Biz and the other dialects would stem from different languages. Perhaps for this reason, he did not have any followers in the study of the structure of the root.

In Lakarra (1995), we point out various restrictions on historical B roots – ****VC** and ****CV** in autonomous monosyllabic lexemes, ****T₁TV** in disyllables were not permitted – and we suggest that they could be explained as originating from the root **CVC**; thus important lexical and morphological results were achieved:

A1) additional combinations with known roots:

- **bel* (before *harbel* ‘blackboard’, *ubel* ‘bruised’, *orbel* ‘fallen leaf’, etc.): *sabel* ‘belly’, *gibel* ‘liver, behind’, etc.

A2) new roots (historically fossils):

- **ger*: *okher* ‘mischievous, twisted’, *akher* ‘male goat’, *puzker* ‘fart’.
- **han*: *ahuntz* ‘goat’, (*h*)*andots* ‘ram’, *ahari* ‘ram’.

A3) previously unknown loanwords: *zemai* ‘threat’ (Old Sp *menaza*), *alu* ‘vulva’ (Lat. *aluus*).

B1) reduplication:

- *adar* ‘horn, branch’ < **da-dar*; *eder* ‘beautiful’ < **de-der*; *odol* ‘blood’ < **do-dol*.
- *ahal* ‘to be able’ < **na-nal* (Mitxelena **anal*), *ohol* ‘board’ < **no-nol* (Mitxelena **onol*).
- *zezen* ‘bull’ (< **ze-zen*) cf. *gi-zen* ‘fatty part of meat’ (see below), etc., similar to *go-gor* ‘hard’, etc.³⁰

B2) prefixes

- **gi*: *gibel*, *gizen* ‘fatty part of meat’, *gihar* ‘lean’, *gizon* ‘man’, etc.
- **la*: *labar* ‘cliff, precipice’, *labur* ‘short’, *lagun* ‘friend, companion’, *labain* ‘slippery’.
- **sa*: *sabel* ‘belly’, *samin* ‘bitter, pain’, *samur* ‘tender’, *sabai* ‘attic, loft’.

5.3 Canonical root structure and formal etymology

In a language isolate with relatively late and limited documentation, tracing the etymology of inherited terms is much more complicated than tracing that of loanwords. For that reason, the study of phonotactic restrictions of roots, the structure of families of words, and the “etymologies of the root models”, as Uhlenbeck (1942) suggested, may

contribute to a safer and deeper reconstruction than the atomism that underlies the slogan “every word has its own history”.

For the study of root models, we classify (Lakarra 2008d) the words documented historically – not reconstructions, even those clearly and universally accepted like **e-thor* ‘to come’, **e-dan* ‘to drink’, **e-khar* ‘to bring’, etc. – in five groups: (1) loanwords, (2) later variants, (3) compounds and derived forms, (4) forms due to onomatopoeia and phonosymbolism, and (5) of unknown etymology. Later, productivity, phonotactic, and geographical filters are applied to those included in (5), the only ones potentially belonging to the oldest stages of the language.

In the last 15 years, disyllabic forms have been reclassified from (5) to (1) or (3) – rarely to (2) or (4) – and progress in research is moving in this direction. Adding the non-controversial reconstructions to the list of roots, we would obtain the result that monosyllables of unknown etymology would increase by almost 100% – with hardly any loanwords or derived words – in contrast to disyllabic cases. Thus, the clear difference between CVC and the other models is increased even more, ruling out any disyllabic forms CRS for Old Proto-Basque (OPB). In reality, given that the geographical filter is established based on distribution in modern dialects, the results correspond *at the earliest* to OCB or to later stages, since we have been very lenient when it comes to filtering innovations (cf. Lakarra 2008d and here Section 9).³¹

A formal etymology does not provide the exact origins of specific words, but its value as an initial diagnosis seems clear: if, for example, *fede* belongs to the CVCV type, which is a root type with multiple loanwords and very few inherited words (none with *f*), it is difficult for this word to be included in (3) or (5); if *otso* ‘wolf’ is VCV and *-so* ‘*big, older’ is repeated in *amaso* ‘grandmother’, *alabaso* ‘granddaughter’, and *atso* ‘old woman’, it is highly likely that it derives from **hortz-so* [‘fang’ + ___], with loss of *r* in the consonant cluster: cf. **hertz-bu(n)* > *esku* ‘hand’ or **intzaurzedi* > *intsausti* ‘walnut grove’, etc.

5.4 CVC canonical root structure and the phonology of old Proto-Basque

Phonological analysis of the monosyllabic root leads to the proposal of a new consonant system for OPB (cf. Martínez Areta 2006, Lakarra 2011b, 2017). Mitxelena (*FHV*) suggested that strong consonants in PB could come from old groups, but only the research related to CRS has provided sufficient proof (= etymologies) based on sonorants and sibilants to think that this is the correct direction in which to go.

On the one hand, in a CVC root, there is no internal position for consonants, precisely that in which the Mitxelenian system maintained the fortes/lenes opposition; namely, in OPB the sibilants would have had four allophones and two phonemes (one dorsal and the other apical), not four phonemes as in the later system, since there is no contrast but rather complementary distribution between fricatives and affricates. In reality (cf. Section 4.2), alternations like *gatz* ‘salt’/*gazi* ‘salty’ show that previously there were fricatives also in final position, and Latin loanwords like *gorputz* (< *corpus*) are witness to the fact that, at the start of the Common Era (i.e., after LPB), affrication applied in final position. Insofar as we know, word-medial affricates come from consonant clusters – see *otso* earlier – or from affrications in final position of the first element: *atzo* ‘yesterday’ < *hatz* ‘trace, behind’ (cf. *haz-i* ‘to grow, seed’) + *-o* ‘COMPLETIVE’.

The argument is similar for liquids and nasals, with the exception that there are no rhotics in word-initial position: cf. *baiNo* ‘but’/*baina* ‘except, save’ < **ba(da)din* ‘if it

were' + *-no* 'until'/+ *-a* to *beLe* 'crow, raven' < **bel-le* (cf. *bel-tz* 'black'); *erro* 'root, teat' [< *to 'hang up'] < **e-ra-don* [with *nr* > *R*] (**eradon* > **edaron* > **enaron* > **eanron* > *erro(n)*; cf. *errun* [to 'lay eggs'], *arrau(n/l)tza* 'egg'); see Lakarra 2017, in progress-b and Begiristain in progress).

The consonant system of OPB would then be as follows: *th, kh, b, d,*³² *g, l, n, r, ʒ, ʃ, h*, i.e., five plosives, two sibilants, three sonorants, and *h*.³³ Insofar as vocalism is concerned, we find no reason to modify Mitxelena's reconstruction (*a, e, i, o, u*); with respect to diphthongs, it is possible that in OPB there were none since in LPB much fewer would be reconstructed than those documented historically, with it being plausible to see previous hiatuses emerging as a result of the deletion of old consonants for almost all of them (cf. Lakarra 2010). There could be many more diphthongs arising from the loss of intervocalic consonant than those Mitxelena reconstructed, such as in *-do.i/lohi*. See the end of Section 6.3 on the development of consonants from LPB onward.

5.5 The Proto-Basque lexicon

Although there is interesting etymological information dating from the Middle Ages (*Izpea* 1051 'subtus penna', etc.) together with lines of argument like *Andalucía* < *Landaluzea* 'country, field (*landa*) long (*luze*)' or *alabanza* 'similar (*anza*) to the daughter (*alaba*)' (Larramendi 1745), etc., scholarly B etymology begins with Mitxelena (1950) and the establishing of B sound laws in the late 1950s (see also Mitxelena 1973). Previously, B words were mentioned or utilized improperly in comparisons between Basque and multiple other languages, but the study of their formation and evolution was never even a secondary goal, so that today their value is almost exclusively historiographical (see Agud and Tovar 1988–1995 and Section 5).

The collection of Mitxelenian etymologies by Arbelaiz (1978) and vol. 15 of Mitxelena's *Collected Works* (2012) are still the best available etymological repertoires. M is the author of some 1,400–1,500 etymologies, including both loanwords and inherited words. These etymologies established the foundation of his later diachronic work, as well as of any alternative proposals to Mitxelena's reconstruction. Unfortunately, Trask's (2008) dictionary remains far from complete after his untimely death, and even if he had completed it, there was little new to be expected beyond Mitxelena, of whom Trask always remained indebted, a popularizer of Mitxelena's work.³⁴

The *EHHE* (= Lakarra, Manterola, and Segurola 2017) includes over 2,500 entries, organized into 200 word families. The *EHHE* is a modular work designed to expand according to the percentage of words historically corresponding to each letter. Entries are selected and examined hierarchically according to criteria like age of documentation, historical and modern use, extent of dialectal distribution, abundance of compounds and derivations, particular historical-philological reasons (such as *hapax legomena* or literary language), and so on. The historical part of the microstructure includes, besides the information in the *Diccionario general vasco (DGV)* and other lesser sources, copious information on the protohistorical era – essentially medieval – absent or much scarcer in the *DGV* and other Basque dictionaries. As regards etymology, it goes much farther than the Mitxelenian paradigm thanks to the use of advances in phonological and morphological analysis and in philological documentation but, above all, thanks to the use of formal (see Section 5.3) and comparative etymology (later in this section).

Although problematic as in any protolanguage, we can attempt to establish when and where PB was spoken. It is impossible to determine what the language spoken by the Paleolithic or Neolithic tribes (nor even by one of them) in the area was, but nor is there

any evidence that the remote ancestor of B was ever spoken outside the historical territory of the B language described in Section 2.³⁵ Meanwhile, if *haitz* ‘rock, crag’ appears to be in the word family *aizto* ‘knife’, *aitzur* ‘hoe’, etc., this does not sustain the old-fashioned idea that Basque is a Neolithic language: as Gorrochategui (1998, 2002) explained, there would be no more reason for this than for any other languages, like German, in which the same sort of thing is found. What is more, it is very likely that ‘rock, crag, etc.’ were not the old meanings of *haitz*, which appears to be a compound of **han* ‘big’, cf. *han-di* ‘big’ [*< *to become big*] and animal names like *ahari* ‘ram’, *akher* ‘male goat’, *ahuntz* ‘goat’ (see details earlier), etc.

Many plant and tree names have been taken from Latin or Romance (*porru* ‘leek’, *kip-ula* ‘onion’, *piper* ‘pepper’, *baba* ‘bean’, *leka* green bean, pod’, *olho* ‘oat’, *gerezi* ‘cherry’, *mertxika* ‘peach’, etc.) but not other plants or trees such as *haritz* ‘oak’, *arte* ‘holm oak’, *gari* ‘wheat’, *garagar* ‘barley’, *ardantze* ‘vineyard’, and animals such as *behi* ‘cow’, *ahuntz* ‘goat’, *asto* ‘donkey’, *zezen* ‘bull’, *behor* ‘mare’, *zaldi* ‘horse’, *idi* ‘ox’, *ardi* ‘sheep’, (dialectal) *flea*, *ahardi* ‘sow’, or *txerri* ‘pig’, and not just wild ones like *hartz* ‘bear’, *otso* ‘wolf’, *orein* ‘deer’, *orkatz* ‘roe deer’, etc. It is worth emphasizing that in the names of colors – the peak of Dixon’s (1977) adjective hierarchy – we find participles (*gorr-i* ‘red’, *zur-i* ‘white’, *hor-i* ‘yellow’ [cf. *e-thorr-i* ‘come’, *har-i-tu* ‘take’ (1545) with a pleonastic participial suffix], derivatives (*bel-tz* ‘black’), compounds (*ur-din* ‘grayish’ [*< *to become water*]) and loanwords (*berde* ‘green’, *marroi* ‘brown’, *azul* ‘blue’, *gris* ‘gray’), i.e. ways of substituting for adjectives in languages with closed-class types; see Dixon (1977).

The same critique of IE languages by Benveniste (1935) can be applied to Mitxelena’s etymologies: i.e., we find monosyllables, disyllables, and polysyllables; protoforms with initial and final vowels and consonants, with and without consonant clusters; etc. It is difficult to find a system in them or to believe that many of his etymologies are contemporary.

New analyses have explored the phonology and morphology (phonetic changes, restrictions, relative chronologies, etc.) of the words reconstructed by Mitxelena more deeply. Thus, his reconstructed **ardano* ‘wine’ and **enazur* ‘bone’, trisyllabic forms – recall that he thought that the OB roots were disyllabic – may lead even further: to **e-dara-dan-o* and **berna-zur*, respectively. In the former, we get **dan* from the verb *e-dan* ‘to drink’ and the same prefixal amalgam (*ar-*) in *arbin*, *arran-*, *arrats*, etc., old verbs with applicative/directional + causative. In the latter, insufficiently considered phonetic changes such as *R – R > Ø – R* (cf. *ezker* ‘left (hand)’/ *eskuin* ‘right’ in **herz-bu(n)-ger*/**herz-bu(n)-on(e)* (**erskuin* after assimilation and simplification) or **b- > Ø – not* just in front of *o/u* as Mitxelena contended – would lead us to identify a first disyllabic element taken in a loanword (*berna* ‘leg’ < Lat *perna* ‘ham’; cf. Eng *bone* = Ger *Bein* ‘leg’), as well as the inherited root *zur* ‘lumber’, both autonomously known.

