

Inheritance versus lexical borrowing: a case with decisive sound-change evidence.

In an earlier posting I raised the question whether sound change ever helps us distinguish between inherited and borrowed words in interesting cases. There are at least a few such cases; here is one I happen to know about.

There is a family of words relating to political leadership attested in widely separated languages of the non-Anatolian, non-Tocharian clade of the Indo-European family, as follows.

Sanskrit: *rā́t* (Rigveda only) and *rājā́* (n-stem) ‘king’; *rājñī́* ‘queen’; *rājyám* ‘lordship’ (once in the Rigveda, scanned as three syllables).

Latin: *rēx*, *rēg-* ‘king’; *rēgīna* ‘queen’; *rēgnum* ‘kingship, kingdom’.

Celtic: Gaulish *-rīx* in personal names (*Dumno-rīx* ‘World-king’, *Ver-cingeto-rīx* ‘Super-hero-king’, etc.) and pl. *-rīges* in tribal names (*Bitu-rīges* ‘World-kings’, etc.);

British Celtic *-rīx* in the personal name *Votepo-rīx* ‘The-king-is-his-refuge’;

Old Irish *rí*, *ríg-* ‘king’; *rígain* ‘queen’; *ríge* ‘kingship’.

From these cognates it is easy to reconstruct a preform **rēg-* ‘king’ (the ancestor of all the masculine nouns except the Sanskrit n-stem *rājā́*, which was probably backformed to ‘queen’); from the Vedic and Irish cognates it is equally easy to reconstruct **rējñih₂* ‘queen’, and it seems likely that Latin *rēgīna* has somehow been remodelled from the same preform. (Of course in the last common ancestor of these languages ‘king’ must have meant something very different in practice from what it meant in mediaeval Europe, for example!) The Irish noun reflects a Proto-Celtic **rīgiom* (the ending is reconstructable because the OIr. noun is a neuter io-stem); since the corresponding Sanskrit noun appears to be an archaic relic, we might hazard reconstructing a preform **rējíóm* ‘kingship’, though the formation is commonplace enough that the two nouns might have been created independently. Latin *rēgnum* is clearly an independent innovation.

There are also some Germanic nominals which appear to belong to the same family of words, as follows.

Gothic *reiks* ‘ruler’ < Proto-Germanic **rīk-*;

Goth. *reiki* ‘rule, kingdom’, Old Norse *ríki*, Old English *rīce*, Old Saxon *rīki*, Old High German *rīhhi* < PGmc. **rīkiją* (ending reconstructable because the noun is a

neuter ja-stem in all the daughters);

Goth. *reikeis (attested in weak gen. pl. *reikjane*) ‘exercising authority’, ON *ríkr* (stem *ríkj-*) ‘powerful’, OE *rīce* ‘powerful, rich’, OS *rīki* ‘powerful’, OHG *rīhhi* ‘powerful, rich’ < PGmc. *rīkijaz ‘powerful’ (ending reconstructable because this is a ja-stem adjective in all the daughters).

The PGmc. *k of these forms is not surprising, because that is the expected outcome of PIE *ǵ (or velar *g, for that matter); and we expect *r to survive without change in Germanic. But the vowel of the root doesn’t fit; inherited *ē should have remained *ē in PGmc. and in Gothic, becoming *ā in all the other daughters (with various further developments, especially in OE). The following examples establish that regular sound correspondence (note that the length mark is not written over Gothic long *e* and *o* because there are no corresponding short vowels):

PIE *g^wēn ‘woman’ (OIr. *bé*, Jasanoff 1989) >→ PGmc. *k^wēniz ‘wife’ > Goth. *qens*, ON *kván*, OS *quān*, OE *cwēn*;

PIE *sēmi- ‘half’ (Greek ἡμι- /hɛ:mi-/ , Lat. *sēmi-*) > PGmc. *sēmi- > OHG *sāmi-*, OS *sām-* (OE *sam-* was generalized from position before consonant clusters and disyllabic words, where the vowel was shortened by regular sound change, Luick 1921:187-8);

PIE *seh₁- ‘sow’ (Lithuanian *sėti*, Old Church Slavonic *sėti* ‘to sow’; Lat. perfect *sēvisse* ‘to have sown’), *séh₁m̃ ‘seed’ (Lat. *sēmen*, OCS *sěmę*, Lith. pl. *sėmenys*), collective *séh₁mō >→ PGmc. *sēana ‘to sow’, *sēmō ‘seed’ (reflecting the PIE collective, Jasanoff 2002:35-8) > Goth. *saiān* (with lowering of *ē before a vowel), ON *sá*, OE *sāwan*, OS *sāian*, OHG *sāan*, *sāwen* ‘to sow’, *sāmo* ‘seed’;

PIE *h₂wéh₁- ~ *h₂wéh₁- ‘blow’ (of the wind; Skt. *vāti*, Homeric Gk. ἄρει /áɛ:si/ ‘(wind) blows’) >→ PGmc. *wēidi ‘(wind) blows’ > Goth. *waiip*, OE *wāwep*;

PIE *d^héh₁ti- ~ *d^héh₁téy- ‘act of putting’ (Avestan *zraz-dāti-* ‘belief’, lit. ‘putting faith’; Vedic Skt. *vásu-d^hiti-* ‘bestowal of goods’, Gk. θέσις /t^hésis/ ‘act of putting’) >→ *d^heh₁tí-s > PGmc. *dēdiz ‘deed’ > ON *dáð*, OS *dād*, OHG *tāt*, OE *dāð*; Goth. *missa-deps* ‘misdeed, sin’;

PIE *méh₁ns- ~ *méh₁ns- ‘moon, month’ (Vedic Skt. *mās* ‘moon, month’; Gk. μήν /mɛ:n/, Lat. *mēnsis*, OIr. *mí*, *mís-*, all specialized as ‘month’) → *méh₁nōs ~ *méh₁nos- ~ *m_hns-’ ~ (locative) *m_hnés(-i) (shift from an acrostic to an

amphikinetic accent paradigm; cf. Tocharian B *meñe*, Lith. *mėnuo*, *mėnesis* ‘month’, Latvian *mēness* ‘moon, month’) >→ PGmc. **mēnō* ‘moon’ (n-stem), **mēnōþ-* ‘month’ > Goth. *mena*, ON *máni* (poetic), OS, OHG *māno*, OE *mōna* ‘moon’; Goth. *menops*, ON *mánaðr*, OHG *mānōd*, OE *mōnaþ* ‘month’.

