

# 19 Contact and Prehistory: The Indo-European Northwest

---

THEO VENNEMANN

The Indo-European languages, spread in ancient times over large parts of Eurasia and in recent times over large parts of the world, have emanated, according to the views of nearly all specialists, in prehistoric times from a small territory somewhere near the seam between Asia and Europe. Since all other parts of the double continent suitable for human settlement were already populated by non-Indo-European peoples by the time of the Indo-European expansion, all or nearly all historical Indo-European languages must have been in contact with non-Indo-European languages in prehistoric times, and this is certainly so for Europe north of the Central Divide, the mountain ranges separating central, west, and northern Europe from southern Europe: the Pyrenees and the Alps.

While this much is clear, it is less clear what the prehistoric non-Indo-European contact languages of Indo-European might have been. It is, indeed, least clear for the Indo-European languages north of the Central Divide. For while we possess suggestions, most of them vague and controversial, in the writings of ancient authors concerning possibly non-Indo-European contact languages, and in part even texts or individual names and expressions from those languages, these materials relate almost exclusively to the situation in southern Europe: Pelasgian in Greece, Etruscan in Italy, Iberian in Spain, to mention but three of those languages. For Europe north of the Central Divide there exists hardly anything, and for a huge territory comprising most of the German, French, English, and North Germanic speaking countries, there is close to nothing at all.

How does one identify possible earlier substrates of languages in territories where no such indigenous languages have survived? One looks for historical analogs and derives a rule of thumb. One analog is the question of the substrate of English in England. Here we know that the substrate was the language which survives in the mountainous regions of the west, Brittonic or Cymric, surviving as Welsh in regions of Wales. Another analog is the question of the substrate of English in Ireland. Here we know that the substrate was the Irish language which survives natively in a few remote parts of Ireland, the Gaeltracht. Yet another analog is the pre-Celtic Pictish language of the Celtic Isles: It survived longest in northern

Scotland and possibly in certain parts of Ireland. Then there is the language spoken in Bavaria before her Germanization in the sixth century. It was the Raeto-Romance language now only spoken in remote parts of the Alps. Clearly the rule of thumb for the identification of earlier substrates is this: Substrates of intruding languages survive longest on economically uninteresting or hard-to-access fringes. As the examples show, these fringe areas usually lie on the opposite side of where the invasion began.

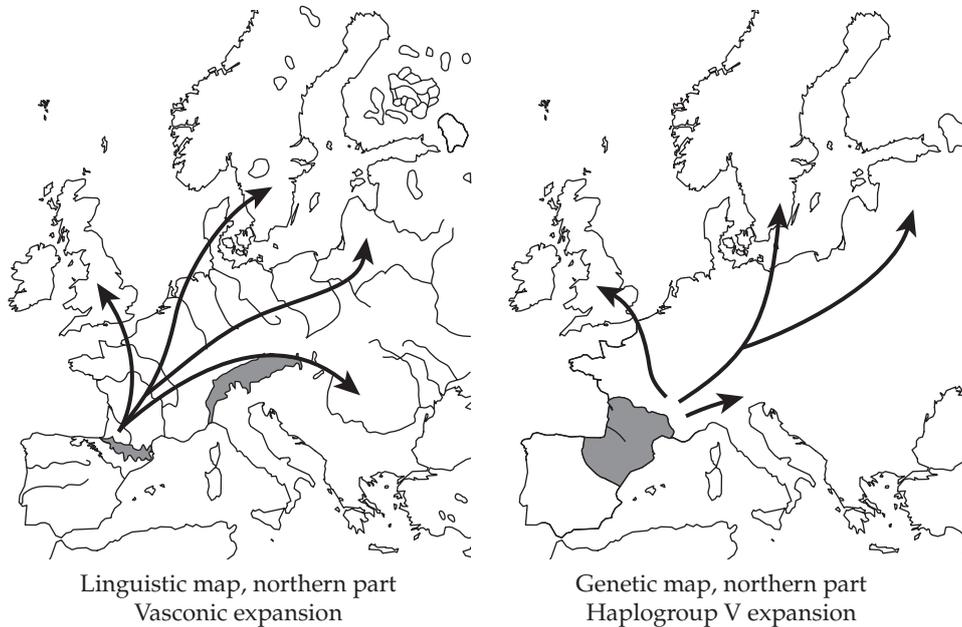
So let us look for a non-Indo-European language on the economically uninteresting or hard-to-access fringes of Europe north of the Central Divide, especially in the far west because the Indo-Europeans intruded into Europe from the east. The area that first comes to mind is the Alps. But the pre-Indo-European language there, Raetian, was most likely related to Etruscan and was thus itself the language of intruders, viz. from the Aegean. That leaves us with the Pyrenees. And there we find a non-Indo-European language that is not assumed by anyone with convincing arguments to be the language of recent intruders but is generally assumed to be indigenous to that part of the world: Basque. Our rule of thumb tells us that Basque is a survivor of the substrate of the Indo-European languages north of the Central Divide.

This is as it should be: During the last glaciation Europeans north of the Central Divide could only survive in southern France (and, farther away, in the Balkans). When the ice receded, Europe north of the Central Divide was repopulated by those survivors. The language of southern France was Basque; the southwest was Basque even in Roman times. The Vasconicity of Europe north of the Central Divide was demonstrated with linguistic means in the early 1990s. It was corroborated by genetic research soon afterwards. The two maps, one linguistic, one genetic, are almost identical.

In the present paper, only linguistic influences will be discussed, to be more precise: a selection of structural influences. For the numerous lexical influences, especially in the toponymy of Europe, reference is made to Freche (1995), Appelt (1998), Röder (2000), Böhm (2003), Welscher (2005), Vennemann (2006a; 2006b; 2008; and most of the chapters in Vennemann 2003).

## 1 Vigesimality

A prototypically structural property of languages is the way the larger numerals are constructed in them.<sup>1</sup> Since the number of natural numbers is infinite and even the number of large numbers used in languages of advanced civilizations is too large to name them individually by means of simplex words, principles by which to construct designations of large numbers starting from the individually named small numbers are needed. What we therefore find in such languages is an isolatable portion of the grammar serving this purpose: a primitive vocabulary of its own designating the smallest numbers and a set of rules for the sole purpose of constructing the larger numbers needed by the speakers of the language, in principle any natural number of arbitrary magnitude ("ad infinitum"). Some of



**Figure 19.1** Northern and eastern Vasconic expansion

Linguistic map, northern part, Vasconic expansion<sup>a</sup>

Genetic map, northern part, Haplogroup V expansion<sup>b</sup>

a Cf. Vennemann (2003: xv) (originally prepared for Vennemann 1996)

b "Map of Europe depicting the most likely homeland of haplogroup V and its pattern of diffusion" from Torroni et al. (1998: fig. 4)

these rules may be of the word formation kind, others purely syntactic, depending to some extent on the typological character of the language. Ideally such a system has simplex numbers from 1 to some number  $n$  (the "basic unit of counting"), then builds multiples of  $n$  filling up the gaps by adding the simplexes one by one, until  $n$  times  $n$  is reached. Then the addition continues. The multiples  $n$  times  $n$ ,  $n$  times  $n$  times  $n$  etc. may be given names of their own. But the system really only becomes fully workable if the number zero is added and a written notation is developed. Mixed systems are common.<sup>2</sup>

A curious feature of the western Indo-European languages is the occurrence of vigesimality, i.e. counting with 20 as a basic unit,<sup>3</sup> either alongside or – partly – instead of decimality, counting with 10 as a basic unit.<sup>4</sup> I say curious because the basic unit of counting in English is ten (*-teen, -ty*), and this has been so since Proto-Indo-European times.<sup>5</sup> We count from 1 to 10, then add 1 (*eleven*) and again 1 (*twelve*) and again 1 (*thirteen*) until we reach twice 10 (*twenty*), then further until we reach three times 10 (*thirty*), then *forty, fifty* and so on up to 10 times 10 (*one hundred*); etc. English is decimal, and this is an inherited Indo-European feature.

Yet we also find in English a modicum of vigesimality. There existed, in Great Britain until 1971, the peculiarity that there were 20 shillings to the pound. There even exists a word for '20 of its kind', *score*, borrowed from Scandinavian. Many speakers know it as the second word of President Abraham Lincoln's Gettysburg address of 19 November 1863, which begins with the sentence, "Four score and seven years ago our fathers brought forth on this continent a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal." The reference is to the year of the Declaration of Independence of the United States of America from the British Empire on 4 July 1776. Thus Lincoln was looking back 87 years, and this is what *four score and seven years* means. Christians may remember psalm 90, verse 10: "The days of our years are threescore years and ten; and if by reason of strength they be fourscore years, yet is their strength labor and sorrow."

In French one counts decimally up to 60: *dix, vingt, trente, quarante, cinquante, soixante*. Then, strangely, 70 is not *septante*<sup>6</sup> but *soixante-dix*, i.e. 60 [and] 10, and, even more strangely, 80 is *quatre-vingt*, 4 [times] 20, and 90, *quatre-vingt-dix*, 4 [times] 20 [and] 10.