Comparative etymology – parallels in the formation and development of B words –³⁶ may lead us to discover the longed for “motivation” of Benveniste (cf. de Lamberterie 2000). Thus, for example, *bi* could not always have been ‘two’ but, rather, something like ‘above, over’ or, even better, ‘franchissement’ as Benveniste (1954) saw it in the family of Lat *pons-pontis*, etc.;³⁷ cf. *zubi* ‘bridge’ [*< *lumber-over*], *azpi* ‘beneath’ [*< *hatz-bi* ‘traces/fingers-over’], *ibi* ‘ford’ [*< *hur-bi* ‘*water-over’], etc. (cf. Lakarra 2015c). It is strange that different domestic animal names – *zaldi* ‘horse’, *idi* ‘ox’, *ardi* ‘sheep’ (dialectally ‘flea’), *ahardi* ‘sow’ – are formed based on *di/-di* (*< *din*) ‘to come’, as in certain African languages (cf. Dimmendaal 2011).³⁸

The development of Monosyllabic Root Theory (Lakarra 1995) facilitates the discovery of very old morphological processes of word formation, like reduplication and

prefixation, and allows one to broaden old lexical families – a crucial instrument in reconstruction – or the establishment of other unknown ones:

* <i>das</i>	:	<i>lats</i>	:	<i>adats</i>	:	<i>aldats</i>	:	<i>jatsi</i>	:	<i>arrats</i>
		‘stream’		‘mane’		‘incline’		‘to descend’		‘dusk’
* <i>den</i>	:	<i>lehen</i>	:	<i>eden</i>	:	<i>ezen</i>	:	<i>eten</i>	:	<i>arren</i>
		‘before’		‘poison’		‘sting’		‘to break’		‘please’
* <i>dur</i>	:	<i>lur</i>	:	_____	:	(<i>h</i>) <i>andur</i>	:	<i>urri</i>	:	(<i>tx</i>) <i>inaurri</i>
		‘land’		‘		‘cruel’		‘scarce’		‘ant’

Naturally, the products of these processes belong to strata prior to, for example, the transparent *goibel* ‘sad’ or *ikertzaile* ‘investigator’ (composition and suffixation); see Section 8 on the *a quo* of prefixation and reduplication. New loanwords have been discovered, sometimes such apparently pure words as *alu* ‘vulva’, *eskatu* ‘to ask for’, *zemai* ‘threat’, *alhatu* ‘to graze’, *alhaba* ‘daughter’, *ilhoba* ‘grandchild; niece/nephew’ (both of the last two with the suffix *-ba* of family words), *oihu* ‘shout’, *aupa* ‘go!’, etc. (the latter five from Gascon; cf. Lakarra 2015a), the families of other previously known ones (such as the cited *berna*) have been extended, and other ones assumed by Mitxelena have been confirmed (*abagada-une* ‘occasion’ < Sp *vegada*, *dollor* ‘bad, poor’ < Rom *trollo* ‘bad fish,’ etc.), obtaining the exact and previously unknown origin (cf. Lakarra 2008d).³⁹ There is no doubt that more loanwords are yet to be discovered, although the main line of research should not be in that direction, as Mitxelena already intuited 50 years ago.

6 TYPOLOGICAL CHANGE AND CHANGE IN THE CANONICAL ROOT STRUCTURE

6.1 Monosyllabism and the need for a new typology

A monosyllabic root, with reduplications and – above all – with prefixes, is not what we expect of a language like historic Basque, typically associated with an agglutinative structure, SOV order, abundant suffixation on noun, and very rich verb agreement. In fact, in similar languages such as the Turkic or Australian languages – with the well-known exception of grammatical forms and phonosymbolisms – disyllabic roots are clearly in the majority, reaching 100% in the case of Uralic languages (cf. Bakrò-Nagy 1992).⁴⁰

The situation in OPB and much later stages – the geographical filter that monosyllables but not most disyllables go through corresponds at most to OCB (see Section 5.4) – is very different, with scarcely any inflection and VO order. Authors such as Trask (1977) – although later he may have recanted in his 1997 text, but not in the appendix of his “Grammar” chapter – and Gómez (1994) and Gómez and Sainz (1995) gave clear indications that PB was VO (SVO according to Trask, VSO according to Gómez & Sainz), as well as providing arguments for the existence of an “impersonal” phase without agglutinative personal agreement nor TAM in the verb.

6.2 Irregularities and new typology

It is easy to show that the characterization of Basque as a perfect agglutinative SOV-type language (as in Trask 1997, 1998) only corresponds to recent phases in its evolution, but not to previous (prehistoric) others or the oldest reconstructable periods.

Irregularities for an SOV language have been pointed out (cf. Lakarra 2005, 2006a) such as the Noun-Adjective order (which, however, Greenberg 1963 and follower typologists considered irrelevant).⁴¹ We may add that we do not have single CVC roots until late, quite distinct from the CV(C)CV pattern in Uralic (see Bakrò-Nagy 1992), there is no vowel harmony as in Uralic and other agglutinative families, and there are indications of VO order, in contrast to the SOV of Uralic, Turkic, Mongol, etc. Nor does the first-syllable accents of these language families appear to be old in B, which explains the scarcity and late character of suffixes and postpositions (cf. Lakarra 1997a, Sarasola 1997).

In the oldest part of the case system (the “indeterminate declension”)⁴² a biunivocal relationship between form and function does not appear to have existed, unlike in Dravidian and other agglutinative languages, but rather, as in Tibeto-Burman (cf. Bhat 2000) there was a kind of general locative (modern inessive marker *-n*) that can be found in archaisms like *barru-a-n-goak* (15th c.) *Be-n-go* (onomastic); and deictics like *ha-n-dik* ‘from there’, *heme-n-dik*, ‘from now on’ etc.

The final-syllable accent, i.e. on the monosyllabic root of old disyllables, implies the existence of prefixes, too, in the noun phrase, besides those on the verb pointed out by Trask (1977), together with a few others like the already mentioned **da-*. It is not just that, at a certain point, the previously prefixal language became suffixal, but rather that some of these prepositions and prefixes – *za-*, *le-*, *da-*, *de-*, etc.: cf. *basa-tza* ‘muddy place’, *saltzai-le* ‘sell-er’, *etxe-r-a(t)* ‘(to go) home’, *elur-te* ‘snowfall’, have cognates in just as many other words where they are suffixes (cf. Lakarra 2006b), so that we must conclude that they emigrated to the right before becoming fossilized.

Perhaps more attention than that received in Basque grammars (it does not appear in any of them) should be given to the structure of the sociative coordination X-COMITATIVE (*-gaz/-kin*) Y-case = “X-case_i and/with Y-case_i”, as in the following examples:

- Gloria Patri/*Biz gloria Aitearekin semearentzat*/Biz gloria espiritu santuarentzat. ‘Gloria Patri/Glory be to Father and Son/Glory be to the Holy Ghost’ (lit. ‘with the Father for the Son’) (16th-century Gipuzkoan Psalm Miserere: modern ed. by L. Akesolo, *Olerti* 1982).
- *Oguiagaz hura oragaz heroen elicatura*, ‘Water and bread, sustenance for the madmen and dogs’ (lit. ‘with bread, water. . . with dogs for madmen’) (*RS* 1596, no. 246, ed. Lakarra 1996).

According to Stassen (2000)⁴³ the following generalizations are universal:

Tendencies in the casedness – *and/with* correlation:

- 1 If a language is Cased, it will tend to have *and*-status.
- 2 If a language has *with*-status, it will tend to be Non-Cased (44).

Tendencies in the tensedness – *and/with* correlation:

- 3 If a language is Tensed, it will tend to have *and*-status.
- 4 If a language has *with*-status, it will tend to be Non-Tensed (p. 46).

This fits perfectly what, by other means, we reconstruct for OPB (and perhaps later): monosyllabic words, without a case system, or verb inflection, with prepositions and prefixes and without postpositions or suffixes; VO order, not SOV, a closed adjective category, impersonal verb and without TAM, etc., i.e., much more close to the isolate

than to the agglutinative type. This is all very different from what we find in present and historical Basque.

Since 2005 we have argued that diachronic holistic typology can give us some indication about the existence in the history of Basque – as in Munda (Donegan 1993; Donegan and Stampe 1983, 2004) or Tani (Post 2006, 2009) – of a drift from a structure similar to the Mon-Khmer toward one approximating more that of the modern Munda languages:

MUNDA *Phrase Accent*: Falling (initial); *Word Order* : Variable – SOV, AN, Postpositional; *Syntax* : Case, Verb Agreement; *Word Canon* : Trochaic, Dactylic; *Morphology* : Agglutinative, Suffixing, Polysynthetic; *Timing* : Isosyllabic, Isomoric; *Syllable Canon* : (C)V(C); *Consonant* : Stable, Geminate Clusters; *Tone/Register* : Level Tone (Korku only); *Vocalism* : Stable, Monophthongal, Harmonic.

MON-KHMER *Phrase Accent*: Rising (Final); *Word Order* : Rigid – SVO, NA, Prepositional; *Syntax* : Analytic; *Word Canon* : Iambic, Monosyllabic; *Morphology* : Fusional, Prefixing or Isolating; *Timing* : Isoaccentual; *Syllable Canon* : (C)V- or (C)(C)V(C)(C); *Consonantism* : Shifting, Tonogenetic, Non-Geminate Clusters; *Tone/Register* : Contour Tones/Register; *Vocalism* : Shifting, Diphthongal, Reductive

(Donegan and Stampe 2004: 3, 16)

As in other instances of drift, the phonological evolution is consistent with the morphology and basic syntax of the language (see Lakarra 2005): development of nasal vowels, voiceless plosives in initial position, word-initial vowels and open syllables, and so on. Finally, in the same way as Dravidian – which doubles its consonantal inventory in the drift from the protolanguage to modern languages – we go from 11 consonants in Old Proto-Basque to 16 in Late Proto-Basque, and to some 20–22 in the modern dialects.

6.3 Changes in the canonical root structure

In order to arrive at the disyllabism of the majority of roots in modern B, in the drift referred to, we have to assume multiple phonological and morphological processes that conspired to erase and alter the monosyllabic constraints in favor of other larger ones (Lakarra 2009a): V-metatheses (**ha(t)s-la(b)ur* > **hasnaur* > *hausnar* ‘to chew, ruminate’) or C-metatheses (**edazun* > **ezadun* > *eza.un* > *ezagun* ‘known’); dissimilations (**buru-bar* > *burar* > *bular* ‘breast, chest’); and assimilations (*zin-hets-i* > *sin-hetsi* ‘to believe’) – especially against sibilant homorganic consonants and rhotics in roots – as well as contractions of polysyllables (*jabe* ‘owner’ < **e-da-dun-e*); additions of the hiatus-breaking -g- (*hogen* ‘lack’ < **ho.en* < Lat *offende(re)*, *eza.un* > *ezagun* ‘to known’); neutralizations and deletions in final position in the first element (*larre* ‘field’ + *mendi* ‘mountain’ > *Larramendi* (topon.), *buru* ‘head’ + *hezur* ‘bone’ > *burhezur* ‘skull’; and so on.

Moreover, noun prefixes (*sabel* ‘belly’ < **sa-* ‘within’ + *-bel* ‘black’) and verb prefixes (**e-da-ra-dan-* ‘PREP.’ + ‘DIRECTIONAL’ + ‘CAUSATIVE’ + ROOT + -o ‘COMPLETIVE’ > **ardano* > *ardo* ‘wine’) became fossilized, and reduplications stopped being productive (**zen* > *zezen* ‘bull’ but ***bebehi* < *behi* ‘cow’ or ***babalea* < Lat *ballena*), and there is other proof – not just formal indications – of prefix > suffix change (see Section 6.2).

While the morphological changes appear to be older – the fossilization of reduplication is prior to **d-* > *l-*, which, in its turn, is older than LPB, there are no traces of prefixation

in old Latin loanwords (although there are inherited words of IVC shape [*sarats* ‘sauce’ < **sa-latz*, [Lat *salicem* > B *zarika*]; cf. Lakarra 2015b) – phonological changes seem to be more recent or even later than OCB. That is consistent with what we have seen about the later character of disyllabic roots and the survival of multiple CVCs after OPB.⁴⁴

7 GRAMMATICALIZATION AND RECONSTRUCTION OF PROTO-BASQUE

7.1 Introduction

Meillet, Kurylowicz, and other leading diachronic linguists pointed out the importance for linguistic reconstruction of the development from lexical to grammatical morphemes and from grammatical morphemes to even more grammaticalized forms. Whether it is to be understood as a primary process or as following from more basic phenomena, and whether or not it is always unidirectional (cf. Campbell ed. 2001, Fischer et al. eds. 2004, etc.), grammaticalization takes place in languages everywhere, and its study may contribute greatly to Basque historical linguistics, as it has contributed to knowledge of the history of many other languages.

Certain grammaticalizations are known to have taken place in Basque. Among relatively well-understood phenomena, we find the development of articles from demonstratives (see Manterola 2015), the more recent development of the Gipuzkoan interrogative particle *al* and the quite diverse and interesting evolution of the auxiliary verbs, etc.

The holism of the phenomenon (phonetic and semantic erosion) and multiple parallels in different geographical and genetic languages (see Heine and Kuteva 2002) leads us to acknowledge its effects on markers like those of the dative (*-i* < **nin* ‘GIVE’); unfinished aspect (*da-* < *dar* ‘SIT’), the plural *-de* < **den* ‘FINISH’); the prosecutive/ablative and adjectival suffix (*-ti* < **din*⁴⁵ ‘COME’); superiority comparison (*-ago* < **ha* ‘demonstr. of 3rd level’ + *-go* ‘TO PASS’); and the adverbial suffixes of mood (completive) *-to* and *-ro* as well as the causative *ra-* (< **lo-*), from **don* ‘PUT’; the old comparative and distant familiarity suffix (*-so* < **san* ‘TO SAY’); the conjunction (*da*); the modal (*-la*); and the coordinations (*e-ta* ‘and’, *e-do* ‘or’); etc. (cf. Lakarra 2013b, 2017).

It is obvious that without the help offered by analyzing grammaticalization, many of these reconstructions would still be unknown.

7.2 Serial verbs in Proto-Basque?

It seems difficult to accept that, in a language in which *aita-ren-tza-ko-a-k Bermio-ko-e-i har d-i-e-za-z-ki-e-ke-gu* (“we can get those that are for father from those from Bermeo”) is a typical sentence, there would be no verbal or noun inflection at an earlier stage in its history. Nevertheless, we have more evidence besides that already cited (Sections 6.1 and 6.2) to defend the proposal that Basque lacked morphology in an earlier period: thus, the analysis of *dago* or *dakar* not as *d-* ‘3rd pers.’ + *a* ‘pres.’ + *go(n)* ‘to be’ and *d-* ‘3rd pers.’ + *a* ‘pres.’ + *kar* ‘to bring’ but as \emptyset ‘3rd p.’ + *da* ‘IMPERF’ + *go(n)* ‘to be’, etc. plus the testimony of *nago* ‘I am’ (< **ni-dago*, *ni* ‘I’) and *nakar* ‘he brings me’ (< **ni-dakar*) made Trask (1977) and Gómez (1994) assume that at some point the verb was “impersonal” (**dago*, **dakar*) and that only later did personal markers agglutinate.

Trask stated that *da-go* and *da-kar* should be analyzed like that, with the root preceded by a prefix or auxiliary of indeterminate aspect, in a very different way from other SOV

languages. We may add that *da-* (cf. Section 7.5) can be taken to be the grammaticalization of **dar* ‘SIT’ (**e-darr-i* > *jarri* ‘to sit’), the best-known source of this type of marker in verb and the locative on the NP (B *-a/-t*: *Zarautz-a* ‘to Zarautz’, *hibaira-t* ‘to the river’). Therefore, *da-go* [*< *dar- *gon* ‘SIT’-‘STAY’] came to be a kind of asymmetric serial verb (see Aikhenvald 2006). We have testimonies of grammaticalization of typical serial verbs (cf. Lakarra 2008a, 2017) in many nominal cases: dative *-i*, ergative *-k* < **ga* < **gon* ‘to stay’, locative *-a* < **dar* ‘SIT’, *raino* ‘until’ < **r-a(d)in-no* [epenthesis-SIT-COME-GO], etc.

