

Etymology and phonotactics: Latin *grandis* vs. Basque *handi* 'big' and similar problems¹

Theo Vennemann gen. Nierfeld
University of Munich

For Hans-Heinrich Lieb

1. The etymological problem of Lat. *grandis* and Bq. *handi* 'big'

For the Latin *s*-stem adjective *grandi-* (nom. sg. m./f. *grandis*, n. *grande*; Fr. *grand/grande*, Span. *grande*, etc.) 'big', according to one etymological Latin dictionary (Walde/Hofmann 1982: s.v.), several Indo-European etymologies, both ancient and modern, have been proposed. Since these cannot all be true, there is the possibility that they are all false. Indeed the fact that the word only occurs in Latin does not make an Indo-European origin appear likely. Also the structure of the word, especially its root vowel *a* (Proto-Indo-European had few if any words with short *a*), does not suggest an inherited Indo-European word. The authors of another etymological Latin dictionary (Ernout/Meillet 1959: s.v.), after ascribing an ancient association of the word with Lat. *gradus* 'step' to popular etymology, abstain from offering any of the modern

¹Most of the material published here was first presented and discussed as part of my contribution "The influence of Vasconic at the time of its greatest expansion upon Latin and other Indo-European languages" to the First International Conference on Linguistic Contact, Valencia, 19-22 September 1995. It is offered as a birthday gift to Hans-Heinrich Lieb (Freie Universität Berlin) who has, from the beginning, observed my efforts to reconstruct linguistic aspects of pre-Indo-European Europe with sympathy and advice. — As in previous publications I mark reconstructed forms by a raised cross (†) and incorrect forms by an asterisk (*). In cited material I respect the practice of the authors. Philip Baldi deserves thanks for reading a manuscript version of this paper and for asking important questions, which I have tried to answer. Edgar C. Polomé has contributed to my reconstructive efforts by expressing, in personal letters and in general terms, doubt about my thesis that Vasconic languages had an important role in the shaping of the linguistic landscape of early Europe; I certainly would like to demonstrate that I am not "betting on the wrong horse" (to use his words), quite apart from the fact that mine at present appears to be the only horse on the turf.

etymologies and succinctly conclude, "L'étymologie de ce mot 'vulgaire' à vocalisme *a* est inconnue. Le mot indo-européen signifiant 'grand' est représenté en latin par *magnus*."

For Bq. *handi* 'big' too, according to Agud and Tovar's etymological Basque dictionary (1989: s.v.), several etymologies have been proposed, apparently none of them convincing to these authors.² The word occurs as *andi* in the Bizkaian, Gipuzkoan, High Navarrese, and Roncalese dialects, as *handi* in the Lapurdian, Low Navarrese, and Zuberoan dialects.³ The form with aspiration appears to be etymologically correct, whereas the form *andi* reflects loss of aspiration in the deaspirating dialects.⁴ The forms *aundi* (Bizkaian, Gipuzkoan) and *haundi* (Lapurdian, Low Navarrese), considered as basic by some authors, are emphatic variants to other authors including Agud and Tovar (1989) who follow Michelena (1977: 96).

The phonological similarity and the semantic identity of Lat. *grandi-* 'big' and Bq. *handi* 'big' have not remained unnoticed, but no way was known in the past to connect the two items which, after all, belong to two genetically unrelated languages. Assuming for the sake of the argument that the two items might be connected, the only way of connecting them was the assumption that the word passed from Latin, or from French or Spanish, into Basque, as is true for innumerable other words that are the same in Basque and in Romance. The converse way of borrowing could not be assumed. The reason is simple: *grandi-* is found in Latin as early as Plautus (cf. Walde/Hofmann 1982: s.v.); Plautus died in 184 B.C., whereas contact of the Romans with speakers of Basque or related languages only began in the following century. Thus the assumption could only be that Bq. *handi* 'big' might be a Latin loan-word, Lat. *grandis/grande* → Bq. *handi* (cf. Dodgson 1903: 326 n. 1).⁵ However, this proposal could not be accepted

²One etymology which Agud/Tovar (1989) do not even mention is Löpelmann's (1968) derivation from Lat. *tanti* 'so many'.

³I use the dialect names in the same form as Trask (1997: 5); they are derived from the Basque names of the provinces.

⁴Since there is also non-etymological *h-* in Basque, this point is important; cf. Gorrochategui 1995: 42 n. 14.

⁵The idea is offered as a question added to an occurrence of *handia* in the text under investigation: "Du latin *grande* avex élision du *ē*: ou bien de la racine *han, haut* — *gonflē*:" Bq. *andu* (also *(h)aunditu*) 'to swell' is, on the contrary, considered to be derived from *(h)andi* 'big' by Agud/Tovar (1989:

(Agud/Tovar 1989: s.v.). Indeed, if Basque had borrowed Lat. *grandis/grande* (or Fr. *grand/grande* or Span. *grande*), the result would have to be **garand-*, e.g. **garandi*, because Basque breaks up initial *Cr-* clusters in loan-words by means of echo anaptyxis, cf. GRATIA → *garazia* 'grace', GRANU → *garau(n)* 'grain', CRUCE → *kurutze* 'cross'.⁶

In a number of recent articles (Vennemann 1993a, 1994a [appendix], 1995a) I have proposed a theory of the linguistic development of prehistoric Europe according to which Vasconic languages, i.e. languages related to historical Basque which later became extinct, were in prehistoric times contact languages of the West Indo-European languages, in particular the Italic languages before their speakers migrated into Italy, and probably even in Italy itself. One kind of evidence to support this theory consists in a number of Vasconic loan-words in West Indo-European, including Latin. This theory offers the basis for a new interpretation of the relationship between Lat. *grandi-* and Bq. *handi*, namely as a Vasconic loan-word in Latin. I realize that this is, even given the new theory, not a trivial assumption, because the way one would naturally capture this relationship, "Bq. *handi* → Lat. *grandis/grande*" (cf. above "Lat. *grandis/grande* → Bq. *handi*"), does not seem to fit into known patterns of loan-word adaptation. And indeed it does not. I will show on the following pages why it is nevertheless possible to make the etymological connection in the only remaining direction, from Vasconic to Latin.

2. Phonotactic constraints and loan-transfer: Some examples

2.1. Segmental change: OIr. *fás* etc., Bq. *baso* 'waste (adj.)'

It is a well-known methodological rule that in loan-word comparison one has to take into account the systems of the involved languages at the time of the loan-transfer, as well as the changes that have taken place since the time of the loan-transfer both in the giving and in the receiving languages. E.g., in order to understand the loan-transfer Lat. *cella* → OHG *zella* /tʰella/ (NHG *Zelle* 'cell'), one has to consider that /k/

s.v. *andu*²).

⁶This was, of course, the reason why Dodgson assumed an intermediate change of Lat. *grande* into **gande*, for which there exists, however, no parallel or other evidence.

became /tʰ/ before front vowels in Vulgar Latin. And in order to understand the loan-transfer Lat. *tegula* → OHG *ziagal* /tʰiagal/ (NHG *Ziegel* 'tile'), one has to consider that /t/ became /tʰ/ word-initially before vowels in pre-Old High German.

Keeping this methodological rule in mind, even regard to simple segmental changes like the above can prove fruitful in the comparative study of West Indo-European and Basque. I would like to show its fruitfulness with just one example of an unexplained parallel between the two families.

If one looks at OIr. *fás* 'empty, waste (adj.)' (with *fásach* 'desert, wilderness')⁷ and Bq. *basa* (also *baso*⁸) 'wild, waste (adj.)', the similarity of sound and meaning suggests that they are the same word. Yet the phonological differences have to be explained.

First, there is a prosodic difference: The root vowel is long in Old Irish; in Basque there is no quantity, so the root vowel is neutral with regard to length. This difference tells us nothing about the transfer direction because when an Indo-European word is borrowed into Basque it necessarily loses its vowel quantity by neutralization, and when a Vasconic word is borrowed into an old Indo-European language quantity has to be assigned. In particular, the Basque root vowel in an open syllable may have appeared long to speakers of a quantity language.

Second, there is an initial *f* in Old Irish and an initial *b* in Basque. These cannot simply be transformed into each other. However, it is known that OIr. *f* derives from PIE **w-* by onset strengthening; cf. Lat. *vir*, Goth. *wair*, also, with long *i*, Skt. *vīrah*, Lith. *vīras*; but OIr. *fer* 'man'. It is also known that Basque, lacking *w*, has neutralized *w* with *b* word-initially before vowels in favor of *b* whenever words entered from a language making a distinction, e.g. Lat. *valōrem* → Bq. *balhore* 'valor, courage'.⁹ Thus, either a pre-Old Irish root **wās-* 'empty, waste

⁷Kluge/Seebold 1995: s.v. *wüst*, Pfeifer 1989: s.v. *wüst*

⁸There is also a noun *baso* 'forest'. Whether this is the same word or not is uncertain. Cf. Agud/Tovar 1991: s.v. *baso*², *baso*¹. Finally, there are *baso* 'mud' (Agud/Tovar 1991: s.v. *baso*¹) and *baso* 'bubbles on the surface of water' (Agud/Tovar 1991: s.v. *baso*²).

⁹Löpelmann (1968: s.v.). According to Agud/Tovar (1991: s.v.), Bq. *balhore* comes from Castilian *valor*. As for Lat. initial *w-*, it is always difficult to decide whether the change into *b* first occurred in Basque or in Romance, because

(adj.)¹⁰ or a Vasconic root **was-* 'wild, waste (adj.)' would explain the Old Irish word and the Basque word as the same item, namely as possible source of a loan-transfer.

Engl. *waste*, both adjective and noun, is according to Onions (1966: s.v.) a borrowing from Old Northern French, where *wast* was a variant of *g(u)ast*, cf. Prov. *gast*, Port. *gasto*, Ital. *guasto*, all from Romance **wasto*, which apparently continues Lat. *vāstus* 'empty, unoccupied, waste, desert, devastated', whence *vāstāre* 'to make empty, lay waste' and other derivatives (Walde/Hofmann 1982: s.v. *vāstus*). This in turn is related to OE *wēste*, OS *wōsti*, OHG *wuosti* (NHG *wüst*), Pre-West Gmc. **wōsti-* 'empty, unoccupied, waste', further OE *wēsten*, OS *wōstunna*, OHG *wuostinna*, also *wuosti* (NHG *Wüste*) 'desert (noun)' (Walde/Hofmann 1982: s.v. *vāstus*, Kluge/Seebold 1995: s.v. *wüst*).

Lat. *vāstus* and Pre-West Gmc. **wōsti-* are in turn both related to OIr. *fás*. This is unproblematical, because OIr. intervocalic *s* is a regular reflex of *ss*, which in turn is a regular reflex of *st* (Thurneysen 1946: §§ 131, 155.e, 194, 217). Thus *fás* (also *fáss*) is likely to be earlier **wāstV-*. Since Pre-West Gmc. **wōsti-* also goes back to Pre-Gmc. **wāsti-*, the reconstruction yields **wāstV-* for all three language families, though Latin and West Germanic differ in the stem-forming vowel of the adjective. To Walde/Hofmann (1982) and Kluge/Seebold (1995) this is the end of the reconstruction. The former consider a possible relationship of Lat. *vāstus* with Lat. *vastus* 'wide, immensely large, unshapely, bulky' but recommend to keep the two items distinct. The latter simply state, "Weitere Herkunft unklar."

Ernout/Meillet (1959: s.v. *uastus*) only consider a single item *vastus*, with short *a*, and covering both meanings, or groups of meanings. They essentially compare the same Old Irish and West Germanic words as above and emphasize the primacy of the adjectival meaning. Isolating a root **wās-* [sic, with long *ā*], and thus implicitly reconstructing an adjectival base **wās-t-*, they compare Lat. *vānus* (< **wās-n-os*) 'empty, thinned-out, vane' and *vascus* (glossed 'inānis') 'empty, void, deserted, unoccupied'. They do not comment on their reconstruction, e.g. by further stabilizing or identifying their reconstructed root **wās-*.

on this point Basque and, e.g., Castilian behave identically.

¹⁰Also the final *-s* of the Old Irish word requires analysis; see below.

In none of the etymological dictionaries cited is there any discussion of the fact that the group of Indo-European words assembled in them and cited here is restricted to three West Indo-European languages which were in close contact in prehistoric times and share many innovations (Krahe/Meid 1969: §§ 4, 6, 7, 9), and that in Germanic the word is restricted to West Germanic. In my opinion these two facts suggest that we are dealing with a loan complex. Since I believe to have shown for a number of such shared innovations that they can be explained as Vasconic influence, the West Indo-European root **wās-is*, in the context of my theory, most likely a Vasconic loan item. I therefore propose that in the case on hand it was words based on a Proto-Vasconic root **was-* 'wild, waste' surviving in Bq. *basa/baso*, which were borrowed into Italic, Celtic, and West Germanic where they became the basis of adjectival derivation yielding the attested West Indo-European words.¹¹

2.2. Phonotactic segmental change

2.2.1. Engl. *rye*, Bq. *ogi* 'bread, wheat, cereal'

Kluge/Seebold (1995: s.v. *Roggen*) reconstruct for OS, OHG *roggo* PGmc. **rugōn* 'rye' and for OE *ryge* (Engl. *rye*), ON *rigr* a doublet PGmc. **rugi-* 'rye'. Lühr (1988: 291) considers the *i*-stem original and the *n*-stem an innovation. The geminate in OS, OHG *roggo*, according to Lühr, is unexplained; it is attributed to forms with zero grade by Kluge/Seebold, as also required for OFris. *rogga*. They compare Lith. *rugys* 'grain of rye', in the plural 'rye', Old Russ. *rūzi* f. 'rye', and ("vielleicht") Thrac. *brīza* 'rye'. "Weitere Herkunft unklar; die lautliche Nähe von gr. *oryza* f. 'Reis' ... ist auffällig." Finn. *ruis* 'rye' and Eston. *rukkis* 'rye' are Germanic or Balto-Slavic loan-words. The limited Indo-European distribution of the word, the peculiarities of its stem formation, and the lack of consonant shifting suggest that the Indo-European item too is a loan complex.

¹¹I would like to emphasise again that the quantity of the root vowel was itself not borrowed but *assigned*, possibly on phonetic grounds, such as subphonemic lengthening in a stressed open syllable. Such assignment was necessary because vowels in the early Indo-European languages had to be phonologically long or short, whereas in Vasconic phonologically they were neither (at least if we trust the evidence of Basque), even though phonetically there probably was a good deal of variation. Alternative phonological length assignments has even led to doublets, as I have shown with toponymic examples (1994b: 242 and n. 43, 1995a: 48f.).