Nonetheless these are but individual vigesimal concepts and numerals. We have to look beyond English and Modern French to see veritable vigesimal number systems in Northwestern Europe.<sup>7</sup>

### 1.1 Vigesimality in Germanic (Danish)

The only Germanic language with a vigesimal counting system is Danish. The following cardinal and ordinal numbers are taken from Bredsdorff (1970: 74–7):

(1)	10	<i>ti</i>	10th	<i>tiende</i>
	20	<i>tyve</i>	20th	<i>tyvende</i>
	30	<i>tredive</i>	30th	<i>tredivte</i>
	40	<i>fyrre, fyrretyve</i>	40th	<i>fyrretyvende</i>
	50	<i>halvtreds, halvtredsindstyve</i>	50th	<i>halvtredsindstyvende</i>
	60	<i>tres, tresindstyve</i>	60th	<i>tresindstyvende</i>
	70	<i>halvfjerds, halvfjerdsindstyve</i>	70th	<i>halvfjerdsindstyvende</i>
	80	<i>firs, firsindstyve</i>	80th	<i>firsindstyvende</i>
	90	<i>halvfems, halvfemsindstyve</i>	90th	<i>halvfemsindstyvende</i>
	100	<i>hundrede</i>	100th	<i>hundrede</i>
	100	<i>hundrede</i> : Old Danish also: <i>femsynnætiffwæ, femsindetiuqe, femsynnomtiwffwe</i> , i.e. 'five times 20'		

The way these numbers are constructed is not immediately clear. Therefore I cite Bredsdorff's explanations:

The names of the Dutch numerals from 20 to 90 are a queer mixture of the 10-system and the old 20-system . . .

20 *tyve* is by origin the plural of *ti*. The meaning of the singular is still seen in *tredive*, i.e. 'three tens', and *fyrre*, i.e. 'four tens', occasionally pronounced "fyrretyve" [fɔrrəty:və]. Here the ending has in time been confused with the word "tyve."

From 50 the real 20-system begins. The numeral "halvtreds" is an abbreviated form of "halvtredsindstyve," which literally means "half third (i.e. 2½) times ('sinde' . . .) twenty." Similarly "tres(indstyve)" has the original meaning of 'three times twenty'; "halvfjerds(indstyve)" means 'half fourth (i.e. 3½) times twenty'; "firs(indstyve)" means 'four times twenty'; and "halvfems(indstyve)" means 'half fifth (i.e. 4½) times twenty'.

The short forms of the cardinals given in the list above ("fyrre," "halvtreds," etc.) are generally used, but it is possible (e.g. for special emphasis) to use the long forms ("fyrretyve," "halvtredsindstyve," etc.). The ordinals are all formed on the basis of the long forms ("fyrretyvende," "halvtredsindstyvende," etc.). (Bredsdorff 1970: 77)

Ross and Berns write about this way of counting:

The formation of Mod. Danish *tresindstyve* 'three times twenty' and the like is fairly straightforward; it is the *raison d'être* that is difficult. The vigesimal forms develop in Old Danish<sup>8</sup> and are dominant in Modern Danish. In the relevant formations the second element is, naturally, always 'twenty'. (Ross & Berns 1992: 612)

The formulation "The vigesimal system developed in Old Danish" is not quite accurate because the system exists in Danish from the very beginning of its documentation. As a matter of fact, "In Old Danish, vigesimal counting precedes beyond the decades to the early hundreds" (Eliasson 2006: 102). The largest vigesimal numeral cited by Eliasson (2006: 103) is 340, *søtthensindstyve* (i.e. '17 times 20').

## 1.2 *Vigesimality in Celtic*

Insular Celtic vigesimality only developed in historical times. Old Irish<sup>9</sup> – at least written Old Irish – still shows the inherited Indo-European decimality (Greene 1992: 511):

(2)	10	<i>deich</i>	30	<i>trícho</i>	50	<i>coíco</i>	70	<i>sechtmogo</i>	90	<i>nócho</i>
	20	<i>fiche</i>	40	<i>cethorcho</i>	60	<i>sesco</i>	80	<i>ochtmogo</i>	100	<i>cét</i>

However, the vigesimal way of counting may already have been available in the spoken language, because it enters the written language in Middle Irish:

*Fiche* 'twenty' continues its Old Irish form and declension. During this period the vigesimal system begins to be normal, although all the decades up to and including 'ninety' are still attested. There are signs that the precision of the Old Irish system was breaking down. (Greene 1992: 525)

Classical Modern Irish shows the triumph of vigesimality:

*Fiche* 'twenty' plays an increasingly important part in the system . . . All these shifts arise from a tension between the literary standard, which tried to preserve the decades,

and the spoken language, which had undoubtedly gone over to the vigesimal system by this time; no modern Irish dialect has preserved any of the decads above 'twenty', nor is there any trace of them in Scottish Gaelic or Manx. (Greene 1992: 530)

Greene here draws the typical picture of the rise of substratal features from the language of the lower classes to the language of the ruling class and thus to the written language.

### 1.3 Vigesimality in Romance

In Modern Standard French only 70, 80, and 90 are named vigesimally. By contrast, Old French seems to have been thoroughly vigesimal. Price (1992) writes:

Vigesimal forms not now found in Standard French occur at earlier periods. Nyrop (1960: §490) quotes the following attested Old French forms:

- [30] *vint e dis*
- [40] *deus vins*
- [60] *trois vins*
- [70] *trois vins e dis*
- [80] *quatre vins*
- [90] *quatre vins e dis*
- [120] *sis vins*
- [140] *set vins*
- [160] *huit vins*
- [180] *neuf vins*
- [220] *onze vins*
- [240] *douze vins*
- [280] *quatorze vins*
- [300] *quinze vins*
- [320] *seize vins*
- [340] *dis set vins*
- [360] *dis huit vins*

Of these, *six-vingts* is well attested in the seventeenth century . . . The form *quinze-vingts* survives in the name of the *hospice de Quinze-Vingts*, founded by St Louis in 1260 as an asylum for three hundred blind people. There is also a *rue des Quinze-Vingts* in Troyes. (Price 1992: 463–4)

Romance vigesimality is not restricted to French. I simply quote Price's succinct account:

In Franco-Provençal, Maps 1239 and 1240 of the *ALF* [*Atlas linguistique de la France*] give forms corresponding to *trois-vingts* ([tre vɛ̃] etc.) *trois-vingt-dix* for some points in Savoie, with the comment that **this usage is "vieilli."** With reference to the patois of Bagnes, one of the best preserved Franco-Provençal dialects in Switzerland, we are told (Bjerrome 1957: 68) that some vigesimal forms were maintained until recently, "**pour indiquer le nombre de vaches d'un alpage,**" e.g. *wī vē vātse* (= *huit vingts vaches*), *sà vē vātse e demyī* (= *sept vingts vaches et demi*, i.e. 'one hundred and fifty cows') . . .

Occitan generally retains a decimal system. However Palay (1961) gives *trés-bints* (= *trois-vingts*) as well as *chichante* for 'sixty' and, under *cén(t)*, comments: "On emploie souvent, au lieu de *cén*, le comp. *cinq bints*. Various other Gallo-Romance forms (including *deux-vingts* in Haute-Marne, and a parallel form *düvê* in Savoie (ALF Map 1110, Point 965) are quoted in von Wartburg (1922–, 14: 443–444) . . .

Outside the Gallo-Romance area elements of a vigesimal system are well established in Southern Italy (see Rohlfs 1966–1969, §975–976), particularly in Sicily (*du vintini* 'forty', *du vintini e ddèci* 'fifty', etc., up to *cincu vintini* 'hundred'), but also in parts of the southern mainland, e.g. *dua/tri/quattro vintini* in various Calabrian dialects and parallel forms as far north as the Abruzzi. In some cases the system is used for numerals up to 'fifteen times twenty' = 'three hundred', e.g. *quánnici vintini* (Cosenza) and *diecentine* 'two hundred' and *quindice intine* 'three hundred' at Vernole. These forms are of course based on collectives (corresponding to Fr[ench] *vingtaine*), but forms corresponding exactly to the French type are found in Salentine dialects (*quattro vinti*). Widespread though the vigesimal system is in southern Italy, the decimal system coexists with it, and **the use of the vigesimal system is restricted to specific functions, e.g. for stating a person's age or for counting eggs, fruit, etc.** Sporadic forms occur in Ibero-Romance. (Price 1992: 464–6)

The Latin way of counting was purely decimal. Therefore the existence of vigesimality in wide parts of the Romance world is in want of explanation. What is also in want of explanation is that vigesimality is a feature only of western Romance; for Greene (1992: 463) writes, "Rumanian stands alone in having entirely abandoned the Latin forms for the decads. However, although the Rumanian forms do not reflect their Latin equivalents, the decimal system itself remains intact." The passages I have highlighted by boldface in the above quotations from Price (1992) show a close connection of counting by 20 to elemental aspects of life. Non-inherited linguistic features of this sort are typical substratal residues.

#### 1.4 *On the origin of western Indo-European vigesimality*

There is no generally accepted account of the origin of vigesimality in the western Indo-European languages. As may be expected, there are two sorts of theories, one ascribing it to spontaneous indigenous innovation, the other, to contact with other languages.

Since I find it hard to believe that speakers with a decimal system spontaneously switch to a vigesimal one, in my view the only acceptable sort of explanation is one based on language contact. Price (1992: 469) offers Celtic, Norman, and Basque as possible giving languages. Of these, Celtic is most often mentioned, probably because most authors are impressed by the well-known partial vigesimality of French and think of this language as having developed on a Celtic (Gaulish) substratum. Tagliavini (1998), for instance, in his chapter on the Celtic substrate, mentions vigesimality in his main text only with regard to French.<sup>10</sup>

Keltisch ist vielleicht das System der Vigesimalzählung, von dem das Französische in *quatre-vingt(s)* einen Rest bewahrt hat (das Altfranzösische kannte noch *treis-vinz* 'sechzig', *sis-vinz* 'hundertzwanzig'). Gleichwohl gibt es Beispiele dafür auch in anderen

Sprachen, und andererseits existieren Spuren des Typs *huitante, nonante* in französischen Dialekten. (Tagliavini 1998: 101)

[Celtic is probably the vigesimal counting system which is the source of the French remnant which can be recognised in *quatre-vingt(s)* (in Old French there was *treis-vinz* 'sixty' and *sis-vinz* 'a hundred and twenty' as well). Furthermore, there are examples in other languages and there are also traces of the *huitante, nonante* type in French dialects. (Tagliavini 1998: 101).]