7.3 Irregularities in the CRS, old root extensions, and reconstruction of the verb

Lafon (1943) in his work on the Archaic Basque verb (see now Mounole 2011) offered a long series of root structures, including monosyllables in V *da* ‘copula’, VC *utz-i* ‘to leave’, and CVC *e-thorr-i* ‘to come’, disyllables CVCVC *jakin* ‘to know’, CVCV *jagi* ‘to get up’, and trisyllables VVCCVCV *aurtiki* ‘to throw’, etc.

In reality (cf. Lakarra 2008c, 2017), the CRS of verbs was CVC and to that root one could add at least two prefixes (causative *ra-* and directional/applicative *da-*) or combinations of both,⁴⁶ as well as the ‘initial vowel’ (mirror image of the Bantu “final V”).⁴⁷ Besides this, contracted forms of these exist:

e-thorr-i ‘to come’, **e(d)utz-i* > *utzi*, *itxi* ‘to leave’

**e-da-khin* > *jakin* ‘to know’, **e-da-duts-i* > *jautsi* ‘to go down’

e-ra-bil-i ‘to use’

**e-da-ra-don-tz-i* > *jarauntsi* ‘to inherit’, **e-da-ra-gotz-i* > *urgatzi* ‘to help’

The prefixes on these verbs are fossils and have not extended to new roots since the prehistoric era (the end of LPB?) – there are none in verbs taken from Latin – and in the case of the causative it was substituted before the first texts by the suffix *-erazo/-arazi*. There are those who have seen in the destinative *-ra* (*mendi-ra* ‘to the mountain’, *egite-ra* ‘to do’, etc.) the origin of the causative (*e-ra-bil-i* ‘to use’ < ‘*to make walk’), but this is unlikely: the two categories do not correspond to the same network of grammaticalization (cf. Heine and Kuteva 2002) and, what is more, we would have a case of suffix (active) > prefix (fossil) in a language with drift toward agglutination and SOV order. It is, moreover, unnecessary because the old auxiliary PUT **lon* (> *i-ro-*) is enough to explain causative *-ra* after *-o* > *-a* in composition and derivation, and regular VIV > VrV in loanwords and inherited words until the Early Middle Age.⁴⁸

As we have seen (in Sections 7.1 and 7.2), *da-* (< **dar*) had locative and (indeterminate) aspectual value, typical in the grammaticalization of SIT. Trask (1977) found this morpheme in conjugated forms (*dago*, *dakhar*, etc.), but it is also present in some non-conjugated ones, which explains the enormous abundance of *-a-* after yod in those pointed out by Mitxelena (*FHV*) and also in certain adjective root and nouns like *la-bur* ‘short’, *la-bar* ‘edge of cliff’, etc.

The reconstruction of elements preceding the old verb root shows the value of combining the notions of CRS and grammaticalization: we are able to reduce the polyformism assumed by Lafon and others and we are able to get a more precise idea of the old morphology and syntax, which functioned to the left and not to the right⁴⁹ as in historical times (see Trask 1977, Lakarra 2008c, 2017).

7.4 The Proto-Basque verb as a closed class

There has been a fairly widespread belief that the synthetic Basque verb, with multiple tense and agreement markers, has gradually deteriorated since its Golden Age in favor of periphrastic structures borrowed from neighboring languages. While archaic attestations (prior to 1600) document conjugated forms matching some six dozen (synthetic) verbs, currently there are no more than a dozen: *ekarri* ‘to bring’, *etorri* ‘to come’, *eroan-eraman* ‘to take’, *ibili* ‘to walk’, *egon* ‘to be’, *joan* ‘to go’, *jakin* ‘to know’, and little more in addition to the auxiliaries. This has led to the notion (cf. Gómez and Sainz 1995) that, if we were to go back in time, we would come across many more synthetic verbs and perhaps that all the verbs were conjugated synthetically at some more or less remote time.

However, we can state with some certainty that this idea is implausible. To start with, only some of the verbs with the prefix **e-* (> *e-*, *i-*, *j-* in known phonological conditions) are documented in conjugated forms: *ekarri* → *dakart*, *zekarren*, etc., does, but not *erosi* ‘to buy’, *iritsi* ‘to arrive’ or *jarri* ‘to put (down), sit down’, etc. Nor does any verb without that prefix have synthetic forms: *apurtu* ‘to break’, *bidali* ‘to send’, *hartu* ‘to take’, *sartu* ‘to enter’, etc. According to Mitxelena-Sarasola (1987–2005), we would have around 200 inherited words with *e-*, *i-*, *j-* and that is the maximum number of candidates for being old Basque verbs. Nevertheless, most of them do not reflect synthetic forms, and they even have serious problems in their structure for being considered old verbs: for example, *itsusi* ‘ugly’ or *itzali* ‘to turn off, put out’ cannot be analyzed as ***i-tsus-i* ‘ugly’, ***i-tzal-i* ‘to be out’, with root-initial affricates. On the other hand, we do see that over a dozen synthetic verbs only have conjugated forms in the imperative (still productive modern agglutinations), that many others have never developed more than just a small part of their potentialities (*dario* ‘he/she/it flows’ from *jario* but not ***zenerizkidakeen* (perhaps interpretable, if it existed, as ‘they can flow to me from you’), etc. nor even more simple forms outside the 3rd pers.), and, if we pay attention to the internal structure of the verbs (Section 7.3) – prefix *da-*, *ra-* or combinations of them – we can assume that no more than two or three dozen roots were possibly ever conjugated. This is similar to what Pawley (2006) and others have described for a lot of Australian languages, as well as other languages of New Guinea, Siberia, and the Americas, i.e., that there is not an open class of inflected verbs with PERSON and TAM markers, but a few semantically basic verbs with such a structure besides several auxiliaries used with abundant “converbs” for the rest.

This is consistent with what was observed in Sections 6.2 and 7.2 and is additional evidence for the isolate > agglutinative evolution.

7.5 Grammaticalization and (pre)history of Basque morphemes

If we define as “primary grammaticalization” (PG) the conversion of lexical roots into grammatical morphemes and as “secondary grammaticalization” (SG) any other later change of the grammaticalized morphemes into more grammatical morphemes, bearing in mind that CVC was the CRS of the old lexemes, there are important consequences for the history of morphemes. It is, in addition, one of the strongest pieces of evidence for the existence of a CRS CVC in lexical roots. Thus, for example, we find that:

- 1 There can be no disyllabic or polysyllabic PG: *-tate* and *-tasun* are loanwords (Latin *-tatem*) or secondary amalgams (*-tasun* < *-t-ar-zu-n*). Likewise, *-heta* (archaic variant of the toponymic suffix *-eta*) and *-zaha* (the same for *-za*) are not PG but fusions of other morphemes: **he* + *-ta* (< *da*), **-za* + *-ha*, etc.
- 2 There are no simple PGs with codas: **-gan* was not the old inessive morph, but an amalgam *-ga* + *-n* (*pace* Jakobsen, Trask, and de Rijk); sociative *-kin* < **-k-i-de* + *-n*, etc.
- 3 There are no CV lexical morphemes but rather this is the Canonical Morpheme Structure of the PGs: *lo* (< **don*) ‘to sleep’, *su* ‘fire’ (< *sun-/sur-*), etc. have lost $-C_2$ in composition or for other reasons (e.g. reanalysis).
- 4 All CV morphemes come from CVC, i.e. they are PGs. E.g. *da/-da* < **dar* ‘to sit’, *-di/-ti* < **din* ‘to come’, etc.
- 5 V-, -V < CV and C-, -C < CV. The suffix and prefix in the dative *-i/i-* come from **ni-* (< **nin* ‘to give’) and *-t* ‘1st per.sing.’, *-k* ‘2nd per.sing.’ in *-da*, *-ga*. The *e-* in old verbs comes from **Ce-* (**he-*), cf. (*h*)*eta*, **her* ‘close, closed’. The instrumental marker *-z* is reconstructed as **-zV* (cf. *za* ‘pl.’ and **zan* ‘to be’, plus the pleonastic *-az* < **-za* + *za*) and the agreements *n-* ‘1st per.sing.’, *-z-* ‘2nd per.pl.’ in the pronouns *ni*, *zu*, etc.
- 6 $-VC < *(C)V\#C(V)$. This is a subcase of (2) and (5); thus *-ak* (nom.pl., ERG.sg. and ERG.pl⁵⁰) is *-(h)a* + *g(a)*. Manterola (2015) reconstructs as **ha* (without coda, not **har*, as it has been to date) as the demonstrative and third-level article. The finals on $-VC$ that Uhlenbeck (1942) took to be suffixes (*-ats*, *-ar*, etc.) are roots that have lost C- in composition (cf. *adats* ‘mane’, *aldats* ‘incline’, *ordots* ‘male’, etc.), not suffixes or amalgams.
- 7 In $-rV$ (DAT. *-ri*, adlative *-ra*, GEN. *-ren*, etc.), the *-r* is epenthetic between a stem in $-V$ and a case marker V-. It would apparently enter into (4), but *r-* is impossible in roots and words, and it is unlikely that all the $-rVs$ in declension come from $-IV(C)$ by $*VIV > VrV$. $-V$ (< $*CV < *CVC$) is the true suffix with later epenthetic *r*. Note that the stems in $-V$ necessary for the change mentioned could only be developed later in a language with CVC Canonical Root Structure.
- 8 Postpositions and other morphemes may be differentiated: morpheme < CVC root/postposition < words (and constructions greater than CVC: *buru* ‘head’, *begi* ‘eye’, *aurre* ‘front’, *atze* ‘back’, *ondoren* ‘after’). They could also be based on disyllables, and their internal structure (composition, derivation, as well as loanwords) is much more obvious compared to other morphemes; the degree of grammaticalization and the antiquity of postpositions is very little (cf. Hualde & Ortiz de Urbina 2003).

The combined study of CRS and CMS and grammaticalization may offer yet more advances in the reconstruction of PB (see Lakarra 2013b, 2016, 2017, 2018).

8 ADVANCES IN CHRONOLOGY AND PREHISTORICAL PERIODIZATION

Unfortunately, Basque lacks any known relatives, which makes it impossible to use for comparative history (see Section 3). This does not make the study of the history of the language an impossible task, as has sometimes been claimed, but it does require developing other means for doing so. Basque historical linguists were late in embracing the most productive approach in any language isolate, internal reconstruction (which practically

began with Mitxelena), both because the necessary philological work has only been undertaken recently (publication of important texts after 1975, *DGV* in 1987–2005, for example) and because important developments in historical linguistics have had very little impact in the field of Basque studies until quite recently.

Still, we do have important prehistorical and protohistorical phenomena (reduplication, prefixes) and changes ($*d- > l-$, $T- > D-$, $-n- > -h-$, $*h_{2/3} > h_1$, prefix \rightarrow suffix, etc. (see *supra*) that we can attempt to connect in a relative chronology, thereby making some progress in establishing periods and strata in the development of Basque (cf. Lakarra 2015b). Thus, the existence of reduplication for patterns such as $*dVC$, $*nVC$, $*zVC$ and $*gVC$ (see Section 5.4) but not for $*IVC$ is probably related to the prehistorical change (already noted by Mitxelena 1957) $*d- > l-$: i.e., reduplication had ceased to be productive before that rule came into force. If *sarats* ‘willow’ comes from $*sa-latz$,⁵¹ with the prefix *sa-* and root *latz* ‘rough, coarse, harsh’ ($< *datz$), then clearly that type of prefixation survived until a more recent time than reduplication and subsequent to $*d- > l-$, but neither of these phenomena was in force when Basque-Latin contact began, since they are not present in any loanword, however old it may be.

Bearing in mind that there are only aspirated voiceless plosives in initial position in old verb roots – not plain voiceless or pure aspiration: *ekharri* ‘to bring’, *ekhusi* ‘to see’, *ethorri* ‘to come’, not $**ekarri$, or $**eharri$, etc. –⁵² it is possible that such consonants are archaic, maintained without suffering $*T^h > h-$ by the addition of the prefix and not later aspirations like those of many loanwords (*khoroa* ‘crown’, *phike* ‘tar, pitch’ $<$ Latin *coronam*, *picem*, etc.) or inherited words (*khal-te* ‘loss’ $<$ *gal-du* ‘to lose’). As verbs with the prefix $*e-$ ($> e-, i-, j-$ according to known contexts) can only be conjugated, we can infer that the prefix ceased to be added to the CVC bases before $T^h > h-$, so that the verbs that developed thereafter, whatever the C- and both in loanwords and inherited words, lack any synthetic form and are conjugated periphrastically.

Given the characteristics of the corpus of the language (Lakarra 1997a, Ulibarri 2013), absolute internal chronologies are scarce and uncertain. Thanks to the *DGV*, we have many interesting lexical and some morphosyntactic clues: e.g. the first appearance of the interrogative particle *al*, the *antequem* of the Aresti-Linschmann Law on neutral or intensive possessives – widespread up to the 18th century, which then disappeared in diverse ways among the different dialects – and the placing of *oso* ‘very’ (after the adjective or at the end of the phrase up to the same century).⁵³

In phonology, when it comes to studying loanwords, the use of the chronological sequences elaborated by Straka (1954–57) and others in the evolution of Latin-Romance offer a good number of absolute and relative chronologies; cf. Guiter (1989):

200	300	400	1000	1100	1200
p- > b-	n > Ø, nt > nd	l > r	ll > l		
t- > d-	mp > mb, nd > n			nn > n	
k- > g-					
mb > m					

Research regarding chronology in the inherited lexicon has been more limited although it may progress with works such as that of Hualde (in press). For instance, based on the phonetic changes established in *FHV* and other work by Mitxelena and others, we can try to establish that $-n- > -h-$ must be prior to $h_{3/2} > h_1$ (Lat *arena* $>$ OCB $*areha$ $>$ *harea* ‘sand’, OCB $*enuskara$ $>$ $*ehuskara$ $>$ *heuskara* $>$ *euskara* ‘Basque language’)⁵⁴ and prior to those nasal vowels reconstructable for the OCB (*ardâô* ‘wine’, *gatzââ* ‘cheese’,

etc.). In turn, *arrâi* ‘fish’ (still attested at the end of the 16th century) predates *arrai(n)*, as does **lukâika* ‘sausage’ (< Lat *lucanica*) to *lukai(n)ka*. And **hVh* > *ØV(h)* is prior to $h_{3/2}$ > h_1 (**hur-bar-bi* > **huhbah(b)i* > **uhbahi* > *ibahi* > *hibai* ‘river’).