Bq. *ogi* means 'wheat, cereal' (Span. 'trigo', Fr. 'blé') in the High Navarrese, Low Navarrese, Lapurdian, and Zuberoan dialects and 'bread' everywhere (de Azkue 1984: s.v.).¹²

The sounds and meanings of IE **rugi-* 'rye' and Bq. *ogi* 'bread, wheat, cereal' are close enough to suggest the possibility of a connection. The anlauts pose an apparent problem; it disappears when a specifically Basque phonotactic constraint is considered, namely the prohibition against word-initial *r*:

**r* / # _____

There are no old words and few popular loan-words with initial *r* in Basque; usually loan-words with initial *r* receive a prosthetic vowel (Michelena 1977: § 8.1):

Lat. *rege(m)* 'king (acc.)' → Bq. *errege* 'king, queen-bee'

Span. *resto*, Fr. *reste* 'rest' → Bq. *arrasta*, *arresta* (id.)

Span. *ropa* 'dress, clothes' → Bq. *arropa* (id.)

Fr. *révolte* 'revolt' → Bq. *errebolta* (id.)

Cat. *risc*, Fr. *risque*, Sp. *riesgo*¹³ 'risk' → Bq. *arrisku* (id.)

Span. *ruta*, Fr. *route* → Bq. *arruta* (id.)

#*rV₁X* → Bq. *V₂rrV₁X*, where *V₂* is (usually) *a*, except when *V₁* is *e*, in which case *V₂* is (usually) *e*; and *X*' is a Basque adaptation of *X*¹⁴

Most languages containing *r*-sounds do not bar them from the word-initial position. There is therefore no reason to assume that Basque in its prehistory had always been subject to this constraint. Assuming on the contrary that at an earlier stage Basque allowed initial *r*, the question arises of how the modern constraint was implemented. Since Basque has a notorious prehistory and early history of initial consonant weakening and loss, I propose that word-initial *r* too was weakened and lost in this language, and indeed before it came into contact with

¹²Cf. Leizaola et al. (1990: no. 389) for the distribution of *ogi* as a name for the cereal and its competition with *gari*.

¹³Also Span. *arisco*.

¹⁴The prosthetic vowels also differ considerably in the various dialects.

Latin and its successor languages. A possible path of development could be the following (with unvoicing of initial *r*, *r* > *r̥*, as in Classical Attic Greek, cf. Allen 1974: 39f.):

$+r > +r̥ > (+)h > \text{zero}$

If we reconstruct PVasc. $+rugi-$ 'wheat', we can account for both Bq. *ogi* 'wheat, cereal, bread' and (West) IE $+rugi-$ 'rye': The former would continue the Proto-Vasconic word, with typical generalizing and metonymic changes of meaning and with mandatory loss of its initial $+r̥$; vowel height is not a stable feature of Basque (Michelena 1977: § 2.1) and may in the present instance have been influenced by the lowering effect of the weakening rhotic.¹⁵ The latter would be a Vasconic loan-word, with a transference of meaning made possible by the fact that wheat and rye are similar in both appearance and use, plus the late introduction and increasing importance of rye in Central Europe.¹⁶ In view of the fact that neither wheat nor rye have their original home in Europe, it is conceivable that the Vasconic word too is a loan-word, perhaps from a Mediterranean language.¹⁷

Finally, there is the possibility that the $+rugi-$ word was originally used for the type of cereal represented by both wheat and rye. The fact that wheat has in Germanic no real name of its own but is merely referred to by a distinguishing property as 'the white one' points in this direction.¹⁸

Connecting PGmc. $+rugi-$ 'rye' and Bq. *ogi* 'wheat, cereal, bread' is only possible against the background of a theory of language change which provides for phonotactic change, such as the rise of a constraint barring rhotics from the word-initial position in a language which earlier on was not constrained in

¹⁵Perhaps an original PVasc. $+rugi-$ would also yield the attested forms. At least for Germanic this would be so, especially under the assumption that this is a late word (see directly), borrowed at a time when Pre-Gmc. $+o$ had already merged with $+a$.

¹⁶Cf. the description, including historical accounts, and the pictures in Körber-Grohne 1987: 28-46. Wheat and rye are the chief cereals used for baking bread.

¹⁷Löpelmann (1968: s.v. *ogi*) compares Egypt. *qw* 'bread, food', Copt. *oerik*, *ōik* 'bread'. He does not connect the *rye* word.

¹⁸Gmc. $+hwaitja-$ 'wheat' and related formations are Germanic innovations based on ablaut variants of Gmc. $+hwaita-$ 'white' (Onions 1966: s.v. *wheat*, Kluge/Seebold 1995: s.v. *Weizen*).

this way. Assuming such a theory, however, forces the etymologist to consider the possibility of etyma that are less constrained in certain regards than the word under investigation. Latinists reckon with this possibility when trying to reconstruct vowel-initial words of late attestation: Since in Classical Latin *h*, and thus in particular *h-*, was no longer sounded, sometimes the question arises whether a given word of this kind is to be reconstructed with the initial vowel or with initial ⁺*h-* (and thus with ⁺*gh-*).¹⁹ By the same token, with regard to Basque every word now beginning with a vowel has to be checked against the possibility of prehistoric onset deletion.

2.2.2 Gk. *dorkás*, *dórx*, Corn. *yorch*, Bq. *orkatz* 'deer'

Gk. *dorkás*, Gen. *dorkádos* means 'deer, gazelle'. There exist an amazing number of morphological variants, all with the same meaning, including one which gives the appearance of a root noun, *dórx*, acc. *dórkan* or *dórkān*, from which all other variants are derived (see Frisk 1973 and Chantraine 1990, s.v. *dorkás*). But Chantraine points out that *dorkás* is the usual form, the only one from which diminutives, adjectives, and verbs are derived. All this together causes the impression of a loan-word, and that impression is strengthened by the fact that there is no **dork-* 'deer, gazelle' in the other Indo-European languages.

The word is peculiar in yet another regard. It has a variant *zorkás* (also *zórax*) with the same meaning. The two groups cannot be connected phonologically.

There is yet another group of variants: *iórkos* (only one author), *iórkes*, *íyrkes* (only Hesychius).

The forms beginning with *i-* have a counterpart in one Indo-European branch, Celtic: Corn. *yorch*, Bret. *iourc'h* 'deer', Cymr. *iwrch* 'caprea mas'. Both Frisk (1973) and Chantraine (1990) therefore reconstruct PIE **ior̥k-o-*, as does Pokorny (1989: 513). In this way they account for the Celtic forms and the Greek *z-* onset.²⁰ They go on to explain the Greek *d-* onset

¹⁹Neglecting this possibility may easily disconnect words that are in all probability identical or closely related, such as Lat. *ardea* 'heron' (attested since Virgil, cf. Walde/Hofmann 1982: s.v.), Span./Port. *garza* (id.), Gk. *charadriós* (a grallatorial bird). Failure to reconstruct ⁺*hardea* isolates *ardea* (cf. also Ernout/Meillet 1959: s.v.) and also separates the name of this eminently grallatorial bird from the root of *gradus* 'step', *gradi* 'to take steps, walk as on stilts' (hence *grallae* 'stilts'). Cf. Vennemann 1996: 124f.

²⁰It is not quite obvious that they do, because the normal reflex of PIE ⁺*i-* in

(*zorkás* → *dorkás* etc.) by association with *dérkomai* 'I see' in an act of popular etymology (on the part of the Pre-Greek speakers). This leaves the Greek *i*-forms. They are explained as Celtic loan-words.

I must admit that I find this unanimous account in the etymological dictionaries far from convincing.²¹ I also find it incomplete (and consider it likely that it cannot be completed within this standard dictionary reconstruction): The anomalous multiplicity of Greek forms is not accounted for. I also reserve doubts about the Indo-European character of a word that occurs only in two branches of Indo-European, both belonging to West Indo-European, at least in a wider sense. I therefore find confirmed my original impression that we are dealing with a loan complex.

The theory within which I am working suggests that we are dealing, for both Greek and Celtic, with loan-words from Vasconic languages.²² Therefore we have to look into Basque. Under 'deer', Aulestia/White (1990) offer *orein*, *orkatz*, and *basahuntz*. So does Trask (1997: 298), who adds to *orkatz* "(also 'Pyrenean chamois')" and to *basahuntz* "(*basa*- 'wild' + *ahuntz* 'goat')". I will return to *ahuntz* in section 2.4 below. Bq. *orein* 'deer' is an interesting word in its own right as it may be connected with Herodotus's *órys*, gen. *óryos* 'a Lybian animal, perhaps an antelope' (Liddell/Scott 1889: s.v.), or with Gk. *óryx*, gen. *órygos*, supposedly the loan name of an Egyptian and

Greek is *h*-. But there is Gk. *zugón* alongside Lat. *iugum* etc. for which a laryngeal solution is available (cf. Rix 1976: § 68). Therefore they probably would have to reconstruct PIE **H₁ork-o*, with a laryngeal consonant *H*.

²¹It must be said in fairness that the authors of all three dictionaries cited employ a certain amount of modalization ("können ... sein", "vermutlich", "probable", "probablement", "apparemment", "on a supposé").

²²It is, of course, possible that all three languages, Greek, Celtic, and Vasconic, borrowed the word from a fourth. This is what Löpelmann (1968: s.v. *okhats*) suggests. He assumes the word to have been borrowed from an "Aegean" language and ultimately via Semitic (Akkad. *turāhu* 'steinbok, capra ibex') from Sum. *durāḥ* (id.). However, since no "Aegean" language with a word **dorkats* 'deer' seems to be extant, and since a scenario with only three languages is available and has worked in many other instances, the account given above is preferable on methodological grounds (Occam's razor, see also Vennemann 1995a: 40). This account is compatible with the suggestion that the Proto-Vasconic word itself is a Semitic loan-word. Since Akkad. *turāhu* (also once *tarāhu*) has the cognates Syr. *tarrāhā* / *tarrūhā*, Arab. *ta'ruh* with the same meaning (von Soden 1981: s.v.), the word is most likely of Semitic origin, and Sum. *durāḥ* an Akkadian loan-word.

Lybian (also Indian) antelope species (cf. Frisk 1973 and Chantraine 1984, s.v. *orýssō* 'I dig', Löpelmann 1968: s.v. *orein*²³). But *orkatz* 'deer, Pyrenean chamois' strikes me as a close match of Gk. *dorkás*, *dorkádos* 'deer' and, e.g., Corn. *yorch* 'deer', both phonologically and semantically. Since the item is in all likelihood not Indo-European, I assume it to be Vasconic.

Proto-Vasc. ⁺*dorkats* would lose its initial ⁺*d* in the prehistory of Basque, which at an early level of its development did not permit word-initial *d*- (Trask 1997: 126). When borrowed as ⁺*dorkats* 'deer' into Pre-Greek, it would naturally be adapted into the morphology as *dorkás*, *dorkádos* 'deer'. This word in turn may have formed the basis for the back-formation of the root noun *dórx* 'deer'; together, *dorkás* and *dórx* suffice to derive all the other related forms on record. However, *dórx* may also go back to some syncopated variant form ⁺*dorks* in the local varieties of Vasconic, and similar variation may have led directly to the other variants of the Greek 'deer' word with initial *d*-.

The Greek variants with initial *z*- and *i*- for *d*- may reflect the borrowing of variants from varieties of Vasconic in which the destabilization of *d*-, which must have occurred in the reconstructable prehistory of Basque as a prelude to its loss²⁴, was already underway; and the same may be assumed for the initial *i*- of the Celtic words.²⁵ In terms of relative chronology the variants with *z*- and *i*- would thus be systematically later borrowings than those with *d*-.

Should convincing arguments be put forward that a reconstruction with initial ⁺*i*-, ⁺*iorkats* 'deer', deserves to be preferred to account for the Greek and Celtic forms, that too would be harmonious with a Vasconic ancestry: The same prehistoric layer of Basque that lacked word-initial *d*- also lacked word-initial *i*-, or any *i* at all for that matter (Trask 1997: 126); therefore a Proto-Vasconic ⁺*iorkats* too would yield Bq. *orkatz*.²⁶

²³Löpelmann refers to Gk. *órynx*, gen. *óryngos*, which is not in my dictionaries.

²⁴See Lutz 1991 for a comprehensive empirical and theoretical account of destabilization and loss in the phonotactic history of English.

²⁵That *i* is a possible weakening product of *d* is shown by Dutch: *goeie* [χu:ɛə], *kwaai* [kva:ɪə] alongside *goede* [χu:də], *kwade* [kva:də], inflected form of *goed* 'good', *kwaad* 'evil' (Goossens 1974: 79).

²⁶The reconstruction Proto-Vasc. ⁺*iorkats*, and likewise a reconstruction PIE **iork-*, are ruled out as soon as a connection of Gk. *dorkás* and Bq. *orkatz* with

2.3. Phonotactic segmental tendencies: Engl. *gate*, Bq. *ate* 'door'

The possibility discussed in the preceding section has to be reckoned with not only for strict constraints such as that against word-initial rhotics in Basque but also for sporadic losses, which may reflect constraints that did not stabilize. A case in point may be word-initial *g-*, or more generally: word-initial velar plosives, in Basque. There are many Basque words beginning with *g-*, and some beginning with initial *k*²⁷, yet there are also a few that have lost the initial plosive. This sporadic change applies more generally to all plosives, but I will here concentrate on the velars. Michelena (1977: § 12.17) does not want to exclude the possibility that it is actually the voiced plosive that is lost, assuming a stage with a voiced initial in case of apparent loss of voiceless plosives. Examples with velars include (*h*)*obiel* < *goibel* 'clouded sky', *astigar* / *gastigar* 'maple', *oroldio* / *goroldio* 'moss (bot.)', cf. Michelena 1977: 253, 531²⁸; (*h*)*abia* < *kabia* 'nest' (< Lat. CAVEA), *amuts* / *kamuts* 'blunted' (< Occ. *camus* 'blunt'), *okotz* < *kokotz* 'chin', *upa* < *kupa* 'barrel' (Lat. CUPA), *ua* 'cradle' (< Lat. CUNA), cf. Trask 1997: 136. There are also instances of prosthesis, as is to be expected in cases of phonotactic instability: *gakula* < *akula* 'goad'²⁹, *garrathoiñ* < *arrathoiñ* 'rat'³⁰; as a matter of fact, since the etymologies of *astigar* / *gastigar* 'maple', *oroldio* / *goroldio* 'moss (bot.)' are uncertain, they too may be instances of prosthesis rather than loss (Michelena 1977: 253).