In my view Celtic should not even be mentioned as a possible source for western Indo-European vigesimality because (a) Celtic is Indo-European and therefore originally decimal; (b) historical Insular Celtic started out decimal, becoming vigesimal only during the Middle Ages; and therefore, (c) Gaulish, lying chronologically between Proto-Celtic and Insular Celtic, must be assumed to have been decimal, a conclusion supported by the fact that "the only known relevant Gaulish form, *tricontis* 'thirty', fits clearly into the decimal system" (Price 1992: 466).

Holding Norman, i.e. the Vikings and, in the final analysis, Danish responsible for western Indo-European vigesimality<sup>11</sup> is equally unexplanatory because Proto-Germanic, as a branch of Indo-European, was decimal, and so was Proto-Norse. Therefore, even if it is admitted that the Normans had some role in the spread or consolidation of vigesimality, e.g. in southern Italy and Sicily, Danish vigesimality is not the explanans but is itself an explanandum.

That leaves us with Basque. And indeed, since vigesimality in decimal Indo-European cannot be explained with an Indo-European, i.e., by default, decimal substratum, non-Indo-European Basque should be considered first rather than last as a possible substratal source of western Indo-European vigesimality.<sup>12</sup> Basque is vigesimal, and even though Basque texts of any length are not older than the sixteenth century CE there is no indication that it ever was anything but vigesimal.

(3) Standard Basque cardinal numbers (Trask 2003: 127; King 1994: 414)

1 <i>bat</i>	11 <i>hamaika</i>	10 <i>hamar</i>
2 <i>bi, biga ~ bi</i>	12 <i>hamabi</i>	20 <i>hogeï, hogoi</i> <sup>13</sup>
3 <i>hiru, hirur</i>	13 <i>hamahiru, -r</i>	30 <i>hogeitahamar</i> <sup>14</sup>
4 <i>lau, laur</i>	14 <i>hamalau, -r</i>	40 <i>berrogeï</i>
5 <i>bost, bortz</i>	15 <i>hamabost, -bortz</i>	50 <i>berrogeitahamar</i>
6 <i>sei</i>	16 <i>hamasei</i>	60 <i>hirurogeï</i>
7 <i>zazpi</i>	17 <i>hamazazpi</i>	70 <i>hirurogeitahamar</i>
8 <i>zortzi</i>	18 <i>hamazortzi</i>	80 <i>laurogeï</i>
9 <i>bederatzi</i>	19 <i>hemeretzi</i>	90 <i>laurogeitahamar</i> <sup>15</sup>
10 <i>hamar</i>	20 <i>hogeï</i>	100 <i>ehun</i>

21 *hogeitabat*, 22 *hogeitabi*; 105 *ehun eta bost*; 1000 *mila*

1979 *mila bederatziehun hirurogeita hemeretzi*  
1000 900 60-and 19

As one can see, this particular historical system shows some adjustment to the Latin and modern Romance decimal way of counting. In particular the special

role of the numeral 100 is non-vigesimal in spirit. As the Old French and Old Danish systems show, genuine vigesimal counting goes beyond 100, 100 itself simply being 5 times 20. Regional vigesimal variants of numerals beyond and including 100 suggest that in Basque too this limit is a modern innovation and that in older Basque vigesimal counting continued without limit. Eliasson (2006: 103) cites the following examples from Lafitte (2001: 77) and *DGV*, 4.363, 5.541:

- (4) 100 *bortzetan hogoi*  $5 \times 20$   
 120 *seietan hogoi*  $6 \times 20$   
 160 *zortzetan hogoi*  $8 \times 20$   
 180 *bederatzitan hogoi*  $9 \times 20$

Assuming a Basque substratum to be responsible for western Indo-European vigesimality was problematic in the past because the Basque territory was viewed as too small to account for the wide spread of vigesimality in Europe. Naturally, Basque vigesimality could be considered an import from Romance. But that this was untenable was recognized even in the absence of an alternative explanation:

Entwistle (1936: 18) suggests that the Basque vigesimal system “may be of Celtic provenance,” but there is no evidence to support this and, if Gaulish in fact had no vigesimal system and the vigesimal system of Welsh, Irish, etc. is a comparatively recent innovation, then of course Entwistle’s suggestion lacks any foundation at all. (Price 1992: 490, n. 30)

However, the role of Basque has changed completely in the Vasconic theory of prehistoric Europe: Since almost all of western, central, and northern Europe is assumed to have been Basque in this theory, Basque structural patterns may be expected to be found everywhere in this area – that is, the source of these features is Basque.<sup>16</sup>

It must be stressed that vigesimality is merely a set of structural patterns and its importation into another language is not tied to the names of the basic numerals. Uncontrolled foreign language learning, as in language shifting, is initially the learning of foreign words and putting them in the patterns of the native language:

(5) Latin substance and Vasconic form in French vigesimal counting

			‘4’	‘20’
Substance:	Latin		<i>quattuor</i>	<i>viginti</i>
			↓	↓
Form:	Basque	‘80’	<i>laur-</i>	<i>-ogei</i>
			↓	↓
Result:	French	‘80’	<i>quatre</i>	- <i>vingt</i>

Only continued intensive learning may lead to greater approximation even to the structural targets of the new language.

## 2 Two Copulas

Proto-Germanic, as indeed Proto-Indo-European, had only a single copula. This may not strike anyone as peculiar, because the same is true for Contemporary English as well as the English of Shakespeare and of Chaucer. It is also true for historical North and East Germanic, as can be seen in the Old Norse and Gothic present indicative paradigms translating the forms of English *be*:

(6) Old Icelandic	Gothic	
<i>em</i>	<i>im</i>	'(I) am'
<i>es(t)</i>	<i>is</i>	'(thou) art'
<i>es</i>	<i>ist</i>	'(he/she/it) is'
<i>erom</i>	<i>sijum</i>	'(we) are'
<i>eroð</i>	<i>sijuþ</i>	'(you) are'
<i>ero</i>	<i>sind</i>	'(they) are'

But this is not true for Old English.

### 2.1 *Two copulas in Old English and Celtic*

All Old English dialects had, from the time of their earliest attestation, two copulas, each with a complete present indicative paradigm:

(7) Old English, West Saxon		
<i>s</i> -paradigm <sup>17</sup>	<i>b</i> -paradigm <sup>18</sup>	
<i>eom</i>	<i>bēo</i>	'(I) am'
<i>eart</i>	<i>bist</i>	'(thou) art'
<i>is</i>	<i>biþ</i>	'(he/she/it) is'
<i>sind(on)</i>	<i>bēoþ</i>	'(we/you/they) are'

Of these, the *s*- or *eom*-paradigm is recognizably a formal continuation of the Germanic paradigm<sup>19</sup> also reflected in Old Norse and Gothic, while the *b*- or *bēo*-paradigm is an innovation. As to the meaning of these two copulas in Old English, Campbell writes:

*bēo* expresses what is (a) an invariable fact, e.g. *ne bið swylcē cwenlic þeaw* [Beowulf 1940] 'such is not a queenly custom', or (b) the future, e.g. *ne bið þe wilna gad* [Beowulf 660] 'you will have no lack of pleasures', or (c) iterative extension into the future, e.g. *bið storma ġehwylc aswefed* [Phoenix 185–6] 'every storm is always allayed' (i.e. on all occasions of the flight of the Phoenix, past and to come); *eom* expresses a present state provided its continuance is not especially regarded, e.g. *wlitig is se wong* [Phoenix 7] 'the plain is beautiful'. (Campbell 1959: 350)<sup>20</sup>

This un-Germanic twofold paradigm for the copula was explained as a contact phenomenon as early as 1925, when Keller pointed to the formal similarity of the

*bēo*-Paradigm with the *b*-paradigm of the Celtic languages, for which see Pedersen (1976: II. §§636–41), especially the tables in sections 637 and 645, and Lewis and Pedersen (1989: §§476–86), especially the tables in sections 477 and 485. Whereas the Celtic *s*-paradigm rarely shows its *s* anymore, owing to intense phonological change, the *b*-paradigm is clearly recognizable as such in all Insular Celtic languages. As to the origin and meaning of the Celtic paradigms, Lewis and Pedersen (1989) state:

The paradigm of the verb ‘to be’ consists in Italo-Celtic of forms of the roots \**es-* and \**bheu-*. In Celtic a pres. stem \**bhwī-*, \**bhwije-*, derived from \**bheu-*, also appears. This latter present denotes either a praesens consuetudinale or a future, a natural development from an orig[inal] meaning ‘to become’ (Lat. *fiō*). The same root is also used in the subjunctive. The root \**es-* stands only in the pres. and ipf. ind. in Celtic; in Ir[ish] it is not found in the ipf. (Lewis & Pedersen 1989: §476.1)

The paradigms added to the inherited *s*-paradigm in Old English and in Middle Welsh are remarkably similar both as to form and to meaning: All forms in the paradigm of both languages begin with a *b-* followed by a front vowel; and the meanings formulated by the specialists – “(a) an invariable fact . . . or (b) the future . . . or (c) iterative extension into the future” in Old English and “a praesens consuetudinale or a future” for Celtic – are close enough to invite the idea that the innovations did not arise independently. Keller (1925: 59) too emphasized both facts and said specifically that the functional agreement was especially remarkable because it implied “a greater similarity of thinking between Anglo-Saxons and Britons than between Anglo-Saxons and Frisians or Germans.”<sup>21</sup> Indeed, for whereas both Old English on one hand and Frisian and Old Low and High German on the other differ from North and East Germanic showing *b*-forms for the expression of ‘to be’, only Old English has developed them into a separate second present tense paradigm, while the other West Germanic languages combine *b*-forms with *s*-forms in a single paradigm, without any evidence for a split as in Old English.