The old strata and variants are easier to recognize in loanwords (see Mitxelena 1974): *gela* ‘room’ < Lat *cellam* is older than *zeru* ‘sky’ (< *tselu* < Lat *caelum*) due to the fact that the palatalization of *k-* is a Romance phenomenon, so that *gela* must have already existed in B when *caelum* ‘sky’ was first palatalized and went on to become an affricate sibilant later. *Zeru* has *-l-* > *-r-* in contrast to its later variant *zelü*, but in *gela* the *-l-* comes from a fortis lateral (= Lat geminate) like PB **beLe* ‘crow, raven’, not from a singleton Lat *-l-* (*gula* > B *gura* ‘to want’) or B lenis (toponym *Araba*, cf. Rom *Álava*). In *baradizu/paradisu* ‘paradise’, there is a voiced/voiceless C in line with the antiquity of the dorsal/apical sibilant (this in recent loanwords); if we add *paraiso* we would have the modern *-o* with *p-* and *s-* but older *-u* with *-z-* and with *b-*.

The words *sabel* ‘belly’ and *zazen* ‘bull’ (prefixation and reduplication; cf. Section 5.2.2) are much less transparent in their structure than *ogibide* ‘job’ < **‘bread’-‘way’*⁵⁵ (later composition) and *Hirutasun* [‘Trinity’ < *hiru* ‘*three’ + *-tasun* ‘-ity’] (modern suffixation). Prefixation and reduplication are not just fossilized by the time of the first texts written in B, but they do not even apply to the oldest known loanwords. Derivative suffixes, on other hand, are scarce outside the highest literary language still in the 18th century (cf. Lakarra 1997a, Sarasola 1997).

In verbal inflection, the older nature of prefixes with respect to suffixes is clear. Regarding nominal inflection, the number and size of case markers in the noun phrase has grown substantially, due in part to the phenomenon known as *surdeclinaison* (*zaldi-ar-en-tzat* = horse-the-GEN-for; cf. Lafitte 1944). In the transitive auxiliaries of periphrastic forms in irrealis moods, *iron* ‘can’ (from the same root as the causative prefix *-ra-*) appears to have been more widespread at some time, although in modern times it has only survived in eastern areas; **ezan* (today central-eastern) is documented in the far western and southern areas in the 16th and 17th centuries; *egin* ‘to do, make’ is of general use as a main verb, but only in western varieties (Biz, A, G) is it used as an auxiliary. Among the transitive auxiliaries used with dative agreement (**nin*, **edutsi*, **eradun*) the first appears to be the oldest and most grammaticalized (widespread but barely attested in Biz), the second is an innovation from Biz and A, and the third is documented not just in eastern areas, as in the present, but also in G and A, even if only in plural in these latter dialects (cf. Ariztimuño 2013).⁵⁶

As for the article – on which the singular and plural case systems are based – not only was it missing in LPB but also in Aquitanian and in Pyrenean Basque (see Manterola 2015). Its grammaticalization is more recent than in Hispanic Romance languages (after the 8th century).

9 ON OLD COMMON BASQUE⁵⁷

Perhaps as a result of the limited development of many diachronic issues in B or because of the deep-rooted myth of the immovable nature of B, we know much more about the modern dialects than about their origin or about the changes that Post-Medieval B underwent in these dialects. In practice if not in theory, the B dialects are conceived of by many as old or eternal, as if the dialects have not changed at all – despite changes in structure and different changes according to territory. Moreover, the testimony of the old southern dialect (Landucci’s vocabulary, Lazarraga’s texts), the history of Biz and L (with abundant old evidence and without any continuity until Contemporary B), and the R attestations of 1615–1617 (see Lakarra 2014), as well as the practice of standard

diachronic linguistics, have still not put an end to opinions based mainly on the modern situation of the language.

The dialects established by Bonaparte (1866), Mitxelena (1964), and Zuazo (1998) are recent, and there is little difference among them, without offering the possibility of taking us back to the speech of the old Vasconic or Vascoïd tribes (at the beginning of the Common Era or the end of the previous age) or to LPB. Mitxelena (1981) postulated the notion that convergence of the most differentiated forms of speech beginning in PB – which Aquitanian and the Pyrenean B did not undergo – could have taken place in centuries after the weakening and fall of the Roman Empire, when the B-speaking populations resisted Visigoths and Franks.

It is clear that all B dialects have shared many innovations after PB in phonology (voicing of initial consonants, lenition of sonorants, metathesis of /h/, development of nasal vowels, neutralizations and deletion of vowels in final position in the first element of compound and derived words, etc.) and grammar (the article – initially including three degrees of deixis –, intensive personal pronouns, the plural, most definite declension and almost all indefinite declension, the tense system, the distribution of synthetic and periphrastic forms, most of the auxiliaries, the allocutive, verb periphrasis, the development of *-zu* ‘you’ from plural to singular, etc.). We must assume that if all B dialects came directly from PB, their differences would be much greater, almost certainly having become different languages.⁵⁸ We would have to assume therefore that the historical dialects known to us are the result of fragmentation of a common language dating from approximately the 5th to 6th centuries in the Common Era, which would have been the product of a convergence process among distinct and previously more dispersed B forms of speech.

Yet Mitxelena based his ideas essentially on non-linguistic arguments – i.e., that conditions after the fall of the Roman Empire would be those most suited to reducing the dependency on foreign powers and reinforcing cohesion and internal organization (following the historical model of Barbero and Vigil (1965), which most historians currently reject) – in order to demonstrate the need to assume an OCB, without attempting a precise definition of such a protolanguage, and in particular its differences (relevant innovations) with regard to LPB, an indispensable task for its justification from a linguistic point of view and something that has only recently received some attention.

We tentatively present here a series of phonological innovations that could have occurred between LPB and OCB or, at the very least, prior to the fragmentation of the latter and that perhaps may serve to differentiate both protolanguages:

- | | |
|--|--|
| 1 T- > D- ; | 9 * <i>d</i> ₁ > <i>ø</i> /V __ V; |
| 2 *- <i>n</i> - > - <i>h</i> -; | 10 <i>a</i> - <i>o</i> > <i>o</i> - <i>a</i> in the verb; |
| 3 Nasal vowels; | 11 - <i>n</i> > - <i>r</i> / __ # |
| 4 Diphthongs; | 12 * <i>b</i> -, * <i>k</i> - > <i>ø</i> -; |
| 5 *- <i>n</i> - > - <i>n</i> ; | 13 - <i>l</i> - > - <i>r</i> -; |
| 6 * <i>r</i> > - <i>h</i> ; | 14 - <i>i</i> ₂ / <i>-u</i> ₂ > <i>ø</i> / __ #; |
| 7 * <i>hVh</i> > <i>øVh</i> and * <i>hC</i> > <i>øC</i>); | 15 - <i>V</i> ₃ > <i>ø</i> / __ # |
| 8 * <i>e</i> - > <i>j</i> / __ V | |

On the other hand, without refuting the existence of horizontal innovations among the B varieties – which have never ceased to be in contact – it is obvious that the flat dialectal tree commonly used in the field of B linguistics is implausible and antihistorical (cf. Austronesian or IE). Furthermore, it impedes the establishment of historical and geographical timelines⁵⁹ and hierarchies in the evolution of features and varieties.

Dialectal classification must be based on the oldest innovations and the bipartite branchings that they produce. It cannot depend on the number of traits that could be used to separate one dialect from another. Unquestionably, the dating of all the historical dialects of B – Biz, NHNa, and Z, for example – cannot be the same insofar as the particular innovations that each one of them shows are very different in their age (old in the case of Z, more recent in the case of NHNa). Thus, for example, considering phenomena like (a) voicing (or devoicing) of plosives after *l/n*, (b) palatalizations after *(V)i*, (c) dissimilation $a + a > ea$, and (d) grammaticalization of *egin* as AUX, it would appear that:

- (1) is – against what is usually thought – an innovation of R and Z as opposed to the Central-Western voiced consonants (old lenis) and Pyrenean Romance forms of speech, further east than R and Z;
- (2) is a particular and later innovation of Biz; and
- (3)–(4) are innovations that, as well as Biz, affect A and most of G.⁶⁰

The tree derived from these and other innovations would be closer to (2) (below) than to (1), despite the fact that this is the partition with the greatest number of followers (cf. Bonaparte, Lacombe, Uhlenbeck, etc.), especially among enthusiasts who do not make excessive use of the existing philological documentation on old Biz and A (cf. Lakarra 1996, Mounole 2015, Mounole and Lakarra 2017;⁶¹ Mitxelena rejected this classification explicitly on many occasions (1958, 1964, 1981) and there are still no arguments to change this opinion.

Following a suggestion by Mitxelena (*FHV*), in Lakarra (2014) we defended the idea that voicing after *l* and *n* (*alde* ‘side’, *handi* ‘big’) is an archaism and that the innovation is the devoicing found in R and Z. For this, in addition to Mitxelena’s observations, we are supported by the B substrate in Gascon (cf. Rohlf 1977) and the parallel of sibilants in an identical context, which are realized as fricative (= lenis) and not as affricate (=

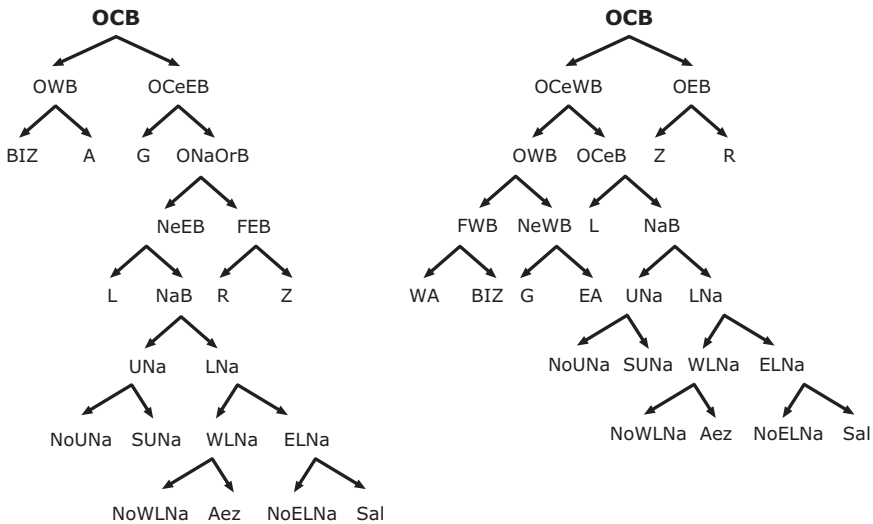


FIGURE 3.1 CLASSIFICATION OF BASQUE DIALECTS

fortes) unlike in the modern Western dialects. For that reason, R and Z are the innovators – this is perhaps the earliest dialectal innovation (= right tree in figure 3.1) that we know of – bearing in mind that affricates were opposed to fricatives and voiceless plosives to voiced plosives as fortes and lenis, respectively, in the previous system; cf. Section 4.2).⁶²

As for the question of the OCB homeland, we believe we must locate innovations chronologically and geographically rather than considering the quantity of existing modern dialects in this or that territory (cf. Janhunen 2009 for Uralic, for example). In particular, in the territory of Bizkaia, Gipuzkoa, and Araba, the dialectal division seems much simpler than in Navarre, in the geographical area that runs from Pamplona toward the north, up to the modern French border. Perhaps that would be the area where OCB developed and from which it spread. It is not advisable to interpret the most modern forms of speech, situated at the lowest level of the tree, as decisive. We should look instead to the root of the tree when searching for the homeland of OCB, given that the first fragmentations are found in the highest branches: see Janhunen's (2009) conclusions on the homeland of Proto-Uralic.

Returning to our case, the place of Samoyedic, Finno-Ugric, and Proto-Uralic would be occupied, respectively, by Zuberoan-Roncalese (= Old Eastern Basque), Eastern Low Navarrese-Salazarese (= Easternmost Navarrese), and OCB, with the oldest isogloss situated between Z-R and NoELNa-Sal – during an era that, for the moment, we cannot specify – so that we should locate in that specific place⁶³ the proto-homeland of OCB, as shown in Figure 3.1 above.

10 CONCLUSIONS

The main proof for genetic relationship among languages lies in the help it offers for the reconstruction of a common protolanguage and for studying the history of the languages in the family. The strength of the demonstration cannot be based on the quantity of alleged superficial analogies without regular phonetic connections or the reconstruction of homologies.

We consider it essential that specialists in Basque, distancing themselves from unjustified allegiance to remote agendas, analyze the facts of B diachrony according to the best philology and the most productive theories and methods of linguistic change and reconstruction, as Meillet and Mitxelena asserted.⁶⁴ Any advances via the expansion of materials – languages or protolanguages related to ours, pre-Latin loanword strata – do not appear any nearer, than they were some decades ago, so it is reasonable and necessary to opt for the application of more efficient theories and methods (cf. Haas 1969), in order to arrive at a more complete and deeper reconstruction of PB and the prehistory of the language. We defend the notion that – besides the usual internal reconstruction methods masterfully used by Mitxelena a half century ago – research on the Canonical Form of roots and morphemes; Diachronic Holistic Typology (subordinate to the search for homologies, not dedicated to pure analogies); and Grammaticalization processes may continue to contribute important advances in reconstruction. Finally, the elaboration of chronologies and periodizations – including the establishment of a minimum number of necessary (intermediate) protolanguages for the reconstruction of the prehistory of the language (as is the case with Old Common Basque, Late Proto-Basque, and Old Proto-Basque) – are unavoidable topics, as we find in any other language or family.

NOTES

- 1 Work connected to the projects “*Monumenta Linguae Vasconum* (IV): Textos Arcaicos Vascos y Euskera Antiguo” (FFI2012–37696) and “*Monumenta Linguae Vasconum* (V)” [FFI2016-76032-P], Consolidated Research Group “Historia de la Lengua Vasca y Lingüística Histórico-Comparada” by the Basque Government (GIC.IT698–13) and UFI (Training and Research Unit) UFI 11/14 at the UPV/EHU.

Acknowledgements: I thank Borja Ariztimuño, Joaquín Gorrochategui, José Ignacio Hualde, Julen Manterola, and Blanca Urgell for numerous and interesting observations and corrections of form and content, although I have not necessarily accepted all of their suggestions; all remaining errors are my own. The map and the dialectal genealogies were produced by Adur Larrea, with the collaboration of Céline Mounole, with important observations from Gidor Bilbao and Ricardo Gomez. I have found Lyle Campbell as rigorous and generous an editor as anyone could want.