Bq. *ate* (in the Low Navarrese dialect *athe*) 'door, gate' has no generally accepted etymology (cf. Agud/Tovar 1990: s.v. *ate*). Various proposals, e.g. the old idea that the word is borrowed Lat. *atrium* 'central room of a house, hall, fore-court',

Akkad. *turáhu* etc. is admitted; cf. note 22 above.

²⁷See Trask 1997: 130f. on the special status of word-initial voiceless plosives in Basque.

²⁸Also *oldei* / *goldio* (id.), cf. Agud/Tovar 1992: s.v. *goldio*.

²⁹According to Löpelmann (1968: s.v. *gakula*), the variant of *akula* arose by contamination with *gako* 'hook'.

³⁰Cp. Span. *ratón* 'mouse', *rata* 'rat', Gasc. *arrat* 'rat', Bq. *aratu* 'rat'. The word occurs only in Romance, Germanic, and Celtic and is of unknown origin. It may be a Vasconic loan-word that was lost in Basque when the constraint against initial rhotics arose. Bq. *arratu* etc. were borrowed from Romance, with obligatory prosthesis because of the constraint. However, since the rat is an immigrant from East and South-East Asia (*Der Große Brockhaus*, s.v. *Ratten*), its European name too may have Asian roots.

as well as the recent suggestion (Löpelmann, s.v. *athe*) that the word is an Arabic loan-word (he relates it to Arab. *fatha* 'opening, hole, sluice', from *fataha* 'to open') are mentioned but not accepted. The authors seem to favor Corominas's view that the word, on account of its antiquity and derivational fecundity, is Proto-Basque. I accept this evaluation and propose that *ate* is a Vasconic word surviving in Basque.

Since the word begins with a vowel in Basque, the question arises whether perhaps it once began with a consonant that was lost in the course of constraint development. Michelena (1977: § 11.10) assumes Bq. *athe* to belong to a type of disyllabic words of the phonological structure VPhV- with initial vowel and aspirated medial plosive for which he considers an origin with initial *h*-likely: *athe* < ⁺*hathe*, with dissimilatory loss of aspiration. In the context of the material cited from Michelena 1977 and Trask 1997 above, an earlier ⁺*gate* also appears likely. Perhaps the two reconstructions are not mutually exclusive (in dialects preserving aspiration): ⁺*gathe* > ⁺*hathe* > *athe*.

Some speakers of English when walking from the main building to the departure area of Bilbao Airport may wonder why the words on the bilingual exit signs are so similar: *Atea 3 / Gate 3, Ateak 1-2 / Gates 1-2*. Could Bq. *ate* and Engl. *gate* be the same word? Engl. *gate* belongs to a group of Germanic words meaning 'hole, door, passage' that show morphological irregularities and have no Indo-European etymology. Onions (1966) has two entries for *gate*.

gate¹ ... opening in a wall capable of being closed by a barrier; barrier itself, especially framework on hinges. OE *ġæt*, *ġeat*, pl. *gatu*, corr. to OFris. *gat* hole, opening, OS. *gat* eye of a needle (LG., Du. *gap*, hole, breach), ON. *gat* opening, passage :- Germ. *zātam* (wanting in HG. and Goth.). Forms with initial *y*-, repr. OE. forms *ġeat*, pl. *ġeatu*, remain in northerly dial. *yett*, *yeat*, and in the surnames *Yates*, *Yeats*; but the standard literary form has been *gate* since XVI [the 16th century]."

gate² ... A. (north. dial.) way XIII (Orm); street (surviving in place-names, as *Canongate*) XV; B. †going, journey XIII; manner of going (see GAIT). - ON. *gata* = OHG. *gassa* (G. *gasse* lane); Goth. *gatwō* :- Germ. **gatwōn* (whence Lith. *gātūė*, Lett. *gatwa*), of unkn. origin (connexion with GET has been assumed by some, as for GATE¹).

The OED is more elaborate, but reaches the same etymological

conclusion: With regard to *gate*¹: "The ulterior etymology is obscure." With regard to *gate*²: "As to the ulterior etymology nothing has been ascertained." In de Vries (1962: s.v. *gata* [corresponding to *gate*²]) it says, "Dunkles wort; zusammenhang mit *gat* ['hole', i.e. *gate*¹] ist fraglich." Also Lehmann (1986: s.v. *gatwo*) says, "Etymology obscure."³¹ Both de Vries and Lehmann cite attempts at connecting the word with etyma meaning 'defecate' or 'anus', but Dutch *gat* 'anus' is merely a vulgar use of *gat* 'hole', and the German etymological dictionaries do not even mention those attempts any more: Kluge/Seebold (1995: s.v. *Gasse*) note that OHG *gazza* (MHG *gazze*, NHG *Gasse*), ON *gata* and Gothic *gatwō* 'city street, plaza' are close matches except that at least the Old Norse form does not agree exactly with the Gothic one; here too the result is "Herkunft unklar." Likewise Pfeifer et al. (1989: s.v. *Gasse*): "Herkunft unbekannt. ... Außergerm. Verwandte fehlen."

In my view we are dealing with a single Germanic word family: There is one and the same root to all these items, *gat-*; and there is a common meaning, 'passage' (which is not that of *get-*, so these roots have nothing to do with each other). The fact that there are unexplained morphological peculiarities in the family, together with the other fact that the family is unique to Germanic, suggests that we are dealing with a loan complex. Since Bq. *ate* 'door, gate' is similar both phonologically and semantically and may, furthermore, derive from an earlier *+gate*, and since this item is by eminent authorities considered indigenous to Basque, I propose that the Germanic words were borrowed in prehistoric times from Vasconic languages and preserve the same root as the Basque word, *+gat-* 'passage'.

2.4. Phonotactic sequential tendencies: Engl. *chamois*, Bq. *ahuntz* 'goat, chamois'

Engl. *chamois*, first attested in 1560 (OED, s.v.), is, of course, a French loan-word, which is itself first attested in the 14th century as a word from Southern France, derived from Gallo-Lat. *camox*, CAMOCE (Gamillscheg 1969: s.v.). The animal is described in the OED as follows:

A capriform antelope (*A. rupicapra* or *Rupicapra tragus*), the

³¹Feist (1939: s.v. *gatwo*) cites half a page of etymologies, but his verdict too is "Etymologie dunkel", which Lehmann merely translates.

only representative of the antelopes found wild in Europe; it inhabits the loftiest parts of the Alps, Pyrenees, Taurus, and other mountain ranges of Europe and Asia. Its size is that of a full-grown goat; it is covered with brown hair, and has horns, about six inches long, which rise straight above the head, bending back so as to form complete hooks. Its agility and keenness of scent make its chase most difficult and exciting.

Note the reference both to the form and the size of the goat in this description. There is indeed a high degree of similarity between the chamois and the goat, even though they are very different species. As for the linguistic side, the OED offers the following:

chamois (['ʃæmɔɪ, 'ʃæmɪ, 'ʃæmwa), *n.* Forms: see below. [a. F. *chamois* (16th c. in Littré), prob. from Swiss Romanic: in Tyrolese *camozza*, *camozz*, Piedm. *camossa*, *camoss*, mod.Pr. *camous*, Rumansch *camuotsch*, *chamotsch* (Diez); It. *camozza*, *camoscio* (cf. *camoscia* chamois leather); Sp. *camuza*, *gamuza*, Pg. *camuça*, *camurça*. Presumably of the same origin as OHG. and MHG. *gamz*, mod.G. *gemse*, but the relations between the Teutonic and Romanic words have not been ascertained, and no etymology is known either in Latin or Teutonic. ...]

Kluge/Seebold (1995: s.v. *Gemse*) address the question of "the relations between the Teutonic and Romanic words" by offering two possibilities: Either the German word is a Romance loan, or both the Romance and the German word are loans, and in this case probably from one and the same substrate language:

Mhd. *gemeze*, ahd. *gamiza*. Entlehnt aus einer romanischen Sprache (it. *comoscio m.*, spl. [spätlateinisch, Late Latin] **camox*) oder aus der Substratsprache, aus der das lateinische Wort vermutlich stammt.

The Bavarian name of the chamois, *Gams* (without umlaut), according to Kluge/Seebold (1995: s.v. *Gams*) may go back directly to the Late Latin form *camox*. But considering the form OHG *gamiza*, an earlier form of *Gams* may well have been **gamuza*. A pre-German suffixal variation between *u* and *i* is also evident in other substratal domains (cf. Vennemann 1995b: § 10).³²

³²It is also a feature of Basque, cf. Michelena 1977: ch. 3.

The reconstruction of an etymon common to all the word forms cited is difficult, which is normal for a loan complex with several target languages (see also Meyer-Lübke 1935: s.v. *camox*). Perhaps the cumulative evidence favors an etymon $+kamuts(a)$, with a laminal affricate *ts* on account of the medieval German spelling *z*.³³ But there is variation at least in the following directions: $+kamutso$ (or $+kamusso$, cf. Hubschmid 1960: 69, $+kammusso$, Gamillscheg 1969: s.v. *camox*) to account for the masculine forms; $+kamussio$ (Gamillscheg 1969: s.v. *camox*) to account for forms with palatal sibilants; $+kamutsio$ etc. to account for variants with palatal affricates (cf. the “paleo-europeo alpine *camocio*”, crossed with Lomb. *camòsc*’, given as a source for It. *camoscio* in Devoto 1968: s.v. *camocio*, also “Rumansch *camuotsch*, *chamotsch*” in the above OED quotation, perhaps also Friaul. *kamuts*, Tyrol. *kyamorts* in Meyer-Lübke 1935).³⁴ Latin with its substitution of the animal name termination *-ok-s* (perhaps *-òk-s*, cf. Walde/Hofmann 1982: s.v.) is a special case.³⁵ For this suffix, Walde/Hofmann invite us to compare Lat. *esox*, *isox* (also Hesych. *isox*) ‘ein Rheinfisch, wrsch. Lachs’ (i.e. ‘a Rhine fish, probably salmon’). I will return to *esox/isox* briefly at the end of this section.

³³OHG *gamiza*, MHG *gemeze* point to a sibilant or affricate source similar in sound to those resulting from the High Germanic (Second) Consonant Shift.

³⁴The variants with *g-* probably are not of great antiquity. According to Corominas/Pascual (1980: s.v. *gamuza*) the Spanish and Portuguese forms are not indigenous to the Iberian Peninsula but imported from Italy in the chamois leather trade; the doublets may have arisen in the loan transfer. The OHG *g-* in *gamiza* etc. would be a normal reflex of either *k-* or *g-* in Romance borrowings, because Old High German did not have a phoneme /k/ but only /g/ and /k^x/ (cf. Vennemann 1993b: 284, n. 57); the same argument would apply to non-Romance borrowings. It is therefore not necessary to set up a separate form $+gamutsa$.

³⁵The assumption that Vasconic $+ts$ (Basque *-tz*) could be borrowed into Latin as *-x* finds support in the names of the Roman inscriptions of Aquitaine, where *x* may appear for sibilants, especially those different from $+s$, including what from a Basque point of view seems to have been the affricate $+ts$, e.g. *Anix-*, cf. Bq. *haritz* ‘oak’; *Belex*, *-belex*, *-bel(e)s*, Bq. *beltz* ‘black’ (*bele* ‘raven’); *Berhax*, Bq. *beratz* ‘soft, benign’; *Bihox-*, *Bihos-*, Bq. *bihotz* ‘heart’; *Osson*, *Oxson*, *-oss(o)-*, *-ox(o)-*, Bq. *otzo* ‘wolf’; but also, after liquids and glides, *Bors-*, Bq. *bortz* ‘five’; *Hars-*, Bq. *hartz* ‘bear’; *Heraus-*, Bq. *herautz* ‘boar’ (Gorrochategui 1984: 359-362, 1995: 43f., Trask 1997: 399f.). Whether this only represents graphic attempts at rendering Aquitanian pronunciations adequately or whether occasional sound substitutions, namely of [ks] for [is], are involved, seems hard to determine.

The basic form may have been ⁺*kamuts*, to which an article ⁺*-a* could be added, as in Basque; since nouns ending in *-a* were likely to be interpreted as feminine by speakers of early Indo-European languages, the coexistence of ⁺*kamuts* and ⁺*kamutsa* would explain why both masculine and feminine adaptations have occurred in the borrowing languages.

The history of research on Lat. *camox*, and a new attempt at coming to grips with the strange termination *-ok-s* (or *-oks*) or *-ōk-s*, are presented by Szemerényi (1989: 141-157). He also makes the most curious of all attempts at explaining the chamois word etymologically: To him the name is pure Indo-European, viz. a Celtic loan-word: **kambo-uksō(n)* "hirschartiges Tier mit krummem (Gehörn)" ("deer-like animal with crooked (horns)"), a compound of Celt. **kambo-* 'krumm, crooked' and PIE **uksōn* 'ox, bull, deer'. Rather than proving *camox* to be Indo-European, in my view this desperate construction makes it appear even more likely that the opposite is true.

This leaves us with the question of what the substrate language may have been from which the item was taken (cf. the quotation from Kluge/Seebold 1995 above). Within the theory with which I am working, it should be a Vasconic language, because according to this theory all of Western Europe (Vennemann 1995a), and the Alps in particular (Vennemann 1997b), were Vasconic when the speakers of Indo-European languages arrived. So we have to look into Basque. Aulestia/White (1990: s.v. *chamois*) offer *orkatz*, which also means 'deer' (Aulestia 1992, de Azkue 1984: s.v. *orkatz*)³⁶, and *sarrio*, which only has this meaning (Aulestia 1992). I have not found *sarrio* in de Azkue 1984. It is obviously the same word as Span. *sarrio* 'izard, male chamois' about which Corominas/Pascual (1980: s.v.) write, "parece tratarse de una palabra ibérica o protovasca IZARR-, cuyo significado originario es dudoso."³⁷ These are, of course, dead ends. Remembering the similarity of the chamois to the goat, one should also check under 'goat'. The result again appears to be disappointing: *ahuntz* 'goat'.³⁸

³⁶For 'deer' Aulestia/White (1990) offer *orein*, *orkatz*, and *basahuntz*. So does Trask (1997: 298), who adds to *orkatz* "(also 'Pyrenean chamois)" and to *basahuntz* "(*basa-* 'wild' + *ahuntz* 'goat')". See section 2.2.2 above.

³⁷Cf. Corominas/Pascual (1980: s.v. *gamuza*): "El antiguo nombre ibérico o celtibero parece haber sido IBEX y el vasco-pirenaico IZARR-."