(8) Frisian	Old Saxon	Old High German
<i>bim</i>	<i>bium</i>	<i>bim</i>
	<i>bist</i>	<i>bist</i>
<i>is</i>	<i>is</i>	<i>ist</i>
<i>sind</i>	<i>sind(um)</i>	<i>birum</i>
		<i>birut</i>
		<i>sind</i>

Schumacher (2007) sees the *b*-forms in Frisian and German as evidence for a separate, contact-induced prehistoric West Germanic *b*-paradigm which was conflated with the inherited *s*-paradigm before the earliest Frisian and German documents but preserved in Anglo-Saxon through the continued contact with Celtic.

This view is considered unlikely by Lutz (2009): Contact-induced grammatical categories do not result from borrowing but through language-shifting. Schumacher's view presupposes that the West Germanic peoples are for the most part Celts who learned Germanic. There is no independent evidence for this to be true – except in the case of the English. Nevertheless the Frisian and German *b*-forms are important: They show that not only Insular Celtic but also Continental Celtic had the *b*-forms, and by implication that the Celtic separate *b*-paradigm originated on the Continent and was taken from there to the Isles.<sup>22</sup>

At this point an important question arises: If we explain with Keller the fact that “a greater similarity of thinking [exists] between Anglo-Saxons and Britons than between Anglo-Saxons and Frisians or Germans” came about because large numbers of Britons shifted to Anglo-Saxon in the centuries after the Conquest, carrying their way of expressing themselves – their “thinking” – into the target language, then how do we explain that the Celts themselves differ from the rest of the Indo-Europeans by this special way of thinking? How did they acquire this un-Indo-European manner of speaking?

The simplest answer would be one beginning with the words, “In the same way.” It would also be the best answer, because it would not require additional or different theoretical assumptions. The double paradigm would then simply exemplify the model which I named “the transitivity of language contact” (Vennemann 2002). It suggests that for Celtic we look into the substratum of Celtic in Central Europe, just as for Old English Keller looked into the substratum of Anglo-Saxon in Britannia. So let us look into Basque and see if that language too has two paradigms for ‘to be’, two copulas.

## 2.2 Two copulas in Basque

Trask (1997: 113) emphasizes the difference between two meanings, eventive and stative, in connection with two verbs ‘to be’: “Note the difference between *moxkortu naiz* ‘I got drunk’ (a little while ago) (eventive, with *izan*) and *moxkortuta nago* ‘I’m drunk’ (stative, with *egon*).” For these two verbs corresponding to English ‘to be’, *izan* and *egon*, one eventive, the other stative, de Azkue (1984, s.vv.) provides, besides French *être* for both, the following Spanish equivalents, each as the first (1°) of several translations, where “(c)” indicates that the item is “común . . . á toda la lengua,” rather than being restricted to one or several of the dialects or individual communities:

- (9) IZAN . . . 1° (c) ‘ser’  
EGON 1° (c) ‘estar’

While both verbs and the semantic distinction thus are *común*, several authors mention that the distinction is more common in the west or in the south of the Basque dialect area, as in the following instances.

*Copula.* All varieties have a copula, used both in equational sentences and in predications of class membership. Southern varieties distinguish two copulas, the second being roughly confined to expressing temporary states or conditions and locations (temporary or permanent). (Trask 1997: 119)

**Izan** is the verb *to be*: **Irakaslea naiz** *I am a teacher*. In the third person it can also express *there is/are*: **Bada eliza bat** *There is a church*. **Izan** is also used as an auxiliary. It is intransitive . . . **Egon** is another verb meaning *to be*, and also *stay* or *wait*: **Isilik dago** *He is silent*; **Zaude hemen!** *Wait here!* It is characteristic of western dialects and is intransitive. (King 1994: 362)

Agud and Tovar (1993, s.vv.) even seem to restrict the occurrence of the second of these two words to the westernmost dialect:

- (10) IZAN ‘ser’, ‘estar’, ‘haber’, ‘tener’, ‘soler’  
EGON, IGON V[izcaíno] ‘estar, quedar’

But the phenomenon is given much space in grammars of Basque. Let me cite just two recent accounts. Sagüés (1994: 45–6) writes about *izan*:

El verbo *izan* desempeña en euskara un doble papel:

– por un lado corresponde al verbo «ser» castellano.

P.e. **ni zaharra naiz** : yo soy viejo  
**zu gaztea zara** : tú/ud. eres/es joven

– y por otro lado actúa como verbo auxiliar.

P.e. **ni etorri naiz** : yo he venido  
**gu etorri gara** : nosotros hemos venido<sup>23</sup>

Etxepare (2003) begins the section about “Copular constructions” with the following description and examples:

Basque makes a distinction between stage-level predications (those which attribute some transitory property to the subject of predication) and individual-level predications (those which attribute some standing property to the subject of predication) in the auxiliary selected to express them. Transient properties are assigned by the verb *egon* ‘be in a location’, whereas standing properties are assigned through the verb *izan* ‘be’. The distinction, which is for the most part limited to western dialects, is reminiscent of the one found in Spanish between *ser* and *estar* . . . *Izan* is also used in equative sentences. (Etxepare 2003: 365)

Since it is not known how old this way of “thinking” in terms of two copular verbs is in Basque, whose oldest extensive documentation did not begin before the sixteenth century, there is no way of proving that the Vasconic substrate of Celtic on the Continent did have this feature. There is, however, indirect evidence for the assumption that it goes back to prehistoric times.

Where we find late prehistoric or early historical – let us say, recent – developments of a second copula, as in Celtic and in Northwestern Romance, its origin in a verb with the meaning ‘to become’ (\**bheu-*, Celtic) or ‘to stand’ (Latin

*stāre*, Romance) is transparent. Similarly in Basque, where an original meaning 'to stay' or 'to wait' can be recognized for *egon*. But beyond that, Agud and Tovar (1991, s.v.) write about *egon*: "Es otra de las palabras vascas importantes<sup>24</sup> cuya etimología no está en modo alguno resuelta."

Another indirect argument follows from the way the development of a second copula – from Latin *stāre* 'to stand' alongside the continuation of inherited *esse* 'to be' – is distributed in the Romance languages.

### 2.3 Two copulas in Northwestern Romance

The two copular paradigms based on Latin *esse* and *stāre* exist in Portuguese, Galician, Castilian, and Catalan as well as, to a lesser extent, in Italian and, even lesser, in Sicilian, not at all in Rumanian. They once existed in French:

Modern French has only one copula. Old French, however, had two, *estre* (ESSE > *essere* > \**essre* > *estre*) and *ester* (STARE > \**estare* > *estar* > *ester*), and distinguished between them in a similar way to other Romance languages. With phonetic evolution, the forms of each verb tended to be confused with one another, with the result that *estre* finally absorbed *ester*; around the same time, most words beginning with *est-* changed to *ét-* or *ê-*. The modern form of the verb is *être*.

The only clear trace of *ester* (or *éter* if we bear in mind the loss of the *s*) in the modern copula is the past participle: instead of the \**étu* one would expect, we find *été*.<sup>25</sup>

The areas south of the Pyrenees, as well as north of the Pyrenees and the Alps, plus to a lesser extent northern Italy, are precisely the regions assumed to have been Vasconic before the Indo-Europeanization of Europe.<sup>26</sup> The null hypothesis accounting for this distribution of the rise of a second copular paradigm – in northwestern Romance and in Continental Celtic – is, on the basis of the theory of a once-Vasconic Western and Central Europe, that this particular way of syntactic "thinking" in terms of two copulas was carried from the Vasconic substrate into those Indo-European superstrates in the process of language shifting.

An argument against the above contact interpretation could be the fact that in Contemporary Basque the use of the two copulas is primarily a feature of the western dialects, so that superstratal transfer from Spanish to Basque should be assumed (cf. Trask 1997: 292–3). The correct interpretation of this distribution in the dialects is, on the contrary, that the extended superstratal contact with French, which reverted to a one-copula syntax centuries ago, led to diminished use of the second copula *egon* in the eastern dialects of Basque.

Needless to say there have been attempts to explain the rise of a second copula in Romance, especially in Spanish, in terms of the internal dynamics of change.<sup>27</sup> For example, Posner (1996: 313) suspects three language-internal "factors" as having contributed to this development. She does not seem to notice, however, that the explanantia may themselves be in need of explanation. Thus, that "another factor that may have played a part is early loss in the Iberian languages of full lexical meaning of *estar*" is actually part of the explanandum rather than

an explanans. Posner does not consider the possibility of external influence. Pfaller (2003) uses the model of “expressive change” of Koch and Oesterreicher (1996) and Detges (2001) to develop the view that the expressive use of the position verbs *sedēre*, *stāre*, and *iacēre* gave rise to a semantic change which, as a consequence of increasingly “normal” use, ran from the positional *Grundbedeutung* ‘sit’, ‘stand’, and ‘lie’ via merely locational semantics toward ‘be’.<sup>28</sup> She does not ask the question why the same kind of change does not affect the corresponding position verbs in German or Russian in same or similar ways, nor does she consider the possibility of external influence.