Abbreviations: C = consonant, R = sonorant, S = sibilant, T = plosive, V = vowel, p = person, sing = singular, pl = plural, ERG = ergative, DAT = dative, ABS = absolutive, GEN = genitive, part = participle, TAM = Tense-Aspect-Mode, PREP = preposition(al), PG = primary grammaticalization, SG = secondary grammaticalization. *Languages:* B = Basque (language), OPB = Old Proto-Basque, LPB = Late Proto-Basque; OCB = Old Common B; IE = Indo-European, Rom = Romance; Lat = Latin, Sp = Spanish, Eng = English, Ger = German; *Basque dialects:* A = Alavese, Aez = Aezkoan, Biz = Bizkaian, EA = Eastern Alavese, ELNa = Eastern Low Navarrese, G = Gipuzkoan, FEB = Far Eastern B, FWB = Far Western B, L = Lapurdian, LNa = Low Navarrese, NaB = Navarrese B, NeEB = Near Eastern B, NeWB = Near Western B, NoELNa = North-Eastern Low Navarrese, NoUNa = North-Upper Navarrese, NoWLNa = North-Western Low Navarrese, OCeEB = Old Central-Eastern B, ONaORB = Old Navarrese-Oriental B, OCeB = Old Central B, OCeWB = Old Central-Western B, OEB = Old Eastern B, OWB = Old Western B, R = Roncalese, Sal = Salazarese, SUNa = South Upper Navarrese, UNa = Upper Navarrese, WA = Western Alavese, WLNa = Western Low Navarrese, Z = Zuberoan. Others: DGV = *Diccionario general vasco* (Mitxelena & Sarasola 1987–2005), *FHV* = *Fonética Histórica Vasca* (Mitxelena 1961).

Symbols: * = reconstructed form; ** = Undocumented and impossible.

- 2 Following Ringe (2003), we understand that the comparison between dialects is not internal but rather comparative reconstruction. In our case, its result would be OCB (see Section 9), but it is far from being approached systematically, most likely because internal reconstruction is much more compelling on almost all fronts.
- 3 It is impossible to demonstrate that languages are not related, that there is *no* genetic relationships, and therefore, it is the “believers” – as Mitxelena used to say – who are obliged to offer proofs (standard ones, and not just any old thing).
- 4 All of them defended the polygenesis of *all* languages, like Boas and Trubetzkoy as well (Lakarra 2008d).
- 5 That is, not non-professionals in linguistics but people who are unfamiliar with the methods and aims of historical linguistics and philology and those who have no experience as regards the real history of any language or family, whether they are linguists by profession or not: this is not uncommon.
- 6 For Vovin (1994), Japanese had the dubious honor of having been the language on whose origins the most ridiculous things had been said; we do not know if he had thought about B when he argued this. In Campbell’s (2013) extensive list of non-proven genetic relationship hypotheses, the extremely high proportion of combinations in which B appears is striking.

- 7 For the current sociolinguistic situation see Barreña et al. (2013); Hualde and Ortiz de Urbina (2003) is the most complete grammar in English on contemporary Basque; Mitxelena-Sarasola (1987–2015) the obligatory lexicographical, historical, and dialectal source; and Hualde, Lakarra, and Trask (1995), Trask (1997), and Martínez Areta (2013) are the most up-to-date monographs on the history of the language (particularly prehistory and internal developments). Gorrochategui, Igartua and Lakarra (eds. 2017), besides being an examination of prehistory and protohistory, is the most complete available treatment of strictly speaking historical eras.
- 8 In the school system, the “D” teaching model (in Basque with Spanish as a subject) is the most common choice now whereas the “A” model (instruction in Spanish with B as a subject) is fairly marginal. The University of the Basque Country awards degrees in both B and Spanish, with B being used at the university level for the first time in 1978; there are both television and radio stations (in the BAC) wholly in B as well as local television and radio stations in the language, which moreover is used increasingly on the Internet.
- 9 Studied in the last third of the 19th century by Luchaire and later by Mitxelena (1954b) and Gorrochategui (1984, etc.). In the southern part, one should add an important inscription found in Lerga (Navarre) and a few others discovered in the historical Vascons’ territory, as well in La Rioja and Soria (cf. Gorrochategui 2011a).
- 10 Rico (1982) points out the curious syntactic order of many Romance glosses, more befitting B than Romance.
- 11 The two best-known dialectal classifications are that of Bonaparte (B, G, SHNa, NHNa, L, WLNa, EaLNa, and Z dialects) and that of Zuazo (1998); see Martínez Areta (2013). These two classifications refer to situations around 1860 and 1990 and were not done from a diachronic point of view. The observations of Mitxelena (1958, 1961–1977, and 1964) are interesting insofar as he distances himself from Bonaparte by differentiating (as Azkue had done previously) Z and R; he also separates, for phonetic reasons, Aez and Sal from LNa and adds the southern dialect (documented in Landucci 1562 and Lazarraga [~1600]). See Section 9 and Lakarra (2014), Mounole (2015), Mounole and Lakarra (2017).
- 12 See Blust (2014) on the Proto-Ongan-Austronesian hypothesis of Blevins (2007); it is difficult not to see similarities among diverse errors revealed there and those committed in her B-IE hypothesis (Blevins 2013).
- 13 Nor does the supposed existence of two datives *-o/-a* in 3rd p., which would come from those old articles, have any real foundations – (*pace* Rijk 1981); actually, *o* → *a*/ C; cf. *deutso* ‘3p.ERG-3p.DAT’ : *deutsala* ‘3p.ERG-3p.DAT + *-la*’, *jako* ‘3p.ABS-3p.DAT’ : *jakan* ‘3p.ABS-3p.DAT + *-la*’, etc. (Mitxelena 1954a).
- 14 Attempts have been made to justify other details but not the *h-*, despite the fact that this is etymologically present (cf. Lakarra 2009b, 2015a, etc.). Moreover, an internal explanation exists (**her* ‘to close’ + *-i* ‘part.’, cf. Lakarra 2010, 2013c), and ILTIR could mean ‘river’, not ‘city’ (cf. De Hoz 2010–2011).
- 15 The influence of a lack of clear and persistent fragmentation in the process of the dialectalization of Basque remains to be studied (cf. IE and Austronesian languages, for example); see Section 9.
- 16 For example, the *-n-* > *-h-* change that Mitxelena understood as simply prior to the first medieval onomastical testimonies (9th–10th centuries) can be taken back at least six centuries (cf. Lakarra 2014) from the dating by Chambon and Greub (2002) of Proto-Gascon in the 5th century, which shares this and other features with its Aquitanian substrate, as has been acknowledged since Luchaire (cf. Gorrochategui 1984).
- 17 Martinet also addressed sibilants, although his argument, which was complex and yielded few results, has not had any followers. On his ideas about old accentuation, see Section 4.2.5.

- 18 Although Martinet and his successors attributed the disappearance of aspirated *fortes* to external influence, this is unnecessary and unlikely, hypercharacterization being sufficient to explain it (**th-*, **kh-* > *h* : **b^o*, **d^o*, **g^o* → *b-*, *d-*, *g-*) after centuries of the weak phoneme systematically adapting loanwords. The rare exceptions (*cauea* > *habia* ‘nest’, **kar* > *harri* ‘stone’) are very problematic; nor are the *g-* and *k-* in deictics in Navarrese speech forms (*gau*, *gori*, *gura* and *kau*, *kori*, *kura* ‘this’, ‘that’ (close), and ‘that’ (far) versus common *hau*, *hori*, *hura*) obviously archaic but, rather, much later innovations (cf. Lakarra 2014, 2017); See Egurtzegi (2018) for more arguments on behalf of [aspirated = fortes] vs [non-aspirated = lenes] plosives and Lakarra (2017) for some new cases for **T^h-* > *h-*.
- 19 This argument should now be revised: Hualde (1997a) proposed *f* < **wh* (see some of the etymologies in Lakarra 2009a) and, therefore, the chronology of these forms would be later, even subsequent to the change *au* > *ai* in Z and R: cf. Z *aihairi* ‘dinner’, *gaiherdi* ‘midnight’ versus general *afari*, *gauerdi*.
- 20 **Ba(d)in* + *-a* / **ba(d)in* + *-no* > *bai(n)a* ‘but, yet’ / *baino* ‘only, but’ and **arran* + *-i* / **arran* + *-no* > *arrain* ‘fish’ / *arrano* ‘eagle’. *-no* < **non* ‘to move’ as in comparisons in other languages (cf. Heine and Kuteva 2002).
- 21 Mitxelena mentioned it occasionally, but little importance has been given to another change (CrV- > CVr): cf. Rom. *trollo* (a little regarded fish, see Corominas-Pascual, s.u.) > **torllo* > **dorllo* > B *dollor* ‘bad, poor’.
- 22 The same thing occurs in older testimonies with occlusives. Such a tendency is much less complete in final position than in initial position, and in sibilants and vibrants than in laterals and nasals. See *FHV* and now Begiristain (2015) and Lakarra (2017).
- 23 See the new southern (peninsular) data on /h/ in demonstratives with a later chronology in Manterola (2015).
- 24 Not all intervocalic *h*’s are, however, etymologically prior. *Lur* ‘ground, soil, earth’ has a late and very minor variant *luur* in Biz (cf. *zoor* ‘debt’ in this same dialect). In the DGV, *luhur* appears as attested in the modern LNa dialect of Baigorri, but not in any other modern or old dialect with /h/. It is thus extremely unlikely that these are old forms (pace Blevins 2013).
- 25 Only this latter one is *h* < **h*. As Janhunen (2007) demonstrates, it is typical for /h/ to originate in different sources (“secondary laryngeals”), both in Uralic and in other languages. Lakarra (2015a) adds three other sources of /h/: (5) **-r* > *-h*, (6) /h/ in Gasconisms, and (7) *hVR-* > *VRh-* in Gasconisms and inherited words. There are, also, some *-b-*, *-d-*, *-g-* > *-h-* in Contemporary Low Navarrese (cf. Camino 2014) and *-r-* > *-h-* in Modern and Contemporary Zuberoan.
- 26 For the OPB and LPB accent, it should be remembered that the monosyllabic root (accentuated at the outset) was not initial but final with the important typological consequences that this implied (prefixes, not suffixes, etc.); cf. Section 6.2 ff.
- 27 In Hualde (1997b), one finds a synthesis of many of Hualde’s works on accentology of the modern dialects (particularly the western ones) that revolutionized studies on the subject. Hualde (2012) himself later continued with studies on accents in varieties such Goizueta (Navarre). Other recent essays on the history of the accent are Martínez Areta (2009), Elordieta (2011a-b), and Egurtzegi and Elordieta (2013).
- 28 Artiagoitia’s (1990) model is much more restrictive (CVC plus a extrametrical -C). Forms with CC-, -CC, and -VV- are common in phonosymbolisms (*brast* ‘abrupt start’, *dzaust* ‘dive’, etc.). Given that the canonical form of these usually (cf. IE, etc.) approaches the mirror image of lexemes, we have here additional proof of the CVC structure of the oldest B lexicon.
- 29 The expansion of derivation is a later phenomenon and belongs in great measure to literary B; see Lakarra (1997a), Sarasola (1997), etc.

- 30 Of great interest – although it corresponds to a much later stage – is the type of complex reduplication *ikusi-makusi* examined by Igartua (2013), whose distant origins could be in Turkish and surrounding languages and which would have been transmitted via Arabic and Romance.
- 31 See the results there corresponding to fossils and loanwords in the most common patterns.
- 32 Mitxelena pointed out that **d- > l*, prior to LPB, may explain the lack of *d-* in inherited terms; now (cf. Lakarra 2006b) we reconstruct multiple **dVC* roots: apart from the already known, **dun* ‘must, to have to’ and **din* ‘to become’, **don* ‘to put, to hang’, **dar* ‘to sit, to get’, **den* ‘to finish’, **dats* ‘to go down’, etc.
- 33 As Meillet and Mitxelena demonstrated on multiple occasions, the number of reconstructed phonemes is the minimum necessary to account for the morphemes in the protolanguage and its historical cognates. There could, of course, have been other phonemes that have not left sufficient traces or evidence of their existence. Of more interest are Igartua’s works on aspiration (/h/), beginning with Igartua (2002), in which he brilliantly related the change of aspiration and the root, or that of Igartua (2008, 2015) on rhinoglottophilia, completing the known etymological character of /h/ – derived from intervocalic /n/ – with numerous typological parallels.
- 34 The work of Morvan (2009) is a parody of proper etymological method: he is unaware of testimonies and previous philological efforts, as well as the potentialities of internal reconstruction, which he attempts to replace by a recourse to supposed Siberian, Dravidian, Uralo-Altaic, and Amerindian genetic relationships, on his whim (see Lakarra 2017).
- 35 Probably in a much-reduced area; see Janhunen (1982 and 2009) on Proto-Uralic.
- 36 To which one should add everything related to grammaticalization; see Section 7.
- 37 In Lakarra (2010), we arrive at similar conclusions in another way: analyzing the formation of numerals in B, in the same way as in many other languages, we observe that counting began on the index finger, ignoring the thumb and, therefore, the middle finger, that which is “uppermost,” was the second one; cf. Epps (2006) for Amazonian languages, de Lamberterie (2000) for IE languages, etc.
- 38 There are insufficiently researched or unknown others like *hor-tz* ‘fang’ (cf. Sp *cano*), *ipurdi* ‘bottom’ < **ibi-erdi* ‘central ford’, *laur* ‘four’ < *labur* ‘short’, *zur* ‘wood’ and its derivative *zuhur* ‘wise’ or *aretx* ‘oak/tree’ in Old Biz, with parallels in IE (cf. de Lamberterie 2000).
- 39 Blanco (2014) marks the beginning of lexicological analysis of the archaic corpus (up to 1600); this analysis of the lexicon in the old and classic eras (1600–1745) is extended by Blanco in his thesis (in progress).
- 40 See Dixon (2002) for the Australian languages and Austerlitz (1976) for the definition of “agglutinative” in the Uralic, Turkic, Mongolic, and other languages of Eurasia: these languages share common features such as suffixation, SOV order, vowel harmony, etc., as well as disyllabism of roots
- 41 Yet, as in English (AdjN), a more harmonic previous order (SOV in Old English, SVO/VSO in PB: Trask 1997 / Gómez 1994). Add to this the lack of an adjective open class, as in Tibeto-Burman (as opposed to the Dravidian agglutinative, see Bhat 2000). As pointed out in Section 5.5, there are reasons to think that the adjective was not an open class in Tibeto-Burman, but, on the contrary, in modern Tani – and in historical B –, adjectives do belong to an open class (cf. Post 2006).
- 42 The two determinate declensions are based on the grammaticalization of the article (after the 8th century for the singular, later for the plural); cf. Mitxelena (1971) and Manterola (2015).
- 43 Stassen, by the way, contends that in B there are no “WITH-language” structures and is therefore badly informed by his B sources: as a *fase sparita* it appears in Old