³⁸One good thing that can be said about Bq. *ahuntz* 'goat' and which warrants

Löpelmann translates *ahuntz* as 'goat, chamois' ('Ziege, Gemse'). Looking therefore at *ahuntz* more closely, and doing so in the light of phonotactic constraints and tendencies, the word becomes more interesting. First, word-initial *k-* (and *g-*) may fall in the history of Basque, as discussed in the preceding section.³⁹ Second, *ahuntz* is reconstructed as ⁺*anu(n)tz* (Michelena 1977: § 5.5⁴⁰, Agud/Tovar 1990: s.v. *auntz*, Trask 1997: 299), with obligatory intervocalic loss of *n* (for which cf. Trask 1997: 139). Assuming both developments in the case of *ahuntz*, one reconstructs ⁺*kanu(n)tz*. In a labial context *m* and *n* may have been confused; so ⁺*kanu(n)tz* may have been ⁺*kamu(n)tz* at an earlier stage.⁴¹ This form, which should be written ⁺*kamu(n)ts* outside the modern Basque context, would be almost identical with that reconstructed for the *chamois* word earlier on. It is worth mentioning that Basque, and therefore probably also the early West European substratal languages related to it, lack gender distinctions; ⁺*kamu(n)ts* and ⁺*kamu(n)tsa*, distinguished as indefinite vs. definite, could therefore be misunderstood as differing in gender rather than in definiteness to speakers of languages encoding gender but not definiteness.

Agud/Tovar and Trask reconstruct ⁺*anu(n)tz*, suggesting that perhaps originally there was only the intervocalic nasal, the preconsonantal one arising by way of compensation when the intervocalic one became unstable, nasalizing its vocalic

the following discussion is that the great Corominas considered the word to be Paleo-Basque (see Agud/Tovar 1990: s.v.).

³⁹A famous example of the loss of *k-* is Bq. *harri* 'stone', which is identified with words for 'stone' in Ibero-Romance, Gallo-Romance, Insular Celtic, and Swiss German which all reflect a pre-Indo-European root ⁺*kan-* 'stone' (Hubschmid 1960: 36), and which is also evident in derivatives and toponyms (cf. Kluge/Seebold 1995: s.v. *Karst*). Most famous among the latter is Carrara, which is described in Der Große Brockhaus as a city world-famous for its production, manufacturing, and shipment of marble, and as situated in a deep valley of the Apuan Alps, four miles away from its harbor, Marina di Carrara, on the Ligurian Sea. Therefore I reconstruct *Carrara* < ⁺*kan-ara(n)* 'stone-valley' (cf. Bq. *aran* 'valley'). The connection of *Carrara* with Bq. *harri* 'stone' etc. had also been made, without my knowing, by Cowan (1984: 446), which I am glad to acknowledge.

⁴⁰Michelena carefully says "*ahuntz*, acaso ['perhaps'] procedente de ⁺*anu-*."

⁴¹There has been confusion in the opposite direction, so to speak: replacement of *n* by *m* after *u*, as in *kuma* 'cradle', a variant form of *ua* (both < Lat. *CUNA*); cf. Trask 1997: 140.

environment. That is certainly a possibility. However, the other possibility, a reconstruction with two nasals, ⁺*kamunts*, would also yield the extant forms, viz. by a dissimilatory development of ⁺*kamunts* into ⁺*kamuts* in the borrowing languages (or even the Vasconic source languages).⁴²

There may be a specific argument in favor of assuming a Proto-Vasconic etymon for the chamois word with a preconsonantal nasal. Trask (1997: 132-135) argues that an earlier form of Basque did not have a bilabial nasal phoneme but that *m*, where it occurs in regular Basque words, was either borrowed together with the word or developed from *nb* or from *b*, from *b* especially when a nasal followed later in the same word. Assuming this tendency to be old in the language (it applies both to native words as well as to loan-words of all periods), we should perhaps reconstruct a Proto-Vasconic ⁺*kabunts*, a form which was changed to ⁺*kamunts* not only in Basque itself but also in those Vasconic languages from which the chamois words were borrowed. This reconstruction would permit a real etymological explanation of the chamois words, as well as of Bq. *ahuntz* 'goat', namely by associating it, as ⁺*kabunts*, with a well-known set of etymologically problematical words (cf. Ernout/Meillet 1959: s.v. *caper*, Walde/Hofmann 1982: s.v. *caper*): Lat. *cap-r-* (*caper*, *capra*) '(male and female) goat', Umbr. *kaprum*, *kabru* 'caprum', identical with Germanic ⁺*haf-r-* in OE *hæfer*, ON *hafr* 'he-goat', Gm. *Haber-* 'goat'⁴³; Gk.

⁴²Perhaps the mysterious *r*-forms such as Tyrol. *kyamorts* find an explanation here, to which Corominas/Pascual (1980: s.v. *gamuza*, n. 1) add Belluno *camorz*, Livinallongo *chamouré*, Gardena *chamoré*, Comelico *chamorθa*, Savoyan *famúr*, and the place-name *Chamourse* in the Dauphiné Alps, also Calabr. *camorcia* 'chamois leather'. The Portuguese *camurça* is declared secondary in Corominas/Pascual (1980: s.v. *gamuza*), namely as either an instance of hypercorrection or as a contamination with *murça* 'mozzetta' or *corço* 'deer', or as an import of one of the Ladinian forms. The latter have been explained as analogical too, namely by contamination with *r*STU's 'bear' (cf. Meyer-Lübke 1935: s.v. *camox*), but this explanation has not been adopted by Corominas/Pascual. If a reconstruction of the chamois word as ⁺*kamunts/kamuntsa* is chosen, the second nasal would permit a phonological explanation of the *r*-forms by progressive nasal dissimilation leading to coda weakening.

⁴³In *Habergeiß* '(common) snipe (*scolopax gallinago*)'; here *-geiß* 'goat' explains the obsolescent *Haber-* 'goat' tautologically, and the entire word is applied metaphorically to the bird on account of its bleating pairing call (Kluge/Seebold 1995: s.v. *Habergeiß*).

kápros 'boar'; OIr. *gabor* 'he-goat'⁴⁴, Gael. *caer-iwrch* 'deer', Ir. *caera* 'sheep' (to the extent that all these are related among each other). With their short *a*, their phonological irregularities, and their restriction to West-Indo-European these words give the appearance of a loan complex. The theory would suggest that the giving languages were Vasconic. If my etymological analysis of Fr. *chamois* and Bq. *ahuntz* is correct, this theoretical suggestion has empirical support. We may in this case even wish to include West Germanic ⁺*s-kāp-a* (OE *scēap*, OFris. *skēp*, OHG *scāf*) 'sheep'; the morphological analysis would suggest an old prefix or compound formation (both morphological techniques are evident in the history of Basque), and the phonological discrepancy between Lat. *kap-* and OHG *-kāp-*, in particular the length difference in the root vowel⁴⁵, is fully within the limits of loan phonology.

Returning to Lat. *esox/isox*, I would like to suggest that this word is best understood as a Vasconic loan, based on an identifiable Bq. *iz-* 'water'. I think (pace Trask 1997: 328) that it contains the same item *iz-* with the reconstructible meaning 'water' which survives in Basque compounds, e.g. *izurde* "water-pig", i.e. 'dolphin' (Bq. *urde* 'pig').⁴⁶ De Azkue's (1984) interpretation of Bq. *izurde* 'dolphin' as "water-pig" strikes me as semantically sound; it is, furthermore, supported e.g. by MHG *merswīn* 'dolphin' (literally "sea-pig"). By contrast, Michelena's interpretation of Bq. *izurde* as a reflex of **gizurde* "man-pig", formed with *giza-*, the combining form of *gizon* 'man'⁴⁷, appears to me semantically aberrant; it also seems to lack comparative support, even if it is phonologically possible. If *giza-urde* is, according to Michelena, actually attested in Basque, then that may simply mean that earlier speakers of Basque had hit upon this popular etymology before Michelena, probably at a time when ⁺*iz* was fading out of the language as a simplex and

⁴⁴This word offers the phonological problem that both plosives apparently show a wrong grade. For the medial plosive the problem is resolved by Thurneysen's (1946: § 227e) sound law whereby *p* is voiced before liquids. For the initial plosive, however, there is no such solution, as Thurneysen himself seems to admit: "*g-* instead of *c-* under the influence of *ga(i)bid* 'takes, seizes'". For a loan-word, which I consider the *cap-* word for the goat to be, such uncertainty in the apprehension of the grade of obstruents is normal.

⁴⁵See note 11 above.

⁴⁶Cf. Vennemann 1994b: 275 n. 78, 1996: 138 n. 19.

⁴⁷Michelena 1977: 116 n. 13, cf. also Arbelaiz 1978: 88. The etymology is cited approvingly by Trask (1997: 328).

became a meaningless *iz-*.⁴⁸ Even if Bq. *izokina* 'salmon' is a Latin loan (Walde/Hofmann 1982: s.v. *esox*, Trask 1997: 371), this in no way precludes an interpretation of Lat. *esox/iso* as a Vasconic loan. Loan words do come back, as any two languages with a long history of contact show.⁴⁹

In short, I would like to suggest that a Vasconic *+izuts* or *+izunts* 'salmon' (with the same termination as in *+kamuts*) was adopted in Latin as *iso*/*esox* 'salmon' and that the derivate Lat. *esocīna* 'salmon' was borrowed into Basque as *izokin(a)* 'salmon', with substitution of *iz-* for *es-*, perhaps at a time when *+iz* 'water' was still alive in the language as a simplex.

2.5. Phonotactic sequential change: Gk. *éthnos* n. 'group of people, class, people', pl. *tà éthnē* 'the heathens', Bq. *enda* 'cast, race'

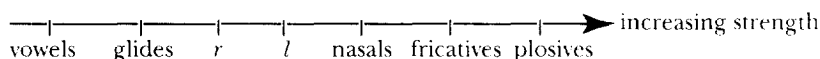
Both the Greek word *éthnos* n. 'group of people, class, people', whose plural *tà éthnē* means 'the heathens', and the Basque word *enda* 'cast, race'⁵⁰ have no convincing etymology (Frisk 1973: s.v., Agud/Tovar 1991: s.v.). Yet both the sound structures and the meanings of the words are very similar, which suggests a common origin.

It is not quite obvious why Bq. *enda* 'cast, race' counts as unexplained. Certainly this word is similar enough to the beginning of Romance words like Sp. *étnico*, *etnografía*, *etnográfico*, etc. to be explained as a borrowing; and furthermore one translation of 'race' into Unified Basque is *etnia* (Aulestia/White 1990: s.v.). The word *etnia* is not well-formed in Basque. It violates a constraint which says that in an intervocalic cluster of consonants the Consonantal Strength as measured on the following scale,

⁴⁸Much like the Proto-Germanic adverb and preposition *uz* 'original(ly), from out of' in Gm. *Urahn* 'ancestor', *Ursprung* 'origin', *Urlaub* 'vacation', etc., or like *mer-* from OE *mere* 'sea' in Engl. *mermaid* for that matter.

⁴⁹E.g. Gm. *Wagon* alongside *Wagen* (via French), Engl. *guerrilla* alongside *war* (via Spanish).

⁵⁰It only occurs in the Bizkaian dialect, cf. de Azkue (1984: s.v.). Agud/Tovar (1991) enter *endeka* 'degenerarse' under *enda*¹ 'casta, raza', but that is likely to be an error: *endeka* is probably borrowed Span. *enteco* 'sickly, infirm, ill, meagre'. Cf. de Azkue (1984) where *endeka* is translated '[Span.] enteco, degenerado, [Fr.] rachitique, dégénéré'. Span. *enteco*, *-a*, together with the verb *entecarse* 'to become chronically ill', goes back to *ético* and ultimately to Gk. *hektikós* 'habitual'.



must increase; in short, with "cs(X)" for 'the Consonantal Strength of segment X' (and the asterisk expressing non-wellformedness):

*C_aC_b / V __ V, unless cs(C_a) < cs(C_b)

I.e., clusters such as *rg*, *nd*, *sk* are permitted, but clusters such as *gr*, *dn*, *ks* are not permitted in a word. Neither are clusters such as *kt* (or geminates, for that matter). One may argue that alongside *st* we find *ts*; but this is not a valid counterargument, because *ts*, *tz*, and *tx* are generally assumed to be monosegmental, i.e. affricates.

Since traditional Basque does not permit medial clusters of more than two consonants, and always places the syllable break between two medial consonants, the above formula expresses the fact that the language is in harmony with the Syllable Contact Law (for which Vennemann 1988: 40 may be compared), whereas *-tn-* in *etnia* would be a bad cluster in terms of this Law. One remedy for bad syllable contacts is metathesis (Vennemann 1988: 55).⁵¹ Applying metathesis to *etnia*, the resulting *ˀentia* would be fairly close to *enda*, especially since plosives have to be voiced after nasals in the Western dialects of Basque (Michelena 1977: § 18.9).

However, I am not content with this explanation myself. The reason is that Gk. *éthnos* n., from which Span. *étnico* etc. and hence Bq. *etnia* are derived, is itself unexplained etymologically.⁵²

Therefore I am inclined not to derive Bq. *enda* from *etnia*

⁵¹Such metatheses are on record, e.g. Span. *ladrillo* 'clinker, brick' → Bq. *ardillu* (d'Urte, cf. Michelena 1977: § 18.5, Agud/Tovar 1989: s.v. *adarailhu*), Fr. Span. *sable* → Bq. *salbe* 'sand'.

⁵²Frisk (1973: s.v.) says "nicht sicher erklärt". Chantaine (1990: s.v.) compares a Greek adjective *othneios* 'foreign' and noun *éthos* n. 'habit' and suggests that the root contains digamma and goes back to **swedh-*, derived from the 3rd person reflexive pronominal **swe*, Gk. *hê*. The same proposal is made without reservation by Watkins (1985: s.v. *ethnic*). The problem with this etymology is that the meaning of the Greek word is not reflexive at all but tends to mean 'foreign group', hence in the plural (*ta éthnē*) 'the heathens'. I therefore doubt that an Indo-European origin of Gk. *éthnos* has been ascertained.

and thus eventually from Greek but to consider it a genuine Vasconic word, which Basque has inherited from the parent language. If Gk. *éthnos* has no explanation in Indo-European, it must be a loan-word. One source of Greek loan-words, as I have suggested in my 1997a article, is the Central European Vasconic substratum. Since in phonological terms the Greek word represents the lectio difficilior, so to speak, because a medial cluster of plosive plus nasal is less preferred than the reverse, the Greek word must preserve the ancient order, so that a reconstruction PVasc. ⁺*edna* 'group of people, class, people, race' recommends itself. This seems to me to be a plausible source of both the Basque word (with contact metathesis, ⁺*ed.na* > *en.da*, when the above sequencing constraint developed in this language) and the Greek word (with the Greek consonant shift which takes old words with ⁺*d* through a stage of voiced aspirates, ⁺*d^h*, to a stage of voiceless aspirates, ⁺*t^h*⁵³). Also a source PVasc. ⁺*etna* is conceivable, with an adaptation ⁺*etn-* → *et^hn-* in Greek and metathesis and sonorization in Basque: ⁺*et.na* > ⁺*en.ta* > ⁺*en.da*.