#### 2.4 *Two copulas in Irish English*

Finally, one may ask if “thinking in terms of two copulas” is at all transferable from one language to another, namely from a substratum to its superstratum. That is, one would like to see a bona fide case from a recent or present-day contact situation. Fortunately, there is such a case.

The two-copula system of Old English was simplified in Early Middle English.<sup>29</sup> What has remained to the present day is a mixed paradigm (*am*, *are* etc., *to be*, *been*) combining forms from the old *s*- and *b*-paradigms in a new paradigm with a uniform copular meaning. This is the situation in the standard language of England. Not so in Irish English. There we see again the rise of a second copula with an habitual meaning, *do be* (+ *-ing*-form of the main verb), cf. Hickey (2007: 141):

- (11) Irish: *Bionn sé ag caint léi.*  
 [is-HABITUAL<sup>30</sup> he at talking with-her]  
 Irish English: *He does be talking to her.*  
 ‘He talks to her repeatedly.’

See also Hickey (2007: 173) and the list of examples on pp. 216–17, which includes sentences without *-ing*-forms:

- (12) *I do be up in a heap . . . I don’t be able to get out at all.*  
*All the dances we have now are céili dances. They do be very good.*

The formal side of this situation is not completely uniform. There are alternative ways of expressing the habitual-copular meaning, with different regional distributions or preferences. The chief alternative is invariable *be* (+ *-ing*-form of the main verb), e.g. *I be there every day* (Hickey 2007: 183). Another is inflected *be*, e.g. *The week-day be’s a quiet day* (Hickey 2007: 231). Yet another is unstressed *do* (used also in the past, *did*) + infinitive of the main verb (Hickey 2007: 216n):

- (13) *He does ‘fish.*  
*That’s what we did ‘say to each other.*  
*I did ‘never hear of that at all.*

This multiple formal representation of the habitual is not damaging to the idea of the transfer of “thinking in terms of two copulas” from substratal Irish to superstratal English in Ireland. Hickey explains this situation as a natural consequence of language shifting: When using English the Irish, with two copulas in their native language, search for a way of expressing the meaning of the second copula, the *b*-copula, in English, which does not offer an equivalent, and hit upon different solutions, which become different ways of regularly expressing habituality in developing Irish English, even for those speakers who are monolingual in Irish English. Hickey formulates this succinctly in terms of category and exponence:

With the habitual, it is particularly important to distinguish between category and exponence. The existence of the category habitual in Irish would have triggered a search for categorial equivalence among the Irish speakers during language shift. The exponence of the habitual in Irish does not have anything like a formal match in English, neither in its standard form and nor in the habitual structures which have arisen in Irish English historically. (Hickey 2007: 214n.)

Hickey discusses the use and the history of the expression of habituality in Irish English at great length (2007: 213–37). He leaves no room for doubt that it is a genuine innovation owed to the Irish substrate rather than, as has also been claimed, a continuation of the Old English double paradigm through Scots (2007: 227–31).

### 3 Accent

Proto-Indo-European had a variable accent. In three and only three of the Indo-European language subfamilies was this way of accentuation given up in favor of a fixed initial accent, namely a first-syllable accent: in Germanic, Italic, and Celtic:

Germanic: “Den überkommenen idg. Akzent hat das Germanische grundlegend verändert: es hat die Möglichkeit des ‘freien’ Akzentes völlig aufgegeben und ihn festgelegt auf die jeweils erste Silbe eines Wortes (Anfangsbetonung oder Initialakzent) . . . Aus dem starken Verfall der Endungen ergibt sich gleichzeitig, daß die Natur des germ. Wortakzentes eine vorwiegend dynamische gewesen sein muß” (Krahe & Meid 1969: vol. 1, §27).

[Germanic: ‘The traditional Indo-European accent underwent a major change in Germanic. The “free accent” was abandoned and accent was fixed on the first syllable of a word (initial accent) . . . As a result of the considerable reduction of inflections the nature of the Germanic word accent must have also been mainly dynamic’ (Krahe & Meid 1969, vol. 1. §27).]

Italic (Latin): “There is little disagreement that the prehistoric accent of Latin was a stress accent, and that this fell on the first syllable of the word. Its effects are seen in the loss or weakening of vowels in the unaccented syllables, which is typical of strong stress in some other languages” (Allen 1970: 83).

Celtic (Irish): Initial accent cannot be doubted in the case of Old Irish: “Words [in Old Irish] susceptible of full stress take this on the first syllable . . . The stress is expiratory and very intense, as may be seen from the reduction of unstressed syllables” (Thurneysen 1946: §36).

The Germanic first-syllable accent is evident in all early Germanic languages, and in their metrical systems based on alliteration; it must therefore be assumed to date back to Proto-Germanic. The situation is less straightforward for Italic and Celtic. “Latin is the only Italic language for which we have any information” (Sihler 1995: §246). For Continental Celtic there exists no direct evidence either. Gaulish place names show final accent (falling on one of the last three syllables) in their later phonological development. Brittonic has a final accent, which fell on the ultima, the penult, or the antepenult in the various sub-branches at different stages of development (Pedersen 1976: vol. 1, §180). But Schrijver (1995: 19–20) presents evidence that the predecessor of British had initial stress. “This strongly suggests that the system of initial stress that we find in Irish goes back at least to PInsCl [Proto-InsularCeltic] times.”

Celtologists are divided between the view that Proto-Celtic accentuation is unknown and the view that a Proto-Celtic first-syllable accent may be inferred. I side with the latter because the attested diversity can be best explained as deviations either caused by a later Romanization (the case of Gaulish toponyms) or by developments toward the universally preferred final accentuation, with final accent – ultimate or penult – as the commonest pattern (cf. Hyman 1977). Furthermore, Mercado (2007) has made a case that similarities between Italic and Old Irish meters are owed to common inheritance, which suggests that the accentuation in the two branches was the same at the time when those meters originated: “Cisalpine Celtic and Old Irish . . . point to a Proto-Celtic trochaic-dactylic colon most likely cognate with that of Proto-Italic,” this colon forming the basis of “the prehistoric accentual meters of initial-stressing Italic and Celtic” (Mercado 2007: Abstract).<sup>31</sup>

Clearly the assumption that not only Proto-Germanic but also Proto-Italic and Proto-Celtic had first-syllable accent yields the simplest overall account of the historical phonology and metrics of Germanic, Italic, and Celtic. These three were neighboring languages in prehistoric times, as impressively demonstrated by Krahe,<sup>32</sup> spoken most likely in a coherent central European area between southern Scandinavia and the Alps. The shift to initial accent is therefore best accounted for as a sprachbund phenomenon.<sup>33</sup> Since there is no language-internal explanation for the development of initial accent in this sprachbund, language contact with non-Indo-European languages has been held responsible for nearly a hundred years (since Feist 1913: 375).<sup>34</sup> Salmons (1992) shows that prosodic properties are transferred with special ease from substrata to their superstrata.<sup>35</sup> I have given arguments in Vennemann (1994: §7.6) for the assumption that the substrate of this Indo-European initial-accent sprachbund, the Old European language reflected in the Old European toponymy, had initial word accent, and in the same and

several other articles, that this language was Vasconic, which is also the position taken in the present chapter.

It would be welcome support for this explanation of the west Indo-European initial accentuation if it could be proved that the surviving member of the prehistoric Vasconic language family, Basque, had initial word accent at a sufficiently early stage of development. Since there is no attestation of very early Basque, the evidence can only be indirect. Following Martinet (1964) and Michelena (1977), I adopted this view in Vennemann (1994). However, Hualde has argued in a number of articles, most recently in Hualde (2003 and 2007), that all accentual systems observable in Basque (for which see Hualde 1999; 2003), including the initial accent systems reconstructed by Martinet und Michelena, have developed from an earlier stage at which the language lacked word accent and only possessed phrase-level prosody (Hualde 2007), a kind of system known from Modern French, though with different contours. This result does not, of course, preclude the existence of a word-accentual system with initial accent at a still earlier stage. The fact that consonant clusters occur almost exclusively between the first and second syllable points to first-syllable accent.

In concluding this section I would like to mention that Etruscan, a non-Indo-European, non-Vasconic language of northern and central Italy, was subject to first-syllable accent too. This is shown by the reductions, especially syncopations, observable over several centuries: These affected especially the second syllables of words, but also later syllables; they never affected the first syllables of words (Bonfante 1990: 335). Etruscan, the language of intruders from the Aegean area, had spread over a Vasconic substratum exactly as the western Indo-European languages,<sup>36</sup> and the substratum speakers likewise took their native initial accent into the intruding language they had to learn.

## 4 Conclusion

The occupation with language contact is not a recent phenomenon. Yet many linguists have shunned away from contact explanations in historical linguistics, and in some circles of Indo-European studies references to substratal influence amount to academic suicide. As late as 1998 I could assess the situation as follows.

Gewiß hat sich die Indogermanistik *auch* mit den nicht-indogermanischen Zügen der indogermanischen Sprachen befaßt, aber doch längst nicht mit derselben Intensität wie mit den deutlich indogermanischen. Auffällig ist auch, daß die Befassung mit den nicht-indogermanischen Zügen als wenig ehrenhaft zu gelten scheint. Man spricht, nicht selten abfällig, von der "Substrattheorie," und wenn ich es richtig höre, so meint die Verwendung des Ausdrucks "Theorie" in diesem Zusammenhang soviel wie "nur eine Theorie" – so als ob nicht alles in den empirischen Wissenschaften nur eine Theorie wäre und als ob nicht umgekehrt die stratale Beeinflussung aller Sprachen zu den beinharten Ergebnissen der Sprachkontaktforschung gehörte.