- Biz, A, G, and, at least, in Na and L oral ballads (see Lakarra 2008a). See Lakarra (in progress-b) for more consequences of the reconstruction of the COMITATIVE for PB morphology and syntax (sociative, modal adverbs, ‘abstracts objects’, dative flags, etc.); for parallels, see Lord (1973, 1993: West African languages) and Chapell et al. (2011: Chinese).
- 44 Note that the existence of disyllables and polysyllables in Aquitanian inscriptions does not demonstrate widespread root disyllabism; only later – after disyllabic inputs (not outputs) became the majority in word formation – could we speak of disyllabism as widespread CRS. Cf. Feng (1997) and Duanmu (1999) for monosyllabic to disyllabic change in the history of Chinese. See Lakarra (2018) for a preliminar analysis of some *-n / -r / -l / -h / -ø* alternations (and derived etymologies) in archaic CVC-word formation.
- 45 The oldest allomorph (present as *fase sparita* in adjectives like *hordi* ‘drunk’, *geldi* ‘still’, *handi* ‘big’, etc.) also concurs with the remote future/potential *-di* (*daidi*, *leidi*, etc.) as in other languages; cf. Heine and Kuteva (2002 s.u. COME) for parallels of all these grammaticalizations.
- 46 Combinations of applicative/directional prefixes + causative are polymorphous (*ar-*, *jar-*, *inar-*, *ihar-* or *ur-*, as well as *eroan/eraman* ‘to take’) as in Bantu but with the difference that in that family both suffixes and their combinations are not very old fossils as in B but instead are completely functional (cf. Good 2005, etc.).
- 47 Considered of unknown origin; it could come from a preposition similar to the English *to* and similar forms in other languages. Of identical origin seem to be the V- in the negation **eze*, the supposed epenthetic *-e-* in local cases of consonant declension or the conjunction *e-ta*, mentioned above; see Lakarra (in progress-b).
- 48 There are more than enough reasons for the destinative *-ra* to be an epenthetic *-r-* + *-a* ‘case marker’, in the same way as *-ri* (< **-r-i*) in DAT or *-ren* in GEN (< **-r-e-n*), i.e., such allomorphs are later reanalyses; see Section 7.5.
- 49 That is, it had prepositions and prefixes (both in verbs and in nouns, but not postpositions and suffixes) as in the historical period.
- 50 The *split ergativity* in plural of demonstratives shows (cf. Manterola 2015) that the distinction erg.pl./nom.pl. is secondary (Late Medieval), given that the article – the base of both the singular and the plural in declension – is also a later development among them (almost certainly after the 9th century).
- 51 Although this looks similar to Latin *salix* ‘willow’, it is not a loanword in Basque: 1) the sibilant in old Latin loans is *z-* (not the *-tz* of this case); 2) Latin-Romance words borrowed are nearly always borrowed in the accusative form, which would be *salice(m)* in this case; 3) *-ice* or *-icV* does not palatalize (nor change to a sibilant nor change the sound of the voiceless stop in words that pass into Basque); and 4) there already exists a completely regular loan in Basque from that Latin or Indo-European form, *zarika* ‘willow’. There are other arguments as well, but in short, Latin cannot be the source of this Basque word.
- 52 Although *e-ho(n)* ‘to grind, crush’ does exist, though from **e-non*, with **-n-* > *-h-*, subsequent to **T^h* > *h-*; this verb does not possess any synthetic forms in contrast to what happens with *ekharri* ‘to carry’, *ekhusi* ‘to see’, *ethorri* ‘to come’.
- 53 These last two chronological results clash directly with the “testimony” of the Iruña-Veleia inscriptions (dating from the 3rd–4th centuries according to their supposed “discoverers”. They are not, of course, the only (linguistic, epigraphic, or any other type) oddities present in what constitutes one of the greatest modern European hoaxes; see Gorrochategui (2011a and 2011b). Gorrochategui (2002) offers some interesting approaches to dating Basque.

- 54 The name of the language – from **enausi* ‘to speak’ (Irigoyen 1977), perhaps better **enotsi-hara* – has nothing to do with the old ethnonym *Auscii* as has been suggested. The nasal vowels in [êûskera] (< enusquera >, Esteban de Garibay, 16th century) confirm the etymological character of the *h-* in *heuskara* ‘B language’: cf. *Z harea*, *R âria* ‘sand’, *ainzto* ‘knife’, etc.
- 55 The connection between ‘job’ and ‘bread’ has to do with ‘the way to get bread’ extending in meaning to ‘to earn money’.
- 56 As is argued in Lakarra (2014), converting supposed “elections” like B *egin* ‘make’: Central **ezan* ‘can’: Eastern **iron* ‘can’ or B *eutsi* ‘to hold up’: Central **nin* ‘give’: Eastern **eradun* ‘to have for’, etc. into a series of innovations contributes to dating and to giving history to features that have usually been addressed as belonging to the timeless essences of the dialects. Thus in the first case, Biz shares the first three processes with the language as a whole, in the fourth with most of the forms of speech (all of them except Eastern ones), in the fifth with the Western ones as a whole, and in the sixth with some parts of A and G: it does not appear that, diachronically, such a choice confers it with a distinct personality, neither within the Western forms of speech nor within all of them as a whole.
- 57 See a fuller treatment of the methodological questions explored here in Lakarra (2014); we are far from having achieved the treatment that the numerous aspects and implications of the issue demand.
- 58 We can suppose a family of two elements (PB and Aquitanian; cf. Campbell 2011) through a series of similar considerations, including parallels in the languages like the Germanic languages (see now Stiles 2013); however, in reconstructive practice such an option has not been especially important; it seems preferable to assume that Aquitanian is the brother of OCB and Pyrenean B and not LPB, the source of all of them.
- 59 The lack of attention to the chronology of innovations is noticeable in the case of G, whose existence prior to Larramendi (18th c.) is debatable; nevertheless, it is typical to find it counterposed to Biz, as if their modern differences were *ab initio* and not much greater in recent centuries than in earlier ones.
- 60 There has been no research on dating the origin of G as a distinct dialect, but clearly it is one of the most recent ones, based on the defining features of Zuazo (1998: 217):
- | | |
|--|--|
| 1 Instability of the organic <i>-a</i> ; | 5 Change <i>f</i> > <i>p</i> ; |
| 2 Root <i>-e-</i> in pres. of * <i>edun</i> ‘(to) have’; | 6 <i>nor</i> ‘who’/ <i>zein</i> ‘which’ > <i>zein</i> ‘who, which’; |
| 3 Root <i>-e-</i> in present of <i>izan</i> ‘to be’; | 7 Conjugated forms as <i>nijoa</i> ‘I’m going’ from <i>joan</i> ‘to go’; |
| 4 Change <i>d</i> > <i>r</i> [intervocalic]; | 8 Interrogative particle <i>al</i> . |
- G shares with A (2), (3), (4), and (5), and (8) does not appear until 1785; (1) also appears to have spread in the 18th–19th centuries, the point of becoming a differentiating and marking characteristic with respect to the other Southern dialects. However, the innovations shared by G and A and with B are much older.
- 61 The bipartite classifications of Bonaparte and Lacombe (synchronic) and Uhlenbeck (linked to supposed polygenesis) – both similar to our left tree – have little to do with the conduct of diachronic dialectology and linguistics; see Mitxelena (1964), Lakarra (1996 and 2014), among others.
- 62 Lately, Camino (2011, 2014, etc.) has also maintained that the first split involves eastern forms of speech and has offered some indications of proof in this regard.
- 63 This hypothetical model of fragmentation of OCB could be compatible with ongoing historical work (cf. Pozo 2016) that contends that, in the 5th century, an important

political entity emerged between Pamplona and the Pyrenees, directly related to the later Kingdom of Pamplona.

- 64 Although for reasons of space I have only referred tangentially here to the philological part, the importance of its development for advances in B diachrony is essential. Besides the monumental Mitxelena and Sarasola (1987–2005) and Lakarra, Manterola and Seguro (2017), see among others Mitxelena (1958), Gorrochategui (1984), Lakarra (1997a), Mounole and Lakarra (2017), Ulibarri (2013), and Urgell (2013).

REFERENCES

- Agud, Manuel and Antonio Tovar. 1988–1995. *Materiales para un diccionario etimológico de la lengua vasca (A-Orloi)*. Donostia-San Sebastián: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology*, 7 vols.
- Aikhenvald, Alexandra Y. 2006. Serial Verb Constructions in Typological Perspective. *Serial Verb Constructions. A Cross-Linguistic Typology*, ed. by Alexandra Y. Aikhenvald and Robert M. W. Dixon, 1–68. Oxford: Oxford University Press.
- Arbelaiz, Juan José 1978. *Las etimologías vascas en la obra de Luis Michelena*. Tolosa: Kardaberatz.
- Ariztimuño, Borja. 2013. Finite verbal morphology. In Martínez Areta, 359–427.
- Artiagoitia, Xabier. 1990. Sobre la estructura de la sílaba en (proto)vasco y algunos fenómenos conexos. *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 24, no. 2: 327–349.
- Austerlitz, Robert. 1976 [1970]. L'aglutination dans les langues de l'Eurasie septentrionale. *Études Finno-ougriennes* 13: 7–12.
- Azkue, Resurrección María de. 1923–1925. *Morfología vasca*. Reed. Bilbao: La Gran Enciclopedia Vasca.
- Bakró-Nagy, Marianne Sz., 1992, *Proto-Phonotactics. Phonotactic investigation of the Proto-Uralic and Proto-Finno-Ugric consonant system*. Wiesbaden: Studia Uralica 5, Harrassowitz Verlag.
- Barbero, Abilio and Marcelo Vigil. 1965. Sobre los orígenes sociales de la Reconquista: cántabros y vascones desde fines del Imperio Romano hasta la invasión musulmana. *Sobre los orígenes sociales de la Reconquista*, ed. by Abilio Barbero and Marcelo Vigil, 13–103. Ariel: Barcelona 1974.
- Barreña, Andoni, Ane Ortega, and Estibalitz Amorrortu. 2013. The Basque Language Today: Achievements and Challenges. In Martínez Areta, 11–29.
- Begiristain, Alazne. 2015. Gogoetak Mitxelenaren errobo sistemaz eta honen historiaurreaz. Ikuspegi berri baterantz. BA Thesis, UPV/EHU.
- Begiristain, Alazne. In progress. *FHV-ko kontsonante-taldeekiko atalaz*. MA Thesis, UPV/EHU.
- Benveniste, Émile. 1935. *Origines de la formation des noms en indo-européen*. Paris: Maisonneuve.
- Benveniste, Émile. 1954. Problèmes sémantiques de la reconstruction. Reed. *Problèmes de linguistique générale*, ed. by Émile Benveniste, 289–307. Paris: Gallimard, 1966.
- Bhat, D. N. Shankara. 2000. Dravidian and Tibeto-Burman: A Typological Comparison. *International Journal of Dravidian Linguistics* 29: 9–40.
- Blanco, Endika. 2014. *Euskara Arkaikoaren Lexikoiaz*. Master Thesis, UPV/EHU.

- Blevins, Juliette. 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, Juliette. 2013. Advances in Proto-Basque reconstruction. *International Congress of Historical Linguistics* (Oslo).
- Blust, Robert. 1988. *Austronesian Root Theory: Essay on the Limits of Morphology*. Amsterdam: Benjamins.
- Blust, Robert. 2014. Recent Proposals Concerning the Classification of the Austronesian Languages. *Oceanic Linguistic* 53: 300–391.
- Brandstetter, Renward. 1916. *An Introduction to Indonesian Linguistics*. London: The Royal Asiatic Society.
- Calvet, Jean-Louis. 1981. *Lingüística y colonialismo*. Madrid: Ed. Jucar.
- Camino, Iñaki. 2011. Pirinioak, ekialdea eta euskal mintzoak. In Sagarna et al., 773–822.
- Camino, Iñaki. 2014. Ekialdeko euskararen iraganaz. In Epelde, 87–153.
- Campbell, Lyle. 1988. (Art.-Review) *Language in the Americas* by Joseph H. Greenberg. *Language* 64: 591–615.
- Campbell, Lyle. (ed.). 2001. *Grammaticalization: a critical assessment*. (Special issue of *Language Sciences*, vol. 23, numbers 2–3.).
- Campbell, Lyle. 2011. La investigación histórica de las lenguas aisladas, o ¿es raro el vasco? In Lakarra, Gorrochategui and Urgell, 23–40.
- Campbell, Lyle. 2013 [1998], *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Chambon, Jean-Pierre and Yan Greub. 2002. Note sur l'âge du (proto)gascon. *Revue de Linguistique Romane* 66: 473–495.
- Corominas, Joan and José Antonio Pascual. 1980–1991. *Diccionario crítico etimológico castellano e hispánico*, 6 vols. Madrid: Gredos.
- De Hoz, Javier. 2011. *Historia lingüística de la Península Ibérica de la Península Ibérica en la Antigüedad. II. El mundo ibérico prerromano y la indoeuropeización*. Madrid: CSIC.
- de Rijk, Rudolf P. G. 1981. Euskal morfologiaren zenbait gorabehera. In de Rijk 1998, 211–224.
- de Rijk, Rudolf P. G. 1992. 'Nunc' Vasconice. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 26, no. 3, 695–724. In de Rijk 1998, 347–376.
- de Rijk, Rudolf P. G. 1998. *De lingua uasconum: Selected Writings*. Donostia-San Sebastián: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology*.
- Dimmendaal, Gerrit J. 2011. *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam and Philadelphia: John Benjamins.
- Dixon, Robert M. W. 1977. Where Have All Adjectives Gone. *Where Have All the Adjectives Gone and Other Essays in Semantics and Syntax*, ed. by Robert M. W. Dixon, 1–62. Berlin: Mouton 1982.
- Dixon, Robert M. W. 2002. *Australian Languages*. Cambridge: Cambridge University Press.
- Donegan, Patricia. 1993. Rhythm and Vocalic Drift in Munda and Mon-Khmer. *Linguistics in the Tibeto-Burman Area* 16: 1–43.
- Donegan, Patricia and David Stampe. 1983. Rhythm and the Holistic Organization of Language Structure. *Papers from the Parasession of Phonology, Morphology and Syntax*, ed. by John F. Richardson et al., 337–353. Chicago: Chicago Linguistic Society.