2.6. Phonotactic cluster reduction: OE *sceanca* 'thigh', Bq. *zango*, *zanko* 'leg, foot'

In my 1995a article I compared the etymologically unexplained OE *sceanca*, MHG *schenkel* (NHG *Schenkel*) etc. 'thigh' to the likewise etymologically unexplained Bq. *zango* (Eastern dialect *zank(h)o*) 'leg, foot' and suggested that the difference in the root, ⁺*skank-* in Germanic with its initial cluster and *zang-* in Basque with its single initial sibilant, might be the effect of phonotactic change in prehistoric Basque. Historical Basque does not tolerate word-initial consonant clusters:

*C₂ /# _____

However, there is no reason to assume that the same is true of prehistoric Basque. It is known, e.g. from Middle and Modern Korean (Kwon 1996) and from Old and Middle Indic (Fahs 1989: 30f.), that languages with word-initial clusters may develop such a constraint within a few centuries, using several

⁵³Cf. the argument for ⁺*g* > *k^h* in the comparison of Span. Port. *garza* 'heron' (< ⁺*gardea*) with Gk. *charadriós*, name of a bird identified with various grallatorial birds (Vennemann 1996). See note 19 above.

different mechanisms to implement the new constraint. My 1995a proposal is that Basque reduced its initial clusters on the basis of the same Consonantal Strength hierarchy that it used to sequence intervocalic clusters: Of a word-initial consonant cluster, only the most vowel-like consonant survives:

Cluster reduction rule (for consonant clusters Σ)

PVasc. $\Sigma > C_{\min} / \# \text{ ____}$

where C_{\min} is that consonant in Σ whose Consonantal Strength is least

A special case of this rule is the reduction of initial fricative-plus-plosive clusters FP - to the mere fricative F :

$+FP > F / \# \text{ ____}$ for fricatives F and plosives P

Applying this rule to the case on hand,

$+sk > s / \# \text{ ____}$

and remembering that plosives have to be sonorous after nasals in most dialects (Michelena 1977: § 18.9), we see that a root $+skankV-$ would indeed develop into $/sank^{(h)}V-/$, i.e. $zank(h)V-$, and further into $/sangV-/$, i.e. $zangV-$, in Basque. This is the basis for my proposal that Bq. *zango*, *zank(h)o* 'leg, foot' continues a PVasc. $+skankV-$ and that Gmc. $+skankV-$ 'thigh' is a Vasconic loan-word.

Another example treated in my 1995a article is the Germanic strong (class III) verb $+stink^{w}an$ 'to push', Goth. *stiggan* 'to push' (with a change of meaning in OE *stincan* etc. 'to stink'). Applying the special case

$+st > s / \# \text{ ____}$

of the above initial cluster reduction rule, a PVasc. $+stunk^{w}a-$ 'to push' could be the basis both for Bq. *zunka* 'thrust, blow' (preserved only in Eastern dialects where plosives do not become sonorant after nasals, cf. Michelena 1977: § 18.9)⁵⁴

⁵⁴Owing to another segmental constraint, Basque does not allow labialized velars.

and, as a Vasconic loan-word in Germanic, for the ablauting Gmc. $+stink^{w}an-$ / $+stank^{w}a-$ / $+stunk^{w}um-$ / $+stunk^{w}an-$ 'to push'.

2.7. Applications

2.7.1. OHG *scarpf/sarpf* 'sharp', *skulan/suln* 'shall'

Old High German contains two words with a peculiar variation of initial *sk-* and *s-* for which no generally accepted explanations exist. Both of them give the impression of being non-Indo-European loan-words, and since such words may be of Vasconic origin, an effect of the same constraint as in the preceding section must be considered.

The word OE *scearp* (Engl. *sharp*), OFris. *skerp*, *skarp*, MLG MDutch *scharp*, *scherp* (Mod.Du. *scherp*), OS *skarp*, ON *skarpr*, OHG *scarpf* (NHG *scharf*) 'sharp' is by Kluge/Seebold (1995: s.v. *scharf*) connected to Latv. *skāibs*, Mir. *cerb* 'cutting' (Pres. Part. belonging to *cerbaim* 'I cut') and further to PIE $*sker-$ 'to cut'. "Die Abgrenzung ist im einzelnen wegen weit auseinanderfallender Bedeutungen unklar (so z.B. die Beurteilung von ae. *sceorpan* 'schmerzen, schaben, schneiden')." The OED (s.v. *sharp*) offers a different picture: "The Teut. root $*skerp-$: *skarp-*: *skurp-* appears also in OHG. *scurfan*, MHG. *schürfen* to cut open (mod.G. to poke a fire), OE. *scearpe* scarification, *scearpian* to scarify. The Teut. root $*skrep-$: *skrap-* (see *scrape* v.) appears to be related; no cognates outside Teut. are known."

This does not appear to be the end of the uncertainty noted by Kluge/Seebold. There is also in Germanic a verbal root $+kerb-$ (late MHG *kerben*, MLG MDutch *kerven*), as a strong verb in OE *ceorfan*, cf. further ON *kyrfa* 'to carve', Dan. *karve* 'to notch, indent', Swed. *karfva* 'to notch, carve'. Both Kluge/Seebold (1995: s.v. *kerben*) and the OED (s.v. *carve* v.) point to a traditional association with Gk. *gráphein* 'to write', originally 'to scratch or engrave'. Neither note the phonological and semantic similarity to the *sharp* word.⁵⁵

⁵⁵There are more etymological riddles connected to this group which I cannot all follow up in this paper. E.g., Gmc. $+grab-a-$ (str. verb, class VI) 'to dig' (Goth. OHG *graban*, ON *grafa*, OS *grava*, OFris. *grēfa*, *grīova*, OE *grafan*) can, according to Kluge/Seebold (1995: s.v. *graben*), only be directly compared to Latv. *grebt* 'to scrape out, to hollow out' and to OCS *greti* 'to row, to dig' and further be associated with PIE $*g^hr̥ebh-$ 'to grasp, to grip, to seize'. It seems to me that the concept of digging is semantically closer to the idea of using a sharp instrument than to that of gripping. The OED (s.v. *grave*v.) says,

There is yet another peculiarity connected to the *sharp* word which Kluge/Seebold do not mention but the OED does: There exists a form OHG MHG *sarpf*, *sarf*, early MDutch *sarp* 'sharp'. Köbler (1994) lists *sarpf* together with *scarpf* (s.v. *skarpf*⁵⁶), and likewise for the derivatives. By contrast, the OED says, "The OHG. and MHG. *sarpf* (early mod.Du. *sarp*) sharp is prob. unconnected." This is understandable in terms of Indo-European and Germanic etymology where such an alternation (of *sk-* and *s-*) has no place. But apart from that, assuming two unrelated words *scarpf* and *sarpf* of exactly the same meaning in the same language is in my opinion an illegitimate appeal to chance as a mode of explanation.

I do not have an account accommodating all of the peculiarities mentioned. But I think that a few things can be said about them.

First, the phonological irregularities in the above correspondences, e.g. in terms of Grimm's Law, as well as the restriction to some of the Western Indo-European languages, point to a loan complex.

Second, the variant *sarpf* of *scarpf* 'sharp' opens up additional connections: (1) With *s-* > *h-* / # ___ V, the same root form ⁺*sarp-* 'sharp' can be identified in Gk. *hárpē* f. '(1) sickle, (2) bird of prey, kite'⁵⁶ and its likely derivatives (cf. Frisk 1973: s.v. *hárpē*, *harpázō*): *hárpasos* (name of a bird of prey), *hárpax* f. 'robbery', *hárpax* m. 'robber'⁵⁷, *harpázō* 'to snatch away, rob', *hárpagos*, *harpágē* 'hook', etc. Frisk (1973: s.v. *hárpē*) compares as "wahrscheinlich urverwandt" Gk. *hárpē* to OCS *srǫpǫ* 'sickle'⁵⁸,

"Cognates are found in OS. *greba* I dig (also, I row), *grobu* ditch, Lettish *grebju* I scrape. Connexion with Gr. *γράφειν*, to write, is no longer accepted by philologists." This connection must indeed be rejected if the compared items are understood as native Indo-European; in a loan-complex involving several borrowing languages such want of precision, especially on the phonological side, is exactly what must be expected.

⁵⁶According to Frisk (1973: s.v. *hárpē*), the 'bird of prey' word is derived metonymically from the 'sickle' word, after the sickle-shaped claws. I have no objection against the assumption that the 'bird of prey' word was derived metonymically from an old 'claw' word. However, since claws are older than sickles, perhaps the 'claw' meaning is basic to the group, the 'sickle' meaning being metaphorically derived from the 'claw' meaning. The original meaning may be 'sharp (object)', see below.

⁵⁷"Eigentlich 'mit Krallen versehen' (?)" (Frisk 1973: s.v. *hárpē*). I think this is a simple metaphor, cf. Gm. coll. (*sich* dat.) *krallen* w. acc. 'to take/snatch away, take possession of (unlawfully)'.

⁵⁸Russ. *serpǫ* 'sickle' (Walde/Hofmann 1982: s.v. *sarpǫ/sarpō*).

Latv. *sirpe* 'sickle' and possibly Lat. *sarpere* 'die Weinstöcke beschneiteln (to lop vines)' (for which cf. Walde/Hofmann 1982: s.v. *sarpiō/sarpō*) and OHG *sarpf*.

Third, an item that so far seems to have been overlooked in the discussion of the above words is Span. *zarpa* meaning 'claw' but designating at the same time a variety of vaguely related concepts. In Basque there is a word *zarpa* with several meanings overlapping those of Span. *zarpa*, one of them being '[Span.] espolón, [Fr.] éperon (Bot.)', i.e. 'spur (in botanical terminology)'.⁵⁹ The Spanish word *zarpa* 'claw' has several peculiar aspects (Corominas/Pascual 1991: s.v.), most interesting among them in the present context being the fact that it is restricted to Castilian and adjacent Pyrenean dialects and is not attested (with this meaning) before a. 1611. Corominas/Pascual (1991: s.v.) ask the question of whether the word might be Basque in origin ("o deberemos mirar más bien el vocablo vasco como genuino?") but do not answer it in the affirmative. In my view it has to be so answered, because the word is not original in Romance (it does not even occur in Portuguese), and there is no other language around from which both Cast. *zarpa* and Bq. *zarpa* could be assumed to have been borrowed.⁶⁰ I propose that *zarpa* is originally a Basque word meaning 'claw' that was borrowed into Castilian, where it preserved the original meaning, and then gradually receded in its original meaning until it survived with the metaphorical meaning 'spur'.

As for the other **sarp-* words discussed above, I would like to propose that they all go back to the same Vasconic word **sarpV-* that survives residually in Bq. *zarpa* 'spur', i.e., that they are prehistoric Vasconic loans exactly as Span. *zarpa* is a medieval Basque loan. As a likely original meaning of Vasconic **sarpV-* I would like to consider 'claw'; 'spur', 'rob', 'sharp', and 'sickle' would then represent metonymical and metaphorical shifts based on this original meaning. Alternatively, and perhaps preferably, 'sharp' or 'sharp object' could be basic, with 'claw', 'spur', 'rob', and 'sickle' as specific applications.

Is there then a way to connect this group of **sarp-* words to

⁵⁹This meaning occurs only in one community of the High Navarrese dialect. One interpretation of such restricted application makes it appear likely that this meaning is residual and thus possibly very old.

⁶⁰In particular the Castilian word is unlikely to continue Lat. *sarp-*, because the regular reflex of Lat. *s-* in Castilian is *s-* not *z-*.

Gmc. $^+skarp-$ 'sharp' and to Gmc. $^+kerb-$ 'to carve'? In principle, there are, of course, even two ways: inserting the $^+k-$ and deleting the $^+k-$. If it were not for $^+kerb-$, one way of connecting $^+sarp-$ and $^+skarp-$ 'sharp' would be assuming a contamination of Vasconic $^+sarpV-$ with IE $^+sker-$ 'to cut' yielding $^+skarpV-$ 'sharp'. However, if one wants to connect $^+kerb-$ 'to carve' as well, the only way seems to be the other way around: to assume a Vasconic form $^+skarpV-$ 'claw' (or 'sharp (object)'). Bq. *zarpa* 'spur' would phonologically derive by way of implementing the constraint

$^+sk > s / \# \underline{\quad}$

of the preceding section. Gmc. $^+skarp-$ 'sharp' would preserve the original shape of the root. The $^+sarp-$ forms of German and Dutch, Balto-Slavic, Latin, and Greek would have to be explained as borrowings from varieties of Vasconic that had already implemented the above constraint at the times of contact. The $^+kerb-$ group would have to be interpreted as another outcome of the Vasconic process of cluster simplification, leaving the plosive rather than the fricative.⁶¹

The only other case of an *sk-/s-* variation in Germanic is that of the preterite-present Goth. OHG *skal/skulom*, ON *skal/skulom*, OS *scal/sculun*, OFris. *skel, skil*, OE *sceal/sculon* 'owe(s), shall'. Kluge/Seebold (1995: s.v. *sollen*) consider the meaning 'to owe' original (cf. also Gm. *Soll* n. 'debit', *schuld* 'at fault, to blame', *Schuld* f. 'guilt', *Schulden* pl. 'debts', *schulden* 'to owe') and compare Lith. *skelėti* 'to owe'. "Weitere Herkunft unklar." The OED (s.v. *shall*) says, "Outside Teut. the only certain cognates are Lith. *skelėti* to be guilty, *skilti* to get into debt, *skolà* debt, guilt, OPrussian *skallisnan* (acc.) duty, *skellānts* guilty, *po-skulit* to admonish."