Gerade bei den west-indogermanischen Sprachen, bei denen die nicht-indogermanischen Kontaktsprachen (im Gegensatz etwa zu den Dravidia-Sprachen als Kontaktsprachen des Indischen) nicht ohne weiteres zu erkennen sind, kann von einer ernsthaften Bemühung der Indogermanistik um eine Erschließung der vor-indogermanischen Substrate und weiterer vor-indogermanischer Kontaktsprachen kaum die Rede sein. Ich meine, daß die andere Abteilung der Indogermanistik, in deren Skopus diese Probleme gehören, die gleiche Anstrengung der Indogermanisten verdient wie die ohnehin wesentlich weiter ausgebauten erste. (Vennemann 1998: 130–1)

[True, Indo-European studies has concerned itself with the non-Indo-European elements in Indo-European languages but by no means with the same intensity as with the clearly Indo-European elements. It is also remarkable that the investigation of non-Indo-European elements is not held in such high esteem. Not infrequently a dismissive tone is found in discussions of “substrate theory” and, if I have observed this correctly, the use of the word “theory” in this context means “just a theory” – as if everything in empirical science was not just a theory anyway. It is as if mutual influence among languages did not belong to the core insights of language contact studies.

Especially in the case of the western Indo-European languages, where the non-Indo-European contact languages cannot be so easily recognized (in contrast, for instance, with Dravidian in contact with Indic), one can say that Indo-European scholars are not particularly concerned with establishing what the pre-Indo-European substrate and contact languages were. I believe that the part of Indo-European studies which is concerned with these matters deserves the same attention as does that which focuses on inherited features of Indo-European languages. (Vennemann 1998: 130–1)]

It seems, however, that the increasing endeavor to develop contact linguistics as a principled subdiscipline of linguistics, which began more than half a century ago (Weinreich 1953) and has picked up enormous momentum in recent years (e.g. Thomason & Kaufman 1988; Goebel et al. 1997–8; Mufwene 2001; Thomason 2001; Winford 2003),<sup>37</sup> has changed the climate in favor of a more open attitude, witness, for example, the conference “Languages in Prehistoric Europe” (University of Eichstätt, 4–6 October 1999) and the resulting volume (Bammesberger & Vennemann 2003) which brought Indo-Europeanists and scholars of other fields together as speakers, discussants, and authors and proved, if nothing else, that the old reluctance to confront the topic seems not to exist any longer.

I hope to have shown in the present paper that dealing with the question whether, and if so to what extent, languages related to Basque may have substratally influenced in prehistoric times the west Indo-European languages north of the Great Divide, just as they did south of it, especially south of the Pyrenees (e.g. Baldinger 1963;<sup>38</sup> Garvens 1964), is a serious and promising enterprise which may take us a long way toward understanding why those languages show so many sprachbund-like commonalities in their grammar, lexicon, and toponymy.

## NOTES

- 1 This section is for the most part a condensed English version of Vennemann (2003: ch. 17, §17.1.3).
- 2 For a comparative introduction to number systems see the chapter “The number sequence” in Menninger (1992).
- 3 From Lat. *vigēsimus*, -a, -um ‘twentieth’.
- 4 From Lat. *decimus*, -a, -um ‘tenth’.
- 5 This is evident in the descriptions of the number systems of the Indo-European languages in Gvozdanović (1992), where deviations from decimality are always recognizable as secondary. Cf. also Beekes (1995: 212): “The numerals of PIE [Proto-Indo-European] can be reconstructed down to the last detail . . . The numeral system is based on the counting of decimals.”
- 6 The decimal variants *septante*, *octante/huitante*, *nonante* occur regionally; cf. Price (1992: 461–3 et passim). See also the Internet site “septante, octante, huitante, nonante,” [www.langue-fr.net/index/s/septante.htm](http://www.langue-fr.net/index/s/septante.htm) (accessed 12 May 2008).
- 7 A brief survey of vigesimality in Europe may also be found in Menninger (1992: 64–9).
- 8 The Old Danish forms are given in Ross and Berns 1992: 616–19.
- 9 I limit my quotations to Irish. The other Insular Celtic languages show a similar picture, according to Greene 1992.
- 10 Only in a bibliographical note is vigesimality in southern Italy mentioned.
- 11 Rohlfs (1971: §§ 96–8), discussing various possibilities between monogenesis and polygenesis, favors the idea of a Norman origin, even though he stresses the occurrence of vigesimality in older varieties of Spanish as a complicating factor and mentions Basque “im geographischen Übergang vom Französischen zum Spanischen” (p. 132). Menninger (1992: 110) tabulates the vigesimal system of Basque but considers the Normans the most likely importers of vigesimality into French (1992: 67).
- 12 Meillet (1964: 414) did reckon with a non-Indo-European source of western Indo-European vigesimality but remained unspecific: “Comme des traces plus nettes encore de système vigésimal se retrouvent dans le domaine celtique, on se demande si ceci n’est pas dû à une survivance d’un usage pré-indo-européen.”
- 13 The eastern forms have (*h*)*ogoi* also in the following numerals.
- 14 Here and in the following numerals *-tahamar* is shortened from *eta hamar* ‘and ten’.
- 15 More explicit forms occur for 60, 70, 80, and 90; e.g. for 90 *lauretan hogei eta hamar* ‘four-times twenty and ten’. See Eliasson (2006) for an explanation of the termination *-etan*.
- 16 Basque features generally become more sparse the farther east one looks. For example, in Balto-Slavic, which shows much of the Vasconic Old European hydronymy, traces of vigesimality are only found in the western languages and are explained as borrowings from Germanic and Venetian (Comrie 1992: 722–3 with n. 3).
- 17 I call this paradigm the *s*-paradigm because it is, with certain irregularities mentioned directly, the etymological continuation of the present of the Indo-European copula, based on the root  $^*h_1es-$ : sing.  $^*h_1és-mi$ ,  $^*h_1és-si$  >  $^*h_1ési$ ,  $^*h_1és-ti$ , plur.  $^*h_1s-més$ ,  $^*h_1s-té$ ,  $^*h_1s-énti$  (cf. Sihler 1995: §492).
- 18 I call this paradigm the *b*-paradigm because it is based on the Indo-European root  $^*bheu-$  and Proto-Indo-European  $^*bh$  became  $^*b$  in Germanic (and Celtic, see below). Whereas  $^*h_1es-$  inflected as a present/imperfect,  $^*bheu-$  was aoristic, probably because its original meaning was ‘to become’ rather than ‘to be’ (Sihler 1995: §491).

- 19 With some irregularities, explanations for which are rarely attempted in the handbooks (but see Onions 1969: s.v. *be*): The *eo* of *eom* may be analogical after *bēo*, the *ea* and *-t* of *eart* after the present perfects (e.g. *þu þearft* 'you need'); but the *-r-* of *eart* is obscure, possibly of very early Scandinavian influence.
- 20 The *DOE* (s.v. *bēon*) too lists three different types of usage distinctions for the *s-* and the *b-*paradigm, (a) present vs. future, (b) statal vs. actional, and (c) non-durative vs. durative: "The distinction is abandoned in early Middle English, earlier in Northern than in Southern and Southwestern texts (cf. *MED* s.v. *bēn*, *OED* s.v. *be*, Jost 1909: 139–40, Brunner 1962: 277–9" (Lutz 2009: 233, n. 19).
- 21 This is Lutz's translation of Keller's "eine größere Ähnlichkeit des Denkens bei Angelsachsen und Briten . . . als bei Angelsachsen und Friesen oder Deutschen." Keller was relying on the semantic study of the Old English double paradigm by Jost (1909), which has recently found support in the assessment of the *DOE*, s.v. *bēon*.
- 22 That individual forms may be borrowed into an existing paradigm in a situation of non-substratal contact is shown by Lutz (2009: 237–8) with Old Norse *erom* '(we) are', which entered the Anglian *s-*paradigm as a new Einheitsplural *aron* and in the course of time ousted the inherited *sind(on)*.
- 23 I omit further explanations of the use of *izan*.
- 24 Among them *izan*, cf. Agud and Tovar (1991, s.v.).
- 25 From "Romance copula" cf. [http://en.wikipedia.org/wiki/Romance\\_copula](http://en.wikipedia.org/wiki/Romance_copula) (accessed 5 May 2008).
- 26 In large parts of Italy there was an intermediate time of Etruscan domination.
- 27 The classical historical account for Spanish is Bouzet (1953). For work within generative grammar cf. Batllori (2006) and earlier articles. Recent work on copulas in general and in Romance includes Pustet (2003), Maienborn (2005; 2007), and several articles in *Theoretical Linguistics* 31.3 (2005).
- 28 Spanish *yacer* 'to lie' only showed the beginning of this semantic change in the Middle Ages but did not complete it.
- 29 Cf. Lutz (2009: 233, n. 19), who refers to *MED* s.v. *bēn*, *OED* s.v. *be*; Jost 1909: 139–40; Brunner 1962: 277–9.
- 30 The corresponding non-habitual copula form would be *tá*.
- 31 One may speculate that the development of fundamental metrical techniques of the new initial-accent languages was influenced by the meters of the initial-accent substrate language.
- 32 E.g. in Krahe (1954: 71–2, 171); Krahe & Meid (1969: vol. 1, §§3–7).
- 33 This is also the basic assumption in Salmons (1992).
- 34 Salmons (1992) too assumes such contact in the case discussed above but, differing from the position taken here, with Uralic languages.
- 35 The wisdom of everyday language acknowledges this: Speakers of second languages speak it "with an accent."
- 36 That Etruria was part of the prehistoric Vasconia is shown by its toponyms, e.g. the name of the Arno river which Krahe (1964: 46) reckoned among the Old European hydronyms and in which I recognize Basque *aran* 'valley' (Vennemann 1999: §3.1.2; 2006b: 975).
- 37 There is even a separate Internet site "Language contact," [http://en.wikipedia.org/wiki/Language\\_contact](http://en.wikipedia.org/wiki/Language_contact) (accessed 16 July 2008).
- 38 Baldinger shows that connections with Basque (Aquitania) are found as far west as Galicia and Portugal.