- Donegan, Patricia and David Stampe. 2004. Rhythm and the Synthetic Drift of Munda. *The Yearbook of South Asian Languages and Linguistics 2004*, ed. by Rajendra Singh, 3–36. Berlin and New York: de Gruyter.
- Duanmu, Sam. 1999. Stress and the Development of Disyllabic Words in Chinese. *Diachronica* 16: 1–36.
- Echenique, Maria Teresa. 1983. *Historia lingüística vasco-románica* (2nd ed.). Madrid: Paraninfo, 1987.
- Egurtzegi, Ander. 2014. *Towards a Phonetically Grounded Diachronic Phonology of Basque*. PhD, UPV/EHU.
- Egurtzegi, Ander. 2018. *Herskarien ustezko ahoskabetasun asimilazioa eta euskal herskari zaharren gauzatzea*. Forthcoming in a festschrift edited by Joseba Lakarra & Blanca Urgell.
- Egurtzegi, Ander and Gorka Elordieta. 2013. Euskal azentueren historiaz. In Gómez et al., 163–186.
- Elmendorf, William. 1997. A Preliminary Analysis of Yukian Root Structure. *Anthropological Linguistics* 39: 74–91.
- Elordieta, Gorka. 2011a. Euskal azentuaren bilakaera: hipotesiak eta proposamenak. In Sagarna et al., 989–1014.
- Elordieta, Gorka. 2011b. Euskal azentu eta intonazioari buruzko ikerketa: status quaestionis. In Lakarra et al., 389–428.
- Epelde, Irantzu. (ed.). 2014. *Euskal dialektologia: lehen a eta oraina*. Bilbo: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology*, no. 69.
- Ettxepare, Bernat. 1545. *Linguae Vasconum Primitiae*. Edition by Patxi Altuna, Bilbao: Euskaltzaindia.
- Epps, Patience. 2006. Growing a Numeral System: The Historical Development of Numerals in a Amazonian Language Family. *Diachronica* 23: 259–288.
- Feng, Shengli. 1997. Prosodic Structure and Compound Words in Classical Chinese. *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 197–260. Berlin: Mouton de Gruyter.
- Fischer, Olga, Muriel Norde, and Harry Perridon (eds.). 2004. *Up and Down the Cline – The Nature of Grammaticalization*. Amsterdam and Philadelphia: John Benjamins.
- Forni, Gianfranco. 2013. Evidence for Basque as a IE Language. *Journal of Indo-European Studies* 41, nos. 1–2: 39–180.
- García Uriz, Eneko. 2016. Hitz hasierako herskari ahoskabeak euskararen historian. BA Thesis, UPV/EHU.
- Gavel, Henri. 1920. *Éléments de phonétique basque*. Paris: *Revue Internationale des Études Basques* 12.
- Gómez, Ricardo. 1994. Euskal aditz morfologia eta hitzordena: VSO-tik SOV-ra. *La langue basque parmi les autres*, ed. by Jean-Baptiste Orpustan, 93–114. Baigorri: Izpegi.
- Gómez, Ricardo and Koldo Sainz. 1995. On the Origin of the Finite Forms of the Basque Verb. In Hualde, Lakarra and Trask, 235–274.
- Gómez, Ricardo, Joaquín Gorrochategui, Joseba Andoni Lakarra, and Céline Mounole (eds.). 2013. *Koldo Mitxelena Katedraren III. Biltzarra (Gasteiz, 8–11/X/2012)*. Vitoria-Gasteiz: UPV/EHU.
- Good, Jeffrey. 2005. Reconstructing Morpheme Order in Bantu: The Case of Causativization and Applicativization. *Diachronica* 22: 3–57.

- Gorrochategui, Joaquín. 1984. *Estudio sobre la onomástica indígena de Aquitania*. Bilbao: UPV/EHU.
- Gorrochategui, Joaquín. 1985. Historia de las ideas acerca de los límites geográficos del vasco antiguo. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 19: 571–594.
- Gorrochategui, Joaquín. 1987. Vasco-céltica. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 21: 951–959.
- Gorrochategui, Joaquín. 1998. *Euskararen historiaurreaz zenbait gogoeta. Algunas reflexiones sobre la prehistoria de la lengua vasca*. Vitoria-Gasteiz: UPV/EHU.
- Gorrochategui, Joaquín. 2002. Planteamientos de la lingüística histórica en la datación del euskara. *XV Congreso de Estudios Vascos*, 103–114. Donostia-San Sebastián: SEV.
- Gorrochategui, Joaquín. 2011a. Las armas de la filología. In Lakarra, Gorrochategui and Urgell, 41–70.
- Gorrochategui, Joaquín. 2011b. *Hic et nunc*. Falsificaciones contemporáneas. El caso de Iruña-Veleia. *El monumento epigráfico en fuentes secundarias. Procesos de reutilización, interpretación y falsificación*, ed. by Jordi Carbonell, Helena Gimeno, and José Luis Moralejo, 241–261. Barcelona: Universitat Autònoma.
- Gorrochategui, Joaquín. 2017. “El vasco en la Antigüedad.” In Gorrochategui, Igartua & Lakarra (eds.). Forthcoming.
- Gorrochategui, Joaquín, Iván Igartua, and Joseba Andoni Lakarra (eds.). 2017. *Historia de la lengua vasca*. Vitoria-Gasteiz: Gobierno Vasco. Forthcoming.
- Gorrochategui, Joaquín and Joseba Andoni Lakarra. 2013. Why Basque Is Not, Unfortunately, an IE Language? *Journal of Indo-European Studies* 41, nos. 1–2: 203–237.
- Greenberg, Joseph. 1963 [1990]. Some Universals of Grammar With Particular Reference to the Order of Meaningful Elements. Reed. *On language: Selected Writings of Joseph H. Greenberg*, ed. by Keith Denning and Suzanne Kemmer, 40–70. Stanford, CA: Stanford University Press.
- Güter, Henri. 1989. Elementos de cronología fonética del vascuence. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 23, no. 3: 797–800.
- Haas, Mary. 1969. *The Prehistory of Languages*. The Hague: Mouton.
- Heine, Bernhard and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge, NY: Cambridge University Press.
- Holmer, Nils M. 1970. A Historic-Comparative Analysis of the Structure of the Basque Language: The Principal Linguistic Types. *Fontes Linguae Vasconum. Studia et Documenta* 2: 5–47.
- Houis, Maurice. 1970. Reflexion sur une double correlation typologique. *Journal of West African Languages* 7: 59–68.
- Hualde, José Ignacio. 1991. *Basque Phonology*. London and New York: Routledge.
- Hualde, José Ignacio. 1995. Reconstructing the Ancient Basque Accentual System: Evidence and Hypotheses. In Hualde, Lakarra and Trask, 171–188.
- Hualde, José Ignacio. 1997a. Aitzineuskararen leherkariak. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 31: 411–424.
- Hualde, José Ignacio. 1997b. *Euskararen azentuerak*. Donostia-San Sebastián: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology*.
- Hualde, José Ignacio. 2006. Laringalak eta euskal azentuaren bilakaera. *Andolin Eguzkitza gogoan*. In honor to Professor Eguzkitza, ed. by Beatriz Fernández and Itziar Laka, 497–511. Bilbao: UPV/EHU.

- Hualde, José Ignacio. 2008. Acentuación y cronología relativa en la lengua vasca. *Oihernart* 23: 199–217.
- Hualde, José Ignacio. 2012. Two Basque Accentual Systems and the Notion of Pitch-Accent Language. *Lingua* 122: 1335–1351.
- Hualde, José Ignacio. (forthcoming). Dialektologia dinamikoa. In a volume of *Lapurdum* edited by Irantzu Epelde.
- Hualde, José Ignacio, Joseba Andoni Lakarra, and Larry Trask (eds.). 1995. *Towards a History of Basque Language*. Amsterdam and Philadelphia: John Benjamins.
- Hualde, José Ignacio and Jon Ortiz de Urbina (eds.). 2003. *A Grammar of Basque*. Berlin: Mouton de Gruyter.
- Igartua, Iván. 2002. Euskararen hasperena ikuspegi tipologiko eta diakronikotik. *Erramu Boneta: Festschrift for Rudolf P.G. de Rijk*, Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* no. 44, ed. by Xabier Artiagoitia, Patxi Goenaga, and Joseba Andoni Lakarra, 366–389. Bilbao: UPV/EHU.
- Igartua, Iván. 2008. La aspiración de origen nasal en la evolución fonológica del euskera: un caso de rhinoglottophilia. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 42: 171–189.
- Igartua, Iván. 2013. La reduplicación compleja en euskera: notas acerca de su formación y sus paralelos en otras lenguas. *Fontes Linguae Vasconum. Studia et Documenta* 45: 5–30.
- Igartua, Iván. 2015. Diachronic Effects of Rhinoglottophilia, Symmetries in Sound Change, and the Curious Case of Basque. *Studies in Language* 39, no. 3: 635–663.
- Irigoyen, Alfonso. 1977. Geure hizkuntzari euskaldunok deritzagun izenaz. *Euskera* 22, no. 2: 513–538.
- Janhunen, Juha. 1982. On the structure of Proto-Uralic. *Finnisch-ugrische Forschungen* 44: 23–42.
- Janhunen, Juha. 1997. Problems of Primary Root Structure in Pre-Proto-Japonic. *International Journal of Central Asian Studies* 2, (ed. in chief Choi Hab-Woo). The International Association of Central Asian Studies Institute of Asian Culture and Development.
- Janhunen, Juha. 2007. The Primary Laryngal in Uralic and Beyond. *Sámit, sánit, sátnehámit. Riepmočála Pekka Sammallahtii miessemánu 21. beaivve 2007*, 253, 203–227. Helsinki: Suomalais-Ugrilaisen Seuran Toimituksia = Mémoires de la Société Finno-Ougrienne.
- Janhunen, Juha. 2009. Proto-Uralic: What, Where, and When? *The Quasquicentennial of the Finno-Ugrian Society*, ed. by Jussi Ylikoski, 57–78. Helsinki: *Mémoires de la Société Finno-Ougrienne* 258.
- Krajewska, Dorota. In progress. **Historical syntax of relatives in Basque*. PhD, UPV/EHU.
- Lafitte, Pierre. 1944. *Grammaire basque (navarro-labourdin littéraire)*. Facsimil, Donostia-San Sebastián: Elkar 1979.
- Lafon, René. 1943. *Le système du verbe basque au XVIème siècle*. Donostia-San Sebastián: Elkar 1980.
- Lafon, René. 1948. Sur les suffixes casuels *-ti/-tik* en basque. In Lafon 1999, 199–207.
- Lafon, René. 1999. *Vasconiana*. Bilbao: Iker 11, Euskaltzaindia.
- Lakarra, Joseba Andoni. 1995. Reconstructing the Root in Pre-Proto-Basque. In Hualde et al., 189–206.
- Lakarra, Joseba Andoni. 1996. *Refranes y Sentencias: ikerketak eta edizioa*. Bilbao: Euskaltzaindia.
- Lakarra, Joseba Andoni. 1997a. Euskararen historia eta filologia: arazo zahar, bide berri. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 31: 447–535.

- Lakarra, Joseba Andoni. 1997b. Gogoetak aitzineuskararen berreraiketaz: konparaketa eta barneberreraiketa. *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 31: 537–616.
- Lakarra, Joseba Andoni. 2005. Prolegómenos a la reconstrucción de segundo grado y al análisis del cambio tipológico en (proto)vasco. *Palaeohispanica* 5: 407–470.
- Lakarra, Joseba Andoni. 2006a. Protovasco, munda y otros: reconstrucción interna y tipología holística diacrónica. *Oihenart* 21: 229–322.
- Lakarra, Joseba Andoni. 2006b. Notas sobre iniciales, cambio tipológico y prehistoria del verbo vasco. In Lakarra and Hualde, 561–621.
- Lakarra, Joseba Andoni. 2008a. Aitzineuskararen gramatizarantz malkar eta osinetan zehar. *Gramatika Jaietan. Patxi Goenagari Omenaldia*, ed. by Xabier Artiagoitia and Joseba Andoni Lakarra, 451–490. Bilbao: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* no. 51, UPV/EHU.
- Lakarra, Joseba Andoni. 2008b. *Vida con/y libertad*: sobre una coordinación arcaica y la autenticidad de “Urthubiako Alhaba.” *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 42: 83–100.
- Lakarra, Joseba Andoni. 2008c. Irregularidades radicales y antiguas extensiones a la izquierda: Para la reconstrucción del verbo vasco. Ms. UPV/EHU.
- Lakarra, Joseba Andoni. 2008d. Forma canónica, etimología y reconstrucción en el campo vasco. *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 37: 261–391.
- Lakarra, Joseba Andoni. 2009a. Forma canónica y cambios en la forma canónica en la prehistoria de la lengua vasca. *X. Congreso de Lenguas y Culturas Paleohispánicas, Palaeohispanica* 9: 557–609.
- Lakarra, Joseba Andoni. 2009b. Adabakiak /h/-aren balio etimologikoaz. *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 43, nos. 1–2: 565–596.
- Lakarra, Joseba Andoni. 2010. Haches, diptongos y otros detalles de alguna importancia: notas sobre numerales (proto)vascos y comparación vasco-ibérica (Apéndice sobre *hiri* y *bat-bi*). *Veleia* 27: 191–238.
- Lakarra, Joseba Andoni. 2011a. Erro monosilabikoaren teoria eta aitzineuskararen berreraiketa: zenbait alderdi eta ondorio, *Fontes Linguae Vasconum. Studia et Documenta* 113, 5–114.
- Lakarra, Joseba Andoni. 2011b. Aitzineuskara: egindakoak eta eginkizunak. In Andoni Sagarna, Joseba Andoni Lakarra, and Patxi Salaberri (eds.), *Pirineoetako hizkuntzak: lehena eta oraina*, Iker 26, Bilbao: 617–694.
- Lakarra, Joseba Andoni. 2012. Mailegaketa eta berreraiketa euskararen historiaurrearen ikerketan. In Igartua, 17–74.
- Lakarra, Joseba Andoni. 2013a. Monosyllabic root theory and the reconstruction of Proto-Basque: some aspects and consequences. In Martínez Areta, 173–221.
- Lakarra, Joseba Andoni. 2013b. Aitzineuskararen berreraiketa sakonagorantz: forma kanonikoa, tipologia holistikoa, kronologia eta gramatikalizazioa. In Gómez et al., 274–324.
- Lakarra, Joseba Andoni. 2013c. Gramática histórica vasca o vasco-iberismo. *Palaeohispanica* 13: 567–592.
- Lakarra, Joseba Andoni. 2014. Gogoetak euskal dialektologia diakronikoaz: Euskara Batu Zaharra berreraiki beharraz eta haren banaketaren ikerketaz. In Epelde, 155–241.