This verb has a variant *saln/suln*. Braune/Eggers (1987: § 146 n. 4 and § 374) and Kluge/Seebold (1995: s.v. *sollen*) assume the *sk-* forms as basic and explain the loss of *-k-* as a

⁶¹Such different outcomes of constraint implementation are known from Finnish loan-word adaptation of *st-* first as *s-*, later as *t-* (cf. Vennemann 1995a: 85, n. 129, with reference to Koivulehto 1994: 85). An example of *st-* → *s-* is Gmc. $^+stōda-$ (OHG *stuot*, NHG *Stute, Gestüt*, OSwed. *stōp*) 'mare' → Finn. (old, dial.) *suota* 'herd of mares in heat, herd of horses'; an example of *st-* → *t-* is Gmc. $^+stangō$ (Gm. *Stange*, ON *stōng*) → Finn. *tanko* 'pole, stick, staff, rod'.

consequence of weak sentence accent; Kluge/Seebold refer to a similar development in English. About this parallel as well as proposed explanations, the OED says:

The northern English dialects (including Sc[otch]) have a form *sal*, pa. tense *suld*, with initial *s* instead of *sh*. This does not occur in the remains of ONorthumbrian, but first appears in the 13th c. It is remarkable that a similar form, with *s* irregularly representing OTeut. *sk*, existed as a dialectal variant in OHG. (*sal, sol, sulun*) and OFris. (*sal, sel*), and has ousted the regular form in Ger. (*soll, sollen*) and Du. (*zal, zou*). Some scholars regard the *s* form as representing an OTeut. variant, originating from the euphonic dropping of *k* in inflexional forms like the subjunctive **skli-*. It seems more probable that it was independently developed in the different dialects at an early period, while the *sk-* retained its original pronunciation; in stressless position the *k* might naturally be dropped, and the simplified initial afterwards extended by analogy to the stressed use.

Prokosch (1939: 191 n. 3) believes that the *-k-* was first lost from *sk-* in zero grade forms where the cluster stood before the nuclear liquid prior to the anaptyxis of *u*, from where the *s-* onset subsequently generalized. Braune/Eggers (1987: § 146 n. 4) support the accentual weakening hypothesis by pointing to the further reduction of the form in Swed. Gutn. *al*.

The OHG variants *scarpf/sarpf* and *skal/sal* (etc.) develop differently in time (Braune/Eggers 1987: § 146 n. 4 and § 374): In the oldest documents *sarpf* dominates; *scarpf* only occurs occasionally, e.g. Hildebrandslied v. 64 (*scarpen scurim: dat in dem sciltim stont*⁶²), becomes a little more frequent in the 10th/11th centuries, and begins to oust *sarpf/sar*f in Middle High German. By contrast, *sal* forms are rare in early Old High German but become dominant toward the end of the 10th century.

Since *sk-* dominates in the Germanic preterite-present from the earliest times and can be related to *sk-* verbs in Baltic, the best assumption seems to be that it represents the original onset.

The limited distribution, however, is another problem: The root only occurs in Germanic and Baltic.⁶³ This makes it likely

⁶²With Low Saxonizing *-p-* for *-ph-*. The alliteration warrants the *sk-* onset.

⁶³Prokosch (1939: 192) connects Lat. *scelus, -eris* n. 'guilt, crime' but

that we are dealing with a prehistoric loan-word, borrowed at a time when the verbal and deverbal formation processes (including the ablaut system) were still strong enough to integrate such a borrowed item.

Judging by its meaning, the item may have been borrowed with a social significance. This is perhaps underlined by a word that is not generally associated with the preterito-present⁶⁴ but may nevertheless belong here, and whose social significance is obvious: Gmc. *+skalkaz* m. 'servant' as in Goth. *skalks*. ON *skalkr* (Nícel. *skálkur* 'rogue'), OHG *scalc* (NHG *Schalk* 'wag, rascal'), OS *scalk*, OFris. *skalk*, OE *sealc* (Engl. *shalk* poet. obs.⁶⁵). "Herkunft unklar" (Kluge/Seebold 1995: s.v.). "No cognates outside Teut. have been found" (OED: s.v. *shalk*). Considering the fact that debts used to be (and still are in many parts of the world) a regular cause of servitude, Gmc. *+skalkaz* m. 'servant' could be derived from the preterite-present.⁶⁶ A way to combine the two items is the assumption of a radical laryngeal in the borrowed word, as proposed by Beekes (1988: 99)⁶⁷, e.g. *+skalH-/skulH-*. This would agree well with Germanic root structure and would explain the ablaut vowel of the plural forms (e.g. OE *sculon* not *scēalon*). It would also explain the doublet with root-final *k*, *+skalk-a-z*, because it often happens that a segment threatened by weakening and loss is saved by strengthening. This apparently happened again when the word was borrowed from Germanic or Baltic into Finno-Ugric as *+kelki-*, *+kelke-* 'to have to, to be to', as in Lapp. *gål'gâ* 'to be to, to have to', Hung. *kell-* 'to have to, to be to, to need; to please'

Walde/Hofmann (1982: s.v.), glossing the word 'Bosheit, Verruchtheit, Verbrechen' ('wickedness, infamy, crime'), consider this connection uncertain, while Ernout/Meillet (1959), glossing the word 'mauvaise action, faute, crime', declare the connection to be "pas davantage".

⁶⁴The proposal was made by von Grienberger (1900: 187). Lehmann (1986: s.v. *skalks*) refers to it but rejects it, without giving reasons.

⁶⁵"The word forms the second element in **marho-skalko-z* 'horse-servant', *marshal* n." (OED: s.v. *shalk*).

⁶⁶Without wanting to press the issue, perhaps the word *shilling* finds an explanation along these lines, too: Goth. *skilliggs*, ON *skillingr* (Icel. also *skildingr*), OHG *scilling*, *skillink*, OS. *scilling* (MLG also *schildink*), OFris. *skilling*, *skilleng*, *schilling*. The original meaning could be that of a unit of compensation or taxation, and thus of what one owes. If correct, this would show the *e*-grade of the verbal root in Germanic.

⁶⁷Beekes assumes the root to be Indo-European, despite its restricted occurrence.

was borrowed from Germanic or Baltic into Finno-Ugric as *+kelki-*, *+kelke-* 'to have to, to be to', as in Lapp. *gâl'gâ* 'to be to, to have to', Hung. *kell-* 'to have to, to be to, to need; to please' (cf. Koivulehto 1991: 69 n. 37). The replacement of a laryngeal (in a wide sense, including not only glottal stops and fricatives but also e.g. velar fricatives) with *k*, especially in contact with *l*, does not only occur by way of sound substitution in loan-words but also as a language-internal development. This phenomenon has been described by Lutz (1991: 43-45), cf. the following examples:

OE *hōh* > StEngl. *hough* /hʊk/ 'hollow of the knee'

OE *hēahfore* > dial. *heckfer* alongside StEngl. *heifer*

Late OE *dweorg*, *dweorh* > ME *dweygh*/*dwerf*/*dwerk*, Engl. dial. *dwerk* alongside StEngl. *dwarf*

Engl. dial. *selk* alongside StEngl. *seal* (zool.)

OE *eolh* > ME *elgh*/*elk*, Engl. *elk*

As for the source of the putative loan-word or loan-complex, one place to look would be Basque again, remembering that initial *+sk-* there would be reduced to *z-* /*s-*/ and that intervocalic *-l-* would appear as *-r-*, after apocope as *-ġ-*, which in turn may be generalized (cf. Michelena 1977: chs. 16, 17). It so happens that the Common Basque word *zor* (*zor-*) means '[Span.] deuda, [Fr.] dette', i.e. 'debt'. It could continue an ancient *+skolV* or *+skulV*, perhaps from an earlier *+skolHV* or *+skulHV* (with a laryngeal *H*, see above).⁶⁸ Cesar (*Bell. Gall.* III. 22) describes a kind of Aquitanian brotherhood of 600 men, locally called *Soldurians* (*Soldurii*), who "owe" their lives to friends, in the sense that they share their friends' joys of life but

⁶⁸Even though *zor* 'debt' is monosyllabic in all contemporary varieties of Basque, there is an older form *zoor* on record whose spelling is in certain other words a sign of hiatus, i.e. of disyllabicity (Michelena 1977: § 5.3). This may show that the equation with the Germanic and Baltic words is wrong. However, it may also be a trace of the way the word was simplified, e.g. *+skolHV* > *+skoHV* (metathesis) > *+skoHolV* (anaptyxis) > *+zoHolV* (onset cluster reduction) > *+zoHorV* > *+zoHor* > *zoor* > *zor*. If a change type employed by Löpelmann (1968: s.v. *zor*) who identifies Bq. *zor* with Span. *socorro* 'help, assistance, support, advance payment', is attested, then one could see the hiatus as a consequence of the way onset clusters of a sibilant and a plosive were simplified: *+skolHV* > *+sokolHV* (by echo anaptyxis) > *+sokolV* > *+sokorV* > *+sokor* > *+sohor* > *zoor* > *zor*. In any event, Löpelmann's own etymology appears weak to me for semantic reasons.

no-one in living memory had ever refused to pay.⁶⁹ Since the name is used by the local people (*illi*), the inhabitants of the city of the Sotiates, in Aquitaine, Cesar's Aquitania, the name is likely to be Aquitanian, and thus Vasconic.⁷⁰ Perhaps then the first part of the name *Soldurii* is the same as Basque *zor* 'debt'.⁷¹

2.7.2. Gm. *Strunk*, Lat. *truncus*, Bq. *unxhü* 'tree stock, trunk'

Gm. *Strunk* 'stalk, stem, tree stump', according to Kluge/Seebold (1995), is only attested in Late Middle High German (*strunc*) and in Dutch (*stronk*). Outside the Germania the only comparison ("vielleicht") is Lith. *strūngas* 'cut back, trimmed'. "Weitere Herkunft unklar." Finn. *runko* 'stem' is a Germanic loan-word.⁷²

For Lat. *truncus*, *-i* m. 'stem, trunk (of a tree, of a human body), block, blockhead', related words can only be found in West Indo-European languages (cf. Walde/Hofmann 1982: s.v.); and considering that in Latin there is an adjective *truncus*, *-a*, *-um* 'mutilated, stripped of one's branches, dismembered' (also a verb *truncāre* 'to maim, mutilate, cut off' which is, however, of more recent attestation than the nominal forms and thus probably not their base), they do not all fit well semantically: Lith. *trenkiù*, *trenkti* 'heftig, dröhnend stoßen' ('to push violently, resoundingly'), *i-tranka* 'Anstoß' ('push'), *trankūs* 'holprig' ('rough, bumpy'), *trankšmas* 'Gedränge'

⁶⁹"... cum DC devotis, quos illi soldurios appellant — quorum haec est condicio, uti omnibus in vita commodis una cum iis fruuntur, quorum se amicitiae dederint, si quid his per vim accidat, aut eundem casum una ferant aut sibi mortem consciant; neque adhuc hominum memoria repertus est quisquam, qui eo interfecto, cuius se amicitiae devovisset, mortem recusaret — ..." (Caesar [1990]: III. 22). Walde/Hofmann (1982: s.v. *soldurii*) gloss the name 'die jdm. durch ein Gelübde verpflichteten, die Getreuen', i.e. 'those obligated (indebted) to someone by a vow, the faithful ones' and report on several doubtful etymological attempts. Ernout/Meillet (1959: s.v. *soldurii*) say 'gardes du corps ou vassaux d'un chef gaulois. Nom donné expressément comme gaulois par César, BG 3, 22, 1), without giving etymologies.

⁷⁰Aquitanian is considered an early regional variety of Basque by Michelena (1954), Gorrochategui (1984, 1987), and Trask (1997: 398-403).

⁷¹From a Basque point of view, the second part of the stem could be a compositional head meaning 'having', cf. e.g. Bq. *hobendun*, *hobenduri* 'guilty' (*hoben* 'fault, blame'), cf. Löpeltmann 1968: s.v. *-duri/-duru*. Perhaps a similar formation existed in Aquitanian, so that the literal meaning of *Soldurii* could have been 'those having (a) debt (to pay)'.

⁷²Compare the more recent Finn. *ressi* 'stress' (borrowed Engl. *stress*) for the reduction of the initial cluster.

(‘throng, thronging’), *trinkis* ‘Anstoß’ (‘push’), *trinka* ‘Haublock’ (‘chopping block’), OPruss. *pertrinktan* ‘verstockt’ (‘obdurate, stubborn’), Welsh *trûch* ‘verstümmelt’ (‘maimed, mutilated’); further, “mit vokalischer und Auslautvariation”, Lith. *strungas* ‘gestutzt’ (‘cut back, trimmed’), *striugas* ‘kurz’ (‘short’), *striubas* ‘kurz’. Possible connections with Gm. *Strunk* and with the family of Gm. *dringen* ‘to penetrate’, *drängen* ‘to push, press, throng’ are mentioned; but Kluge/Seebold (1995: s.vv.) only connect with Gm. *dringen/drängen* Lith. *treñkti*, not Lat. *truncus*. Ernout/Meillet (1985: s.v. *truncus*) simply state, “Étymologie incertaine. Peut-être à rapprocher du groupe de *trux* [‘wild, rough, ferocious, grim, stern’]. Ou bien cf. gr. *τρώω* [‘to rub down, wear out’], etc.? Le lit. *trenkiù* ‘je heurte’ est loin pour le sens.”

This group to me does not give the impression of a bona fide Indo-European word family. In particular, Gm. *Strunk* and Lat. *truncus* are suspicious; on one hand they are semantically and phonologically too close to be separated etymologically, on the other hand the identical final root plosives are in violation of Grimm’s Law. Such partial formal mismatch is in other instances taken as a sign of borrowing. One possibility would therefore be mutual borrowing among the West Indo-European languages, perhaps with influences from similar words, such as the family of Gm. *Stumpf*, Engl. *stump*, also Gm. *Strumpf* ‘sock’, which originally meant ‘tree trunk’ (cf. Kluge/Seebold: 1995: s.v.), and Gm. *stumpf* ‘blunt, dull’ (originally ‘maimed, mutilated’, cf. Kluge/Seebold 1995: s.v.), and Gm. *Stubben*, a Low German loan.⁷³ Another possibility is that the entire complex was borrowed from another language. In view of the limitation of the complex to the West Indo-European languages a likely source would be Vasconic.