## REFERENCES

- Agud, Manuel and Antonio Tovar 1990–5. *Diccionario etimológico vasco*. Donostia/San Sebastian: Diputación Foral de Gipuzkoa. (To date 7 vols., A – Orloi.)
- Allen, W. Sidney 1970. *Vox Latina: A Guide to the Pronunciation of Classical Latin*. Cambridge: Cambridge University Press.
- Appelt, Andrea 1998. Zur Deutung des Flußnamens *Lech*. In Wilhelm F. H. Nicolaisen (ed.), *Proceedings of the Nineteenth International Congress of Onomastic Sciences, Aberdeen, August 4–11, 1996: Scope, Perspectives and Methods of Onomastics*, 3 vols. Aberdeen: University of Aberdeen, Department of English, vol. 2, pp. 359–63.
- de Azkue, Resurrección María 1984. *Diccionario vasco – español – francés*. Bilbao: Euskaltzaindia. (Reprint.)
- Baldinger, Kurt 1963. *La formación de los dominios lingüísticos en la península ibérica*. Madrid: Gredos. (Ch. 7 (pp. 161–230), “El Gallego-Portugués y sus relaciones de substrato con la Aquitania,” is available on the Internet, [www.instituto-camoes.pt/cvc/hlp/biblioteca/baldinger.pdf](http://www.instituto-camoes.pt/cvc/hlp/biblioteca/baldinger.pdf). Accessed 16 July 2008.)
- Bammesberger, Alfred and Theo Vennemann (eds.) 2003. *Languages in Prehistoric Europe* (Indogermanische Bibliothek, Dritte Reihe). Heidelberg: Carl Winter.
- Batllori Dillet, Montse 2006. Sincronia i diacronia de *ser i estar*. In *Actes del 13è Colloqui Internacional de Llengua i Literatura Catalanes (Girona, 8 al 14 de setembre de 2003)*. Barcelona: Publicacions de l’Abadia de Montserrat, vol. 2, pp. 43–54.
- Beekes, Robert S. P. 1995. *Comparative Indo-European Linguistics: An Introduction*. Amsterdam: John Benjamins.
- Bjerrome, Gunnar 1957. *Le patois de Bagnes (Valois)*. Stockholm: Almqvist and Wiksell.
- Böhm, Andrea 2003. *Probleme der Deutung mitteleuropäischer Ortsnamen, mit besonderer Berücksichtigung der Toponymie des deutschsprachigen Raumes und einem Ausblick auf den appellativischen Wortschatz des Deutschen* (Reihe Sprach- und Literaturwissenschaften, 8). Munich: Herbert Utz.
- Bonfante, Larissa 1990. Etruscan. In Jeremy T. Hooker (ed.), *Reading the Past: Ancient Writing from Cuneiform to the Alphabet*. Berkeley: University of California Press/British Museum, pp. 320–78.
- Bouzet, Jean 1953. Orígenes del empleo de “estar.” In *Estudios dedicados a R. Menéndez Pidal*, 7 vols. Madrid: CSIC, vol. 4, pp. 37–58.
- Bredsdorff, Elias 1970. *Danish: An Elementary Grammar and Reader*. Cambridge: Cambridge University Press.
- Brunner, Karl 1962. *Die englische Sprache: Ihre geschichtliche Entwicklung*, vol. 2: *Die Flexionsformen und ihre Verwendung*, 2nd edn. Tübingen: Max Niemeyer.
- Campbell, Alistair 1959. *Old English Grammar*. Oxford: Clarendon Press.
- Comrie, Bernard 1992. Balto-Slavonic. In Jadranka Gvozdanović (ed.), *Indo-European Numerals* (Trends in Linguistics, 57). Berlin: Mouton de Gruyter, pp. 717–833.
- Detges, Ulrich 2001. *Grammatikalisierung: Eine kognitiv-pragmatische Theorie, dargestellt am Beispiel romanischer und anderer Sprachen*. Unpublished Habilitation dissertation. University of Tübingen, Neuphilologische Fakultät.
- DGV 1987–2005. Luis Michelena and Ibon Sarasola, *Diccionario general vasco/Orotarik o euskal hiztegia*, vols. 1–16.

- Bilbao: Real Academia de la Lengua Vasca/Euskaltzaindia.
- DOE 1986-. Angus Cameron, Ashley C. Amos, and Antoinette diPaolo Healey (eds.), *Dictionary of Old English*. Toronto: University of Toronto, Centre for Medieval Studies. (To date A–H.)
- Eliasson, Stig 2006. Old Danish vigesimal counting: a comparison with Basque. In Antje Hornscheidt, Kristina Kotcheva, Tomas Milosch, and Michael Riessler (eds.), *Grenzgänger: Festschrift zum 65. Geburtstag von Jurij Kusmenko* (Berliner Beiträge zur Skandinavistik 9). Berlin: Nordeuropa-Institut der Humboldt-Universität, pp. 91–110.
- Entwistle, William J. 1936. *The Spanish Language, Together with Portuguese, Catalan and Basque*. London: Faber.
- Etxepare, Ricardo 2003. Valency and argument structure in the Basque verb. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.), *A Grammar of Basque* (Mouton Grammar Library 26). Berlin: Mouton de Gruyter, pp. 363–426.
- Feist, Sigmund 1913. *Kultur, Ausbreitung und Herkunft der Indogermanen*. Berlin: Weidmann.
- Freche, Katharina 1995. Zur Methodik der Ortsnamendeutung: Konkurrierende Erklärungen ausgewählter Fluß- und Siedlungsnamen des Kreises Göppingen. *Sprachwissenschaft* 20: 396–419.
- Garvens, Fritz 1964. Die vorrömische Toponymie Nordspaniens. Dissertation, University of Münster/Westphalia.
- Goebl, Hans, Peter H. Nelde, Zdeněk Starý, and Wolfgang Wölck (eds.) 1997–8. *Contact Linguistics: An International Handbook of Contemporary Research* (Handbooks of Linguistics and Communicative Science 12.1, 12.2), 2 vols. Berlin: Walter de Gruyter.
- Greene, David 1992. Celtic. In Jadranka Gvozdanović (ed.), *Indo-European Numerals* (Trends in Linguistics 57). Berlin: Mouton de Gruyter, pp. 496–554.
- Gvozdanović, Jadranka (ed.) 1992. *Indo-European Numerals* (Trends in Linguistics 57). Berlin: Mouton de Gruyter.
- Hickey, Raymond 2007. *Irish English: History and Present-day Forms* (Studies in English Language). Cambridge: Cambridge University Press.
- Hualde, José Ignacio 1999. Basque accentuation. In Harry van der Hulst (ed.), *Word-Prosodic Systems in the Languages of Europe* (Empirical Approaches to Language Typology, EUROTyp 20–4). Berlin: Walter de Gruyter, pp. 947–93.
- Hualde, José Ignacio 2003. Accentuation. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.), *A Grammar of Basque* (Mouton Grammar Library 26). Berlin: Mouton de Gruyter, pp. 65–72.
- Hualde, José Ignacio 2007. Historical convergence and divergence in Basque accentuation. In Tomas Riad and Carlos Gussenhoven (eds.), *Tones and Tunes*, vol. 1: *Typological Studies in Word and Sentence Prosody* (Phonology and Phonetics 12.1). Berlin: Mouton de Gruyter, pp. 291–322.
- Hualde, José Ignacio and Jon Ortiz de Urbina (eds.) 2003. *A Grammar of Basque* (Mouton Grammar Library 26). Berlin: Mouton de Gruyter.
- Hyman, Larry M. 1977. On the nature of linguistic stress. In Larry M. Hyman (ed.), *Studies in Stress and Accent* (Southern California Occasional Papers in Linguistics 4). Los Angeles: University of Southern California, pp. 37–82.
- Jost, Karl 1909. *BEON und WESAN: Eine syntaktische Untersuchung* (Anglistische Forschungen 26). Heidelberg: Carl Winter.
- Keller, Wolfgang 1925. Keltisches im englischen Verbum. In *Anglica: Untersuchungen zur englischen Philologie* (= Festschrift für Aloys Brandl), vol. 1. Leipzig: Mayer and Müller, pp. 55–66.