- Lakarra, Joseba Andoni. 2015a. Hiru hasperen haboro. *Eridenen du zerzaz kontenta. Sailkideen omenaldia Henrike Knörr irakasleari*, ed. by Ricardo Gómez and Maria Jose Ezeizabarrena, 349–378. Bilbao: UPV/EHU.
- Lakarra, Joseba Andoni. 2015b. *Saratsola* eta (aitzin)eusk(ar)en geruzak. In Beatriz Fernández and Pello Salaburu (eds.), *Ibon Sarasola. Gorazarre. Homenatge. Home-naje*, Donostia-San Sebastián: UPV/EHU, 419–439.
- Lakarra, Joseba Andoni. 2015c. *Bi* eta bere askazia. Forthcoming in a festschrift.
- Lakarra, Joseba Andoni. 2016. Gramatikalizazioa, morfemen forma kanonikoak eta berreraiketa morfologikoaren bide berriak. *Txipi Ormaetxea omenduz. Hire bordatxoan*, ed. by Gotzon Aurrekoetxea, Jesus Mari Makazaga, and Patxi Salaberri, 175–192. Leioa UPV/EHU.
- Lakarra, Joseba Andoni. 2017. Prehistoria de la lengua vasca. In Gorrochategui, Igartua and Lakarra. Forthcoming.
- Lakarra, Joseba Andoni. 2018. CVC berreraikiaz: aurretik eta atzetik. Forthcoming in a festschrift ed. by Joseba A. Lakarra and Blanca Urgell.
- Lakarra, Joseba Andoni. In progress-a. Prehistoria del comitativo e implicaciones para la reconstrucción de la morfología y la sintaxis protovasca. Ms. UPV/EHU.
- Lakarra, Joseba Andoni. In progress-b. Protogenitivo, protolocativo y sintaxis del protovasco. Ms. UPV/EHU.
- Lakarra, Joseba Andoni, Joaquín Gorrochategui, and Blanca Urgell (eds.). 2011. *II Congreso de la Cátedra Luis Michelena*. Bilbao: UPV/EHU.
- Lakarra, Joseba Andoni and José Ignacio Hualde (eds.). 2006. *Studies in Basque and Historical Linguistics in Memory of Robert Larry Trask (=Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology XL, 1–2)*. Bilbao: UPV/EHU.
- Lakarra, Joseba Andoni, Julen Manterola, and Iñaki Segurola. 2016. Los estudios etimológicos vascos: historia y perspectivas. *Etimología e historia en el léxico del español. Estudios ofrecidos a José Antonio Pascual (Magister bonus et sapiens)*, ed. by Mariano Quirós, 843–869. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Lakarra, Joseba Andoni, Julen Manterola, and Iñaki Segurola. 2017. *Euskal hiztegi historiko etimologikoa*. Bilbao: Euskaltzaindia.
- Lamberterie, Charles de. 2000. Problèmes sémantiques de la reconstruction en indo-européen. *Théories contemporaines du changement sémantique*, ed. by Jacques François, 109–134. Leuven: Peeters.
- Larramendi, Manuel de. 1729. *El imposible vencido. Arte de la lengua bascongada*. Salamanca. Facsimil Hordago, Donostia-San Sebastián, 1979.
- Larramendi, Manuel de. 1745. *Diccionario Trilingüe del castellano, bascuence y latín*, Donostia-San Sebastián. Facsimil Txertoa, Donostia-San Sebastián, 1984.
- Lord, Carol. 1973. Serial verbs in transitions. *Studies in African Linguistics* 4, no. 3, 269–296.
- Lord, Carol. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam: John Benjamins.
- Madariaga, Juan. 2008. *Apologistas y detractores de la lengua vasca*. Donostia-San Sebastián: Fundación para el estudio del Derecho Histórico y Autonómico de Vasconia.
- Manterola, Julen. 2015. *Euskararen morfologia historikorako. Artikuluak eta erakusleak. Towards a history of Basque morphology: articles and demonstratives*. PhD, UPV/EHU.
- Martinet, André. 1950. De la sonorisation des occlusives initiales en basque. *Word* 6: 224–233.
- Martinez Areta, Mikel. 2006. *El consonantismo protovasco*. PhD, UPV/EHU.

- Martinez Areta, Mikel. 2009. El acento protovasco. *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 38: 135–206.
- Martinez Areta, Mikel. (ed.). 2013. *Basque and Proto-Basque*. Frankfurt: Peter Lang.
- Meillet, Antoine. 1925. *La méthode comparative en linguistique historique*, Paris. Reed. Klincksieck 1970.
- Michelena, Luis = Koldo Mitxelena.
- Mitxelena, Koldo. 1950. De etimología vasca. In Mitxelena 1988, 439–444.
- Mitxelena, Koldo. 1951. La sonorización de las oclusivas iniciales. In Mitxelena 1988, 203–211.
- Mitxelena, Koldo. 1954a. Nota sobre algunos pasajes de los *Refranes y Sentencias* de 1596. In Mitxelena 1988, 792–798.
- Mitxelena, Koldo. 1954b. De onomastica aquitana. In Mitxelena 1987, 409–445.
- Mitxelena, Koldo. 1956. La lengua vasca como medio de conocimiento histórico, *Zumarraga* 6, 49–70.
- Mitxelena, Koldo. 1957a. Las antiguas consonantes vascas. Reed. in Mitxelena 1988, 166–189.
- Mitxelena, Koldo. 1957b. Basque et roman. Reed. in Mitxelena 1988, 106–115.
- Mitxelena, Koldo. 1958. Introducción [a Landucci (1562)]. In Mitxelena 1988, II, 762–782.
- Mitxelena, Koldo. 1961/77. *Fonética histórica vasca* (2nd revised ed.). Donostia-San Sebastián: Supplement of the *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* no. 4.
- Mitxelena, Koldo. 1963. *Lenguas y protolenguas*. In Mitxelena, Donostia-San Sebastián: Supplement of the *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* no. 20, 1990.
- Mitxelena, Koldo. 1964. Sobre el pasado de la lengua vasca. In Mitxelena 1988, 1–73.
- Mitxelena, Koldo. 1971. Toponimia, léxico y gramática. In Mitxelena 1987, 141–167.
- Mitxelena, Koldo. 1973. *Apellidos vascos* (3rd ed.). Donostia-San Sebastián: Txertoa.
- Mitxelena, Koldo. 1974. El elemento latino-románico en la lengua vasca. In Mitxelena 1987, 195–219.
- Mitxelena, Koldo. 1977. Notas sobre compuestos verbales vascos. In Mitxelena 1987, 311–335.
- Mitxelena, Koldo. 1979. La langue ibère. Reed. in Mitxelena 1985, 341–356.
- Mitxelena, Koldo. 1981. Lengua común y dialectos vascos. In Mitxelena 1987, 35–55.
- Mitxelena, Koldo. 1987. *Palabras y Textos*, ed. Joaquín Gorrochategui. Bilbao: UPV/EHU.
- Mitxelena, Koldo. 1988. *Sobre historia de la lengua vasca*, ed. Joseba Andoni Lakarra. Donostia-San Sebastián: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology* 10, 2 vols.
- Mitxelena, Koldo. 2011–2012. *Luis Michelena. Obras Completas*, ed. Joseba Andoni Lakarra and Iñigo Ruiz Arzalluz, Bilbao/Donostia-San Sebastián: Supplements of the *Anuario del Seminario de Filología Vasca Julio de Urquijo - International Journal of Basque Linguistics and Philology*, 15 vols.
- Mitxelena, Koldo and Ibon Sarasola. 1987–2005. *Diccionario general vasco. Orotariko Euskal Hiztegia*. Bilbao: Euskaltzaindia.
- Morvan, Michel. 2009. *Dictionnaire étymologique basque-français-espagnol*. www.lexilogos.com/basque_dictionnaire.htm.

- Mounole, Céline. 2006. Quelques remarques à propos de l'histoire des périphrases basques. In Lakarra and Hualde, 723–738.
- Mounole, Céline. 2011. *Le verbe basque ancien: étude philologique et diachronique*. PhD UPV/EHU. Forthcoming in *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology*.
- Mounole, Céline. 2015. Lazarragaren gramatika. Ms., UPV/EHU.
- Mounole, Céline and Joseba Andoni Lakarra. 2017. Euskara Arkaikoa. In Gorrochategui, Igartua and Lakarra. Forthcoming.
- Odriozola, Aitor. 2016. Hitz hasierako bokalismoa: e- hitzen berreraiketa eta bilakabide diakronikoa. BA Thesis, UPV/EHU.
- Olmo Lete, Gregorio del. 2003. *Questions de linguistique sémitique. Racine et lexème. Histoire de la recherche (1940–2000)*. Paris: Jean Maisonneuve, “Librairie d’Amérique et d’Orient”.
- Pawley, Andrew. 2006. Where Have All the Verbs Gone? Remarks on the Organisation of Languages with Small, Closed Verb Classes. *11th Biennial Rice University Linguistic Symposium, 16–18 March*. Unpublished.
- Post, Mark W. 2006. Compounding and the Structure of the Tani Lexicon. *Linguistics in the Tibeto-Burman Area* 29: 40–60.
- Post, Mark W. 2009. Prosodic and Typological Drift in Austroasiatic and Tibeto-Burman: Against “Sinosphere” and “Indosphere.” Ms., The Cairns Institute, James Cook University.
- Pozo, Mikel. 2016. *Vasconia y los vascones, de la crisis del Imperio Romano a la llegada del Islam (siglos V–VIII)*. PhD Thesis, UPV/EHU.
- Rico, Francisco. 1982. *Primera cuarentena y tratado general de literatura*. Barcelona: El festín de Esopo.
- Ringe, David A. 2003. Internal reconstruction. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 244–261. Oxford: Blackwell.
- Rohlf's, Gerhard. 1977. *Le gascon. Études de philologie pyrénéenne* (3rd ed.). Niemeyer: Tubinga.
- Sagarna, Andoni, Joseba Andoni Lakarra and Patxi Salaberri (eds.). 2011. *Pirineoetako hizkuntzak: lehena eta oraina* (= 16. Biltzarra. Iruñea 6–10/X/2008). Bilbo: Iker 26, Euskaltzaindia.
- Sagart, Laurent. 1999. *The Roots of Old Chinese*. Amsterdam and Philadelphia: John Benjamins.
- Salaberri, Patxi. 1991. Toponimia, dialektologiaren ikerbide. *Nazioarteko dialektologia Biltzarra*, 619–645. Bilbao: Euskaltzaindia.
- Salaberri, Patxi. 2011a. Pirinioetako euskal toponimoak: direnak eta diratekeenak. *Pirineoetako hizkuntzak: lehena eta oraina* (= 16. Biltzarra. Iruñea 6–10/X/2008), ed. by Andoni Sagarna, Joseba Andoni Lakarra, and Patxi Salaberri, 977–1005. Bilbo: Iker 26, Euskaltzaindia.
- Salaberri, Patxi. 2011b. El sufijo occidental *-ika* y otras cuestiones de toponimia vasca. *Fontes Linguae Vasconum. Studia et Documenta* 43: 139–176.
- Salaberri, Patxi. 2013. *Apellidos Vascos* eta Mitxelenaren onomastika lanak. In Gómez et al., 673–697.
- Sarasola, Ibon. 1997. Euskal hitz altxorraz. *Anuario del Seminario de Filología Vasca Julio de Urquijo – International Journal of Basque Linguistics and Philology* 31: 617–642.
- Schuh, Russell. 2013. Word Families in Hausa. *Language, Literature and Culture in a Multilingual Society. A festschrift for Abubakar Rasheed*, ed. by Ozo-Mekuri Ndimele,

- M. Ahmad, and H.M. Yakasai, 579–598. Port Harcourt: M. and J. Grand Orbit Communications.
- Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4: 1–54.
- Stiles, Patrick V. 2013. The Pan-West Germanic Isoglosses and the Sub-Relationships of West Germanic to Other Branches. *NOWELE* 66: 1, 5–38.
- Straka, Georges. 1951–1954. Observations sur la chronologie et les dates de quelques modifications phonétiques en roman et en français pré-littéraires. *Revue des Langues Romanes* 71: 247–307.
- Thomason, Sandra G. 1993. Copying with Partial Information in Historical Linguistics. *Historical Linguistics 1989: Papers from the Ninth International Conference on Historical Linguistics*, ed. by Henk Aertsen and Robert J. Jeffers, 485–496. Amsterdam: John Benjamins.
- Thomason, Sandra G. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.
- Tovar, Antonio. 1981. *Mitología e ideología sobre la lengua vasca*. Madrid: Alianza Editorial.
- Trask, Robert Larry. 1977. Historical Syntax and Basque Verbal Morphology. *Anglo-American Contributions to Basque Studies: In Honor of Jon Bilbao*, ed. by William A. Douglass, Richard W. Etulain, and William H. Jacobsen Jr., 203–217. Reno: University of Nevada.
- Trask, Robert Larry. 1996. *Historical Linguistics*. London and New York: Arnold.
- Trask, Robert Larry. 1997. *The History of Basque*. London: Routledge.
- Trask, Robert Larry. 1998. The Typological Position of Basque: Then and Now. *Language Sciences* 20: 313–324.
- Trask, Robert Larry. 2008. *Etymological Dictionary of Basque*, ed. by Max W. Wheeler. Sussex: University of Sussex.
- Uhlenbeck, Cornelius. 1942. Les couches anciennes du vocabulaire basque. *Eusko-Jakintza* 1 (1947): 543–581. French translation of Dutch original by Georges Lacombe.
- Ulibarri, Koldo. 2013. External History. Sources for Historical Research. In Martínez Areta, 89–117.
- Urgell, Blanca. 2013. Euskal filologia: zer (ez) dakigu 25 urte beranduago? In Gómez et al. (eds.), 533–570.
- Vovin, Anatoly. 1994. (Art.-Review) Long Distance Relationships, Reconstruction Methodology and the Origins of Japanese, *Diachronica* 11: 1, 95–114.
- Watkins, Calvert. 1984. L'apport d'Émile Benveniste à la grammaire comparée. *Émile Benveniste aujourd'hui. Actes du Colloque international du CNRS*, ed. by Guy Serbat, I, 3–11. Louvain: Peeters.
- Watkins, Calvert. 1990. Etymologies, Equations, and Comparanda: Types and Values, and Criteria for Judgment. *Patterns of Change, Change of Patterns: Linguistic Change and Reconstruction Methodology*, ed. by Philip Baldi, 289–304. Berlin and New York: Mouton de Gruyter.
- Zuazo, Koldo. 1998. Euskalkiak gaur. *Fontes Lingua Vasconum. Studia et Documenta* 30: 191–233.