If we assume PVasc. +*trunku-* or +*strunku-* ‘stem, trunk of a tree (etc.)’, application of the sound changes leading to the phonotactic restrictions of word-initial syllable heads in Basque would yield *unku* in the Eastern dialects, *ungu* in the Central and Western dialects (Michelena 1977: § 18.9). First, the word-initial cluster +(*s*)*tr-* has to be reduced to the least consonantal

⁷³As a matter of fact, if it were not for Kluge/Seebold’s “Weitere Herkunft unklar”, I would have guessed, in the light of the late attestation of *Strunk* and the appearance of Lat. *truncus* as OFr. *tronc* in the medieval languages (cf. OED: s.v. *trunk*), that *Strunk* is simply OFr. *tronc*, contaminated with Gm. *Strumpf/Stumpf*.

member (cf. sec. 2.6 above), which is *r*; second, word-initial *r* in the resulting ⁺*runku-* has to drop (cf. sec. 2.2.1 above). It so happens that de Azkue (1984) lists a word *unxhü* '[Span.] tronco de arbol, [Fr.] tronçe ou tronc d'arbre', i.e. 'trunk of a tree', but only for the Zuberoan (Souletin) dialect of the French Basque country.⁷⁴ I do not know an etymology for this Basque word. Surely Löpelmann's (1968: s.v.) identification of it as borrowed Lat. *truncus* is problematical. Word-initial clusters of voiceless plosive plus *r* are, to my knowledge, never simply deleted in the loan transfer but treated in one of the following ways: (1) In the earlier history of Basque, such *Pr-* clusters are adapted by anaptyxis, as shown in sec. 1 above, while the plosive itself is often voiced or deleted, so that the result would be likely to be *durinkhü* or *urinkhü*. (2) In the more recent history of Basque, such *Pr-* clusters are preserved, and indeed there is, in the Basque Country of France, *trontzo*, with a variant form *truntzo*, meaning 'stem, stump, log, block', which is borrowed Prov. *trons* (cf. Fr. *tronçe*), from a variant ⁺*truncus* of Lat. *truncus* that is reconstructible on the evidence of the Romance words (cf. also Cat. *trunch*, Span. *tronzo*, Port. *troncho* 'mutilated, cut back, trimmed').⁷⁵ Perhaps then Eastern Bq. *unxhü* is indeed the native remnant of an old Vasconic word which was borrowed in prehistoric times into the West Indo-European contact languages and later reborrowed from the descendants of one of these, the Romance languages, into the one remaining Vasconic language, Basque.⁷⁶

⁷⁴The aspiration of voiceless plosives is a frequent feature of the Eastern dialects (cf. Michelena 1977: § 11.1).

⁷⁵The derived word *trontxoin* 'plate' (Löpelmann 1968) is also mentioned by de Azkue (1984) for one community in the Lapurdian dialect. For the Basque Country in Spain, de Azkue only lists the rare *trontzazerra* 'big saw' and its dissimilated variant *trontazerra*. (I have used the standard sibilant spellings <z, x> for /s, š/.)

⁷⁶If the regional *zungo* 'stem of a tree' (Löpelmann 1968: s.v. *zungo*), 'anvil-block' (de Azkue 1984: s.v. *zungo*) is a compound of *zur* 'wood' and *ungo* (= *unxhü*) 'stem' (Löpelmann) rather than a variant of *zundo* 'chopping block' (de Azkue, with question mark), the word must also have existed in dialects that voice plosives after nasals. But there is no entry *ungo* 'stem' either in Löpelmann 1968 or in de Azkue 1984.

3. The etymological identity of Lat. *grandi-* and Bq. *handi* 'big'

The conclusion to be drawn from the preceding discussion is twofold. First, the examples treated show that in doing etymology, more may have to be considered in the giving and the taking languages than simple segmental changes or assimilations: Phonotactic restrictions may alter languages participating in contact to a degree that loan equations become hard to recognize. This is in principle well known, namely for those cases in which the phonotactic restrictions have been studied on independent grounds and only need to be made fruitful in lexical comparisons. The case on hand is different because the phonotactic restrictions — rather: the changes leading to the relevant restrictions — had *not* been studied on independent grounds, and for a simple reason: There exists no evidence for these changes independently of the presumed loan equations. The procedure looks like a vicious circle: I have reconstructed phonotactic change on the evidence of presumed loan equations, and I go on to explain the same loan equations in the reconstructed framework of changes. But on closer examination the procedure turns out to be normal science: Data and general principles have been used to construct a specific theory. Whether the data are solid enough to call for a systematic account is, as always, disputable; there is always the easy way of calling them accidental. General principles of phonotactic change are likewise subject to dispute; there is always the easy way of calling them into question when their application leads to unexpected results. Since I trust the data and the principles, I am ready to continue working with the theory.

Second, by applying the theory to the initial problem, the question of how the phonological and semantic similarity of Lat. *grandi-* 'big' and Bq. *handi* 'big' has to be explained, it yields a straightforward answer. Latin has not undergone any changes eliminating or creating initial *gr-*; for Basque we do not specifically know of any relevant changes, but we do know that it does not tolerate in its traditional vocabulary either *gr-* or *r-*, and we know in a general way that such phonotactic restrictions result from phonological changes eliminating earlier structures that are not in accordance with the developing constraints. Therefore, the theory suggests that we posit a common etymon ⁺*grandi-* for both Lat. *grandi-* and Bq. *handi*. Since Lat. *grandi-* is not an inherited Indo-European word, and since the

assumption of an unknown third language as the giving language would be a violation of a general principle of scientific simplicity (Occam's razor), the common etymon will be attributed to Vasconic. In other words, I consider Lat. *grandi-* 'big' a Vasconic loan-word. The relationship with Bq. *handi* 'big' can be summarized in the following formulas:

PVasc. **grandi-* 'big' → Lat. *grandi-* 'big' (*grandis, grande* etc.)

PVasc. **grandi-* 'big' > Pre-Bq. **randi-* 'big' > Bq. *handi* 'big'

It might be objected to this reconstruction, as well as to some presented in the above sections, that the postulated phonotactic changes contradict the evidence of the treatment of Latin and Romance loan-words in Basque, in the case on hand the evidence in favor not of loss but in favor of anaptyxis presented in section I above, which I repeat here for convenience:

GRATIA → *garazia* 'grace'

GRANU → *garau(n)* 'grain'

CRUCE → *kurutze* 'cross'.

PVasc. **grandi-* 'big' should therefore, according to this objection, yield Bq. *garandi* rather than *handi*.⁷⁷ This objection is not valid. First, we are dealing with two different periods in the development of Basque; mechanisms of change at work in a language are not always the same. Second, loan-word adaptation is not the same thing as sound change; the motivations and the mechanisms involved are quite different. Sound change is phonetic simplification of words that are, in general, known to all speakers, so that questions of comprehension and identification in the course of the change play a secondary role. By contrast, loan-word adaptation adjusts words to a language that are in general not yet known to all speakers, so that identifiability plays a greater role than simplicity. Therefore, whereas sound change aims at structural simplicity, loan-word adaptation aims at the preservation of as much of the input as is adjustable to the structure of the

⁷⁷I am grateful to Philip Baldi for drawing my attention to this possible source of misunderstanding.

receiving language; indeed, sometimes even the structure of the receiving language is violated in the desire to keep the new word maximally intact.⁷⁸

A single example will suffice to make my point. Modern English has a constraint prohibiting words to begin with the cluster *kn-*:

**kn-*

This constraint is strong enough to prevent the cluster from occurring both in inherited words, i.e. words that were in the language before the constraint developed, and in loan-words post-dating the development of the constraint, but with an important difference: The onset cluster *kn-* is treated anaptyctically in loan-word adaptation:

kn- → *kən-*

Knut → *Canute* [kə'nu:t]

Knesset → [kə'nɛsɪt]

But in the development of the phonological constraint, the cluster was treated by weakening leading to complete loss of the plosive⁷⁹:

kn- > *n-*

knife [knɪf] > [naɪf]

knee [knē] > [ni:]

The loss of the Vasconic onset cluster ⁺*gr-* may therefore have proceeded in the following manner:

⁺*gr-* > ⁺*r-*

⁺*r-* > *h-* (> zero)

Here one could again object that in loan-words initial *r-* is not lost in Basque but treated prosthetically:

⁷⁸A recent paper in this domain, with references to earlier statements, is Singh 1995.

⁷⁹The details of the implementation of the constraint in the phonotactic history of English are complicated; see Lutz 1991: 238-244.

ROSA → *arrosa* 'rose'

REGE → *errege* 'king, queen bee'

FRAGA → *arraga, arrega* 'strawberry'⁸⁰

But again it does not follow that prosthesis also was the change by which word-initial *r*- was ousted from the language. The following development is a possible alternative:

⁺*r*- > ⁺*r̥*- > *h*- (> zero)

Desonorization of word-initial *r*- is recognizable in the Greek writing of *r*- with spiritus asper, which is reflected in the modern spelling with *rh*- (e.g. *rhythm*). And the change of word-initial *r*- into mere aspiration is known from Brazilian Portuguese, e.g. *Rio* [hiu].⁸¹

4. A note on WGmc. ⁺*grauta*- 'big', English *great*

I mentioned in section 1 above that authorities were divided between considering *handi* or *haundi* as the basic variant of the Basque word for 'big'. I there accepted Michelena's decision in favor of *handi*. This was probably a good decision. Larry Trask (University of Sussex, Brighton) wrote in a letter to the Nostratic List on the Internet of 9 March 1998:

The difficulty with taking Basque (*h*)*aundi* as older than (*h*)*andi* is that the shorter form is far better attested in the earliest literature (from 1262, in fact), while the longer form is only well attested in comparatively recent times. The form with /au/ is more or less exclusively western, but the earliest western sources do not show this form.

I naturally do not want to press the point, but it may be of interest that a decision in favor of *haundi* as basic and of *handi*

⁸⁰FRAGA is really the plural of Lat. *frāgum*. Since Basque does not possess an *f* in its native vocabulary, *f* in loan-words receives special treatment: it is lost before consonants; e.g. FLORE → *lore* 'flower, blossom'. The *r* of FRAGA, thus becoming word-initial, is treated prosthetically, according to the rule for word-initial *r*-. Recent loans tend to preserve their *f*; e.g. *faxista* 'fascist', *flauta* 'flute', *fruta* 'fruit'.

⁸¹I would like to thank Peter Auer for kindly reminding me of this pronunciation.

as a simplified variant would have its virtue too. There exists another West Indo-European word meaning 'big' and lacking a plausible etymology: E *great*.⁸² This word is restricted to West Germanic (OE *græt*, OFris. *grāt*, OS *grōt*, OHG *grōz*), and it is given unconvincing etymologies—or none at all—in the OED (s.v. *great*), by Onions (1966: s.v. *great*, "of unkn. origin"), Kluge/Seebold (1995: s.v. *groß*), and Pfeifer (1989: s.v. *groß*). Thus Gmc. **grauta-* 'big' may well be a loan-word, and then a connection with Lat. *grandi-* 'big' would be desirable for three reasons: First, the phonological structure is similar; second, the meaning is identical; third, both Lat. *grandi-* and West Gmc. **grauta-* gradually superseded the inherited Indo-European word for 'big' in historical times, *magnus* in Latin and **mikil-* (OE *micel*, OHG *mihhil*, etc.) in West Germanic. A Proto-Vasconic etymon **graundi-* 'big' would yield rather good results, namely with phonotactically motivated simplifications of the over-heavy syllable *+/graun/* in *+/graun.di-/* by either loss of the glide or loss of the nasal from the coda; for West Germanic, a change of declension class from *i*-stem to *a*-stem and partial application of Grimm's Law would have to be posited, both of them innocuous assumptions.⁸³

PVasc. **graundi-* 'big' > Pre-Bq. **raundi-* > Bq. *haundi* > *handi* 'big'

PVasc. **graundi-* 'big' → Lat. *grandi-* 'big'

PVasc. **graundi-* 'big' → Pre-WGmc. **graundi-* > WGmc. **grauta-* 'big'

Needless to say that even if Bq. *haundi* must be dismissed for etymological purposes, there remains enough similarity between Bq. *handi* / Lat. *grandi-* 'big' and WGmc. **grauta-* 'big' to keep this set of words in mind as a problem for further research.

⁸²I would like to thank Cristian H. Kollmann (Innsbruck) for drawing my attention, in a discussion, to the similarity of Lat. *grandi-* with Gmc. **grauta-* 'big' and for suggesting that they might be the same word.

⁸³First, the *i*-stem adjectives, preserved in Gothic, lost their independent status in West Germanic; cf. Campbell (1959: § 654) for Old English and Braune/Eggers (1987: § 246) for Old High German. Even inherited *i*-stem adjectives were transferred in part to the *a*/*ō*-stem classes, even though mostly to the *ja/jō*-stem class in Old High German. So the assumption of a stem class shift would fit well into the grammatical tendencies of West Germanic. Second, inconsistencies in the treatment of plosive grades in loan transfer are a matter of course.

5. Conclusion

In this paper I have proposed a number of Vasconic etymologies for hitherto insufficiently explained West Indo-European words. The combining characteristic of all examples is the reliance on linguistic interpretations of existing or reconstructed phonotactic constraints and tendencies of Basque. Since we know from languages with longer periods of attestation than Basque that phonotactic constraints come and go, we may—or rather: have to—assume for every given constraint an earlier stage of the language at which that particular constraint was not yet operative. Thus, we may tentatively reconstruct, for any given item of the language under comparison, ancestral forms that yield the attested form by application of the relevant constraints. If one of these reconstructed forms agrees well, both phonologically and semantically, with an unexplained item in a contact language, then, by interpreting that item as a loan-word from the language under comparison, we may select that reconstructed form as the etymon of the attested word.

It goes without saying that this line of etymological reconstruction, even assuming the general validity of my theory of the linguistic prehistory of Europe north of the Alps, has to be followed with the utmost care and judiciousness. E.g., assuming for Proto-Vasconic a range of word-initial clusters similar to that of Indo-European, Bq. *handi* 'big', looked at in isolation, could at least be reconstructed with the following anlauts (not to mention its present one): $^+r-$, $^+br-$, $^+dr-$, $^+gr-$, $^+pr-$, $^+tr-$, $^+kr-$, $^+spr-$, $^+str-$, $^+skr-$. Also single plosives are often lost in Basque, especially $^+t-$, $^+k-$, $^+g-$, and initial $^+d-$ must at some prehistoric time have been deleted, which further increases the reservoir of possible etyma. The choice of PVasc. $^+grandi-$ 'big' from this set of possibilities as the ancestral form of Bq. *handi* 'big' is thus entirely motivated by the comparison with Lat. *grandi-* 'big'.

It need hardly be stressed that for this method to be effective, the requirements of semantic and phonological fit will have to be extremely strict, approaching identity (excepting the focused phonotactically motivated difference and subsequent language-specific developments), and the meanings themselves will have to be specific, and the phonological agreement

substantial, namely of some length.⁸⁴ I hope the reader will find this strictness applied in the cases treated in this paper.

References

- Agud, Manuel — Antonio Tovar
 1989ff *Diccionario etimológico vasco*. Donostia / San Sebastian: Diputación Foral de Gipuzkoa. (So far 5 fasc., A - Iloza.)
- Allen, W. Sydney
 1974 *Vox Graeca: A guide to the pronunciation of Classical Greek*. 2nd ed. Cambridge: Cambridge University Press.
- Arbeláiz, Juan José
 [1978] *Las etimologías vascas en la obra de Luis Michelena*. Tolosa: Kardaberaz.
- Aulestia, Gorka — Linda White
 1990 *English-Basque dictionary*. Reno: University of Nevada Press.