- King, Alan R. 1994. *The Basque Language: A Practical Introduction*. Reno: University of Nevada Press.
- Koch, Peter and Wulf Oesterreicher 1996. Sprachwandel und expressive Mündlichkeit. *Zeitschrift für Literaturwissenschaft und Linguistik* 102: 64–96.
- Krahe, Hans 1954. *Sprache und Vorzeit: Europäische Vorgeschichte nach dem Zeugnis der Sprache*. Heidelberg: Quelle and Meyer.
- Krahe, Hans 1964. *Unsere ältesten Flußnamen*. Wiesbaden: Otto Harrassowitz.
- Krahe, Hans and Wolfgang Meid 1969. *Germanische Sprachwissenschaft I, II*, 7th edn. (Sammlung Götschen 2232–3.) Berlin: Walter de Gruyter. Additional volume III: *Wortbildungslehre*, by Wolfgang Meid, 1967.
- Lafitte, Pierre 2001. *Grammaire basque (navarro-labourdin littéraire)*. Donostia and Bayonne: Elkarlanean.
- Lewis, Henry and Holger Pedersen 1989. *A Concise Comparative Celtic Grammar*, 3rd edn., 2nd impression with the supplement of 1961 by Henry Lewis. Göttingen: Vandenhoeck and Ruprecht.
- Lutz, Angelika 2009. Celtic influence on Old English and West Germanic. In Markku Filppula and Juhani Klemola (eds.), *Re-Evaluating the Celtic Hypothesis*. Special issue of *English Language and Linguistics*, 13: 227–49.
- Maienborn, Claudia 2005. A discourse-based account of Spanish *ser/estar*. *Linguistics* 43: 155–80.
- Maienborn, Claudia 2007. Review of Regina Pustet: *Copulas: Universals in the Categorization of the Lexicon* (Oxford University Press, 2003). *Studies in Language* 31: 232–44.
- Martinet, André 1964. *Économie des changements phonétiques: Traité de phonologie diachronique*. Berne: A. Francke.
- MED 1954–2001. *Middle English Dictionary*, ed. Hans Kurath and Sherman M. Kuhn, 12 vols. Ann Arbor: Michigan University Press.
- Meillet, Antoine 1964. *Introduction à l'étude comparative des langues indo-européennes*. University, AL: University of Alabama Press.
- Menninger, Karl 1992. *Number Words and Number Symbols: A Cultural History of Numbers*, trans. Paul Broneer from the revised (1957–8) German edn. New York: Dover Publications.
- Mercado, Angelo O. 2007. Comparative-historical metrical problem(s): Italic and Celtic. Paper presented at the 26th East Coast Indo-European Conference. New Haven, Connecticut, 16 June.
- Michelena, Luis 1977. *Fonética histórica vasca*. San Sebastián: Diputación Foral de Gipuzkoa.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Nyrop, Kristoffer 1960. *Grammaire historique de la langue française II*, 2nd edn. Copenhagen: Gyldendalske Boghandel.
- OED 1989. *The Oxford English Dictionary*, 2nd edn., prepared by Jane A. Simpson and Edmund S. C. Weiner, 20 vols. Oxford: Clarendon.
- Onions, Charles T. 1969. *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press.
- Palay, Simon 1961. *Dictionnaire de béarnais et du gascon modernes*, 2nd edn. Paris: CNRS.
- Pedersen, Holger 1976. *Vergleichende Grammatik der keltischen Sprachen*, 2 vols. 2nd, unchanged edn. Göttingen: Vandenhoeck and Ruprecht.
- Pfaller, Angelika 2003. Sprachwandel im Spanischen. Magisterarbeit, Universität München, Institut für Romanische Philologie.
- Posner, Rebecca 1996. *The Romance Languages* (Cambridge Language Surveys). Cambridge: Cambridge University Press.

- Price, Glanville 1992. Romance. In Jadranka Gvozdanović (ed.), *Indo-European Numerals* (Trends in Linguistics 57). Berlin: Mouton de Gruyter, pp. 447–96.
- Pustet, Regina 2003. *Copulas: Universals in the Categorization of the Lexicon*. Oxford: Oxford University Press.
- Röder, Katrin 2000. *Struktur und Verbreitung der alteuropäischen Toponymie: Eine Studie am Beispiel der Wurzelformen \*is- und \*ur-*. Berlin: Logos.
- Rohlf, Gerhard 1966–9. *Grammatica storica della lingua italiana e dei suoi dialetti*, 3 vols. Turin: Einaudi.
- Rohlf, Gerhard 1971. *Romanische Sprachgeographie: Geschichte und Grundlagen, Aspekte und Probleme, mit dem Versuch eines Sprachatlas der romanischen Sprachen* (Handbücher für das Studium der Romanistik). Munich: C. H. Beck.
- Ross, Alan S. C. and Jan Berns 1992. Germanic. In Jadranka Gvozdanović (ed.), *Indo-European Numerals* (Trends in Linguistics 57). Berlin: Mouton de Gruyter, pp. 555–715.
- Sagüés Subijana, Miguel 1994. *Gramática elemental vasca: Gramática comparada*, 8th edn. San Sebastian: Editorial Txertoa.
- Salmons, Joseph C. 1992. *Accentual Change and Language Contact: Comparative Survey and a Case Study of Early Northern Europe*. London: Routledge.
- Schrijver, Peter 1995. *Studies in British Celtic Historical Phonology* (Leiden Studies in Indo-European 5). Amsterdam: Rodopi.
- Schumacher, Stefan 2007. Die Deutschen und die Nachbarstämme: Lexikalische und strukturelle Sprachkontaktphänomene entlang der keltisch-germanischen Übergangszone. In Hans Hablitzel and David Stifter (eds.), *Johann Kaspar Zeuß im kultur- und sprachwissenschaftlichen Kontext (19. bis 21. Jahrhundert)*, Kronach 21.7.–23.7. 2006. Wien: Praesens Verlag, pp. 167–207.
- Sihler, Andrew L. 1995. *New Comparative Grammar of Greek and Latin*. New York: Oxford University Press.
- Tagliavini, Carlo 1998. *Einführung in die romanische Philologie* (transl. of *Le origini delle lingue neolatine: Introduzione alla filologia romanza*), 2nd, improved edn. Tübingen: A. Francke.
- Thomason, Sarah G. 2001. *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. and Terrence Kaufman 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thurneysen, Rudolf 1946. *A Grammar of Old Irish*. Dublin: Dublin Institute for Advanced Studies.
- Torroni, Antonio, et al. 1998. mtDNA analysis reveals a major late paleolithic population expansion from southwestern to northeastern Europe. *American Journal of Human Genetics* 62: 1137–52.
- Trask, R. Larry 1997. *The History of Basque*. London: Routledge.
- Trask, R. Larry 2003. The noun phrase: nouns, determiners, and modifiers; pronouns and names. In José Ignacio Hualde and Jon Ortiz de Urbina (eds.), *A Grammar of Basque* (Mouton Grammar Library, 26). Berlin: Mouton de Gruyter, pp. 113–70.
- Vennemann, Theo 1994. Linguistic reconstruction in the context of European prehistory. *Transactions of the Philological Society* 92: 215–84.
- Vennemann, Theo 1996. Zur Frage der vorindogermanischen Substrate in Mittel- und Westeuropa. Paper presented at a colloquium held in honor of Johannes Hubschmid, University of Heidelberg, 18 January.
- Vennemann, Theo 1998. Basken, Semiten, Indogermanen: Urheimatfragen in linguistischer und anthropologischer Sicht. In Wolfgang Meid (ed.), *Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen*

- Gesellschaft, Innsbruck, 22.–28. September 1996* (Innsbrucker Beiträge zur Sprachwissenschaft 93). Innsbruck: Universität Innsbruck, Institut für Sprachwissenschaft, pp. 119–38.
- Vennemann, Theo 1999. Remarks on some British place-names. In Gerald F. Carr, Wayne Harbert, and Lihua Zhang (eds.), *Interdigitations: Essays for Irmengard Rauch*. Bern: Peter Lang, pp. 25–62.
- Vennemann, Theo 2002. Semitic → Celtic → English: the transitivity of language contact. In Markku Filppula, Juhani Klemola, and Heli Pitkänen (eds.), *The Celtic Roots of English* (Studies in Languages 37). Joensuu: University of Joensuu, Faculty of Humanities, pp. 295–330.
- Vennemann, Theo 2003. *Europa Vasconica – Europa Semitica*, ed. by Patrizia Noel Aziz Hanna (Trends in Linguistics: Studies and Monographs 138). Berlin: Mouton de Gruyter.
- Vennemann, Theo 2006a. Ortsnamen von Personennamen? Das letzte Beispiel. In Ursula Götz and Stefanie Stricker (eds.), *Neue Perspektiven der Sprachgeschichte: Internationales Kolloquium des Zentrums für Mittelalterstudien der Otto-Friedrich-Universität Bamberg, 11. und 12. Februar 2005*. Heidelberg: Winter, pp. 269–83.
- Vennemann, Theo 2006b. “A satisfactory etymology has long been available”: notes on Vasconic names outside the Basque country: with particular reference to some British *Arn-* and *Earn-* names and to German *Arnoldsweiler*. In Jose Ignacio Hualde and Joseba Lakarra (eds.), *Studies in Basque and historical linguistics in memory of R. L. Trask* (Anuario del Seminario de Filología Vasca “Julio de Urquijo” 40.1–2). Donostia/San Sebastián: Diputación Foral de Gipuzkoa, pp. 969–92.
- Vennemann, Theo 2008. Zur Etymologie von *Busch* und *boscus* sowie zu *Bus-* und *Bos-* in der Alteuropäischen Toponymie: Mit einem Appendix zur Etymologie von *Büste*. *Sprachwissenschaft* 33: 107–38.
- von Wartburg, Walter 1922–9. *Französisches etymologisches Wörterbuch*. Bonn: Klopp.
- Weinreich, Uriel 1953. *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York.
- Welscher, Iva 2005. *Mitteleuropa und Südosteuropa im Kontext der Alteuropäischen Toponymie*. Diss., Fakultät für Sprach- und Literaturwissenschaften, Universität München.
- Winford, Donald 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.