⁸⁴In the case of Lat. *grandi-* 'big' and Bq. *handi* 'big', the agreeing phonological substance is a sequence of four phonemes, in the case of Gmc. **rugi-* 'rye' and Bq. *ogi* 'bread, wheat, cereal' a sequence of three. For shorter sequences the number of possible comparisons increases quickly, so that connections become too easy to be convincing. To show just one set of this kind: De Azkue (1984: s.v.) lists the word form (*h)atz* (*tz* represents a monophonemic affricate) with eleven meanings. Among these, there exist interesting correspondences with West Indo-European words that are themselves in part of unclear origin. E.g., de Azkue's meaning no. 8 is 'casta, raza', i.e. 'caste, race', listed for only a single community, Ligi (Licq) in the Zuberoan dialect, and thus perhaps a residual form, elsewhere supplanted by other words: *etnia, enda, arraze* (from Span. *raza*); Ital. *razza*, Span. *raza*, Fr. *race*, etc. (hence E *race*, G *Rasse*) are attested late (e.g. Fr. *rasse* a. 1512) and are, according to the OED, "of obscure origin". So this could be a Vasconic etymon, not unlikely for a language of herdsmen, but then two agreeing phonemes is not much. De Azkue's meanings no. 1 'trace, vestige', no. 2 'finger', no. 3 'foot of an animal, paw', no. 4 'scabies', no. 6 'itch' (*(h)atz egin* 'to scratch'), no. 9 'axle, wheel-shaft', no. 10 'sign, mark', no. 11 'rank, position' are in part identified with each other and with no. 8 by Löpelmann and Agud/Tovar. I consider some of these identifications doubtful and would, tentatively, consider the fact that in the West Indo-European languages several phonologically comparable words (some of doubtful origin!) carry individual meanings from this set, e.g. *caste* (which is not Indo-European, even if it derives from Lat. *castus, -a, -um* 'pure'), *trace* (Lat. *trahere*, Engl. *drag* etc. are not likely to be Indo-European), Gm. *kratzen* and Engl. *scratch* ("Herkunft unklar"), Gm. *Tatze* 'paw', Gm. *Pratze* 'paw' (from Ital. *braccio* 'arm', ultimately from Gk. *brachion* 'upper arm', which is likewise unexplained). Perhaps future research will allow a less pessimistic outlook on these strange similarities.

- Azkue, Resurrección María de
 1984 *Diccionario vasco - español - francés*. Reprint (with an introduction by Luis Michelena) [1st ed. 1905, 2nd ed. (reprint) 1969]. Bilbao: Euskaltzaindia.
- Bammesberger, Alfred (ed.)
 1988 *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Carl Winter.
- Beck, Heinrich, et al. (eds.)
 1994 *Reallexikon der Germanischen Altertumskunde*, 2nd ed. Vol. 9, fascicles 1/2. Berlin: Walter de Gruyter.
- Beekes, R. S. P.
 1988 "Laryngeal developments: A survey", in: Bammesberger (ed.) 1988, 59-105.
- Braune, Wilhelm
 1987 *Althochdeutsche Grammatik*, 14th ed. by Hans Eggers. Tübingen: Max Niemeyer.
- Caesar, Caius Julius
 [1990] *Der Gallische Krieg*, ed. Otto Schönberger. Munich: Artemis.
- Campbell, Alistair
 1959 *Old English grammar*. Oxford: Clarendon Press. [1969 reprint.]
- Chantraine, Pierre
 1968-1980 *Dictionnaire étymologique de la langue grecque: Histoire des mots*. 4 vols. Paris: Klincksieck. [Reprint in 2 vols, A-K 1990, A-Ω 1984.]
- Corominas, Joan — José A. Pascual
 1980-1991 *Diccionario crítico etimológico castellano e hispánico*. 6 vols. Madrid: Gredos. [1992 reprint.]
- Cowan, H. K. J.
 1984 "The affinities of non-Celtic Pictish", *Leuvense Bijdragen* 73: 433-488.
- Devoto, Giacomo
 1968 *Avviamento alla etimologia italiana: Dizionario etimologico*. 2nd ed. Florence: Felice Le Monnier. [1995 reprint.]
- Dodgson, E.-S.
 1903 "Analyse des formes verbales de l'Évangile de S. Marc, dans le Nouveau Testament basque de Liçarrague (La Rochelle, 1571)", *Revue de Linguistique et de Philologie Comparée* (Paris) 36: 314-337.
- Dunkel, George E., et al. (eds.)
 1994 *Früh-, Mittel-, Spätindogermanisch: Akten der IX. Fachtagung der Indogermanischen Gesellschaft vom 5. bis 9. Oktober 1992 in Zürich*. Wiesbaden: Dr. Ludwig Reichert.

The Journal of Indo-European Studies

- Ernout, A. — A. Meillet
 1959 *Dictionnaire étymologique de la langue latine: Histoire des mots*. 4th ed. Paris: Klincksieck. [1994 reprint of the 1985 revision.]
- Fahs, Achim
 1989 *Grammatik des Pali*. 2nd ed. Leipzig: VEB Verlag Enzyklopädie.
- Feist, Sigmund
 1939 *Vergleichendes Wörterbuch der gotischen Sprache: Mit Einschluß des Krimgotischen und sonstiger zerstreuter Überreste des Gotischen*. 3rd ed. Leiden: E. J. Brill.
- Frisk, Hjalmar
 1973-1991 *Griechisches etymologisches Wörterbuch*. 2nd/3rd, unchanged ed. Heidelberg: Carl Winter.
- Gamillscheg, Ernst
 1969 *Etymologisches Wörterbuch der französischen Sprache*. 2nd ed. Heidelberg: Carl Winter.
- Glaser, Elvira — Michael Schlaefer (eds.)
 1997 *Grammatica ianua artium: Festschrift für Rolf Bergmann zum 60. Geburtstag*. Heidelberg: C. Winter.
- Goosens, Jan
 1974 *Historische Phonologie des Niederländischen* (= Sprachstrukturen, A.2). Tübingen: Max Niemeyer.
- Gorrochategui Churrua, Joaquin
 1984 *Onomastica indigena de Aquitania*. Salamanca: Universidad de Salamanca, Secretariado de Publicaciones.
 1987 "Die vorrömische Onomastik Aquitaniens", *Beiträge zur Namenforschung, Neue Folge* 22: 343-355.
 1995 "Basque and its neighbors in antiquity", in: Hualde et al. (eds.) 1995, 31-63.
- von Grienberger, Theodor
 1900 *Untersuchungen zur gotischen Wortkunde* (= Sitzungsberichte der königlichen Akademie der Wissenschaften, Phil.-hist. Klasse, 142/8). Wien: Gerold.
- Hickey, Raymond — Stanislaw Puppel (eds.)
 1997 *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th Birthday*. Berlin: Mouton de Gruyter.
- Hualde, José Ignacio, et al. (eds.)
 1995 *Towards a history of the Basque language*. Amsterdam: John Benjamins.
- Hubschmid, Johannes
 1960 *Mediterrane Substrate: Mit besonderer Berücksichtigung des Baskischen und der west-östlichen Sprachbeziehungen* (= Romanica Helvetica, 70). Bern: A. Francke.

Kluge, Friedrich

- 1995 *Etymologisches Wörterbuch der deutschen Sprache*, 23rd ed. by Elmar Seebold. Berlin: Walter de Gruyter.

Köbler, Gerhard

- 1994 *Taschenwörterbuch des althochdeutschen Sprachschatzes* (= Uni-Taschenbücher, 1823). Paderborn: Ferdinand Schöningh.

Koivulehto, Jorma

- 1991 *Uralische Evidenz für die Laryngalthorie* (= Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Phil.-hist. Klasse 566). Wien: Verlag der Österreichischen Akademie der Wissenschaften.
 1994 "Finland", "I. Sprachliches". In: Beck 1994.

Koivulehto, Jorma — Jarmo Korhonen (eds.)

- 1995 *Proceedings of the Öhmann-Symposium, University of Helsinki, 23-24 September 1994* (= Der Ginkgobaum: Germanistisches Jahrbuch für Nordeuropa 13). Helsinki: University of Helsinki, Department of German.

Körper-Grohne, Udelgard

- 1987 *Nutzpflanzen in Deutschland: Kulturgeschichte und Biologie*. 3rd, unchanged ed. 1994. Stuttgart: Theiss.

Krahe, Hans

- 1969 *Germanische Sprachwissenschaft*. 7. Aufl. v. Wolfgang Meid. Bd. I: *Einleitung und Lautlehre*, Band II: *Formenlehre*. Berlin: Walter de Gruyter.

Kwon, Kyung-Keun

- 1996 *Phonologische Entwicklungen im Koreanischen und Deutschen*. Unpubl. Diss., University of Munich.

Lehmann, Winfred P.

- 1986 *A Gothic etymological dictionary, based on the third edition of Vergleichendes Wörterbuch der gotischen Sprache by Sigmund Feist*. Leiden: E. J. Brill.

Leizaola, Fermin, et al.

- 1990 *Euskalerriko Atlas Etnolinguistikoa (EAEL)*, issued by Aranzadi Zientzia Elkarte, Etnologia Mintegia. Part I.2. Donostia / San Sebastian: MUNIBE.

[Liddell/Scott]

- 1889 *An intermediate Greek-English lexicon founded upon the seventh edition of [Henry George] Liddell and [Robert] Scott's Greek-English lexicon*. Oxford: Clarendon Press. [1968 impression.]

Löpelmann, Martin

- 1968 *Etymologisches Wörterbuch der baskischen Sprache: Dialekte von Labourd, Nieder-Navarra und La Soule*. 2 vols. Berlin: Walter de Gruyter.

The Journal of Indo-European Studies

- Lühr, Rosemarie
1988 *Expressivität und Lautgesetz im Germanischen*. Heidelberg: Carl Winter.
- Lutz, Angelika
1991 *Phonotaktisch gesteuerte Konsonantenveränderungen in der Geschichte des Englischen* (= Linguistische Arbeiten, 272). Tübingen: Max Niemeyer.
- Marold, Edith — Christiane Zimmermann (eds.)
1995 *Nordwestgermanisch* (= Reallexikon der germanischen Altertumskunde, Ergänzungsbände, 13). Berlin (Walter de Gruyter).
- Meyer-Lübke, W.
1935 *Romanisches etymologisches Wörterbuch*. 3rd ed. Heidelberg: Carl Winter. [6th ed. (reprint of 3rd ed.) 1992.]
- Michelena, Luis
1954 "De onomastica aquitana", *Pirineos* 10: 409-458. Reprinted in: Michelena 1985: 409-445.
1977 *Fonética histórica vasca* (= Anejos del Anuario Seminario de Filología Vasca "Julio de Urquijo", IV, Serie "Obras completas de Luis Michelena", I). 2nd ed. [1st ed. 1961.] Donostia / San Sebastián: Diputación Foral de Gipuzkoa. [1990 reprint.]
1985 *Lengua e historia*. Madrid: Paraninfo.
- [OED]
1993 *The Oxford English dictionary*, 2nd ed., on compact disc for the Apple Macintosh. Oxford: Oxford University Press.
- Onions, C. T.
1966 *The Oxford dictionary of English etymology*. Oxford: Clarendon Press. [1969 reprint.]
- Pfeifer, Wolfgang, et al.
1989 *Etymologisches Wörterbuch des Deutschen*. 3 Bde. Berlin: Akademie-Verlag.
- Pokorny, Julius
1989 *Indogermanisches etymologisches Wörterbuch*. 2 vols. (2nd vol.: Index by Harry B. Partridge). 2nd ed. Berne: A. Francke. [1st ed. 1959.]
- Rix, Helmut
1976 *Historische Grammatik des Griechischen: Laut- und Formenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Singh, Rajendra
1995 "Rethinking the relationship between phonology and loan phonology", *Folia Linguistica* 29: 395-405.

Szemerényi, Oswald

- 1989 *An den Quellen des lateinischen Wortschatzes* (= Innsbrucker Beiträge zur Sprachwissenschaft, 56). Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Thurneysen, Rudolf

- 1946 *A Grammar of Old Irish*. Dublin: Dublin Institute for Advanced Studies.

Trask, R. L.

- 1997 *The history of Basque*. London: Routledge.

Vennemann, Theo

- 1988 *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.
- 1993a "Zur Erklärung bayerischer Gewässer- und Siedlungsnamen", *Sprachwissenschaft* 18: 425-483.
- 1993b "Die ubischen Matronae Albiahenae und der gallo-römische Mercurius Cimiacinus: Mit einem Anhang über den Weißenburger Mercurius Pro[i]tium", *Beiträge zur Namenforschung, Neue Folge* 28, 271-300.
- 1994a "Die mitteleuropäischen Orts- und Matronennamen mit *f*, *p*, *h* und die Spätphase der Indogermania", in: Dunkel et al. (eds.) 1994, 403-426.
- 1994b "Linguistic reconstruction in the context of European prehistory", *Transactions of the Philological Society* 92: 215-284.
- 1995a "Etymologische Beziehungen im Alten Europa", in: Koivulehto — Korhonen (eds.) 1995, 39-115.
- 1995b "Morphologie der niederrheinischen Matronennamen", in: Marold — Zimmermann (eds.) 1995, 271-299.
- 1996 "German *Eisvogel*, Greek *halkyón*, English *alder*: A study in Old European etymology", *Interdisciplinary Journal of Germanic Linguistics and Semiotic Analysis* 1 (1996), 113-145.
- 1997a "Some West Indo-European words of uncertain origin", in: Hickey — Puppel (eds.) 1997, 879-908.
- 1997b "Der Kastalische Quell, die Gastein und das Rätische: Mit einem Anhang zu Cassandra und Kastianeira", in: Glaser — Schlaefer (eds.) 1997, 479-503.

de Vries, Jan

- 1962 *Altnordisches etymologisches Wörterbuch*. 2nd ed. Leiden: E. J. Brill. [3rd ed. (reprint) 1977.]

Walde, A. — J. B. Hofmann

- 1982 *Lateinisches etymologisches Wörterbuch*. 2 vols. and index. 5th ed. Heidelberg: Carl Winter.

Watkins, Calvert

- 1985 *The American Heritage Dictionary of Indo-European Roots*. Boston, Mass.: Houghton Mifflin. [Comprised by the electronic Macintosh version of the *American Heritage Dictionary* of 1992.]

The Journal of Indo-European Studies