(Other examples with clear PIE pedigrees involve further complications, e.g. because they occur in suffixal syllables or endings.)

So the PGmc. **rīk-* forms must have been borrowed from some other IE language in which **ē* became **ī* by regular sound change, and the only plausible candidate is Celtic. (For the record, the other IE languages in which **ē* regularly became **ī* are the Luvian subgroup, Lydian, and Armenian; the same change happened at least in word-final position in Palaic. On the Anatolian languages see Melchert 1994:217, 263, 311-2, 367; on Armenian see e.g. Schmitt 1981. Contact between Germanic and any of those languages seems very implausible.)

But note that while the **ī* of the PGmc. forms matches the Celtic vowel perfectly, the following **k* does not; it has clearly been affected by part of the Germanic complex of sound changes called “Grimm’s Law”, the part that shifted voiced stops to voiceless stops. (Most of the original voiceless stops had already been shifted to fricatives, so there was only a very limited merger; see Ringe 2008:94-100 for many further examples of both changes and a description of the merger environment.) In other words, Celtic **rīg-* ‘king’ and **rīgiom* ‘kingship’ were apparently borrowed into pre-PGmc. in their Celtic shapes, more or less, and then subsequently became **rīk-* and **rīkiom* by Grimm’s Law. It seems much less likely that PGmc. **k* (after Grimm’s Law had run its course) was judged the best representation of Celtic **g* by Germanic speakers when the words were borrowed; though PGmc. **g* (after Grimm’s Law) was a fricative in most positions, and its pre-Grimm’s Law antecedent was a breathy-voiced stop, the fact that it was voiced should have made it the best representation of Celtic **g* at just about any time. So we have a reconstructable relative chronology:

- 1) **ē* > **ī* in Celtic;
- 2) borrowing of ‘king’ and ‘kingdom’ from Celtic into pre-PGmc.;
- 3) voiced stops > voiceless in pre-PGmc.

(The adjective is a Germanic innovation which must have been created after (2), but still at a fairly early date, because adjectives in **-yo-* ~ **-io-* were no longer productive in

PGmc.) There are other Celtic loanwords in Germanic, recognizable for various reasons, but this is the only set that gives us an interesting relative chronology of changes.

This chronology is startling, because the change of *ē to *ī might not have occurred as early as Proto-Celtic; there seem to be some Hispano-Celtic exceptions in the first long inscription unearthed at Botorrita, and if that finding holds up, the sound change must have spread through an already diversified Celtic dialect continuum, presumably failing to reach northeastern Spain because that area was isolated by the Pyrenees. (For an attempt to interpret the Botorrita inscription by a well-informed specialist see Eska 1989). Yet Grimm's Law was clearly not a late pre-PGmc. sound change: of the forty-odd sound changes whose relative chronology is outlined in Ringe 2008:152, only five must have occurred before Grimm's Law; about twenty others might have (because they are not on lines of inferred relative chronology below Grimm's Law), but it is most unlikely that all of them are earlier. Roughly the same number of sound changes *must* have intervened between Grimm's Law and the PGmc. period. It seems that much of the distinctive phonological development of PGmc. must have occurred fairly late in its prehistory.

A possible alternative is the following. The above argument assumes that the Grimm's Law changes were historically linked as a "chain shift" of stops; but what if they were actually three independent changes? Then the relative chronology of Grimm's Law, Verner's Law, and related changes becomes somewhat different:

- 1) voiceless stops > voiceless fricatives unless an obstruent immediately precedes (GL1);
- 2) voiced stops > voiceless stops (GL 2; must follow 1, which it counterfeeds);
- 3) breathy-voiced stops > voiced fricatives (GL 3);
- 4) voiceless fricatives > voiced fricatives when not word-initial and not adjacent to an obstruent, and the last preceding syllable nucleus is unaccented (VL; must follow 1, which feeds it);
- 5) voiced fricatives > voiced stops immediately following nasals and in some other environments (must follow 3 and 4, which feed it, and 2, which it counterfeeds);
- 6) accent is shifted to the initial syllable of the word (must follow 4, since it destroys the environment for 4).

That gives us a lot more leeway chronologically, because the stops that developed by (5) underwent no further changes in the PGmc. period.

But there is a final piece of evidence which reinforces our initial chronological conclusion. PGmc. **walhaz* ‘foreigner’ (> OE *wealh*, OHG *walah*; pl. ON *Valir* ‘the French’) was clearly borrowed from the Celtic tribal name **Wolkā-*, which appears in Caesar’s Latin as *Volcae*—and it has clearly been affected by the first and earliest part of Grimm’s Law! In fact, there are *no* clear Celtic loanwords in PGmc. that can be shown to postdate Grimm’s Law. There are plenty of pairs of possible loanwords in which PGmc. **b*, **d*, and **g* correspond to Celtic voiced stops, but in virtually every case a preform with a breathy-voiced stop is either certain or possible (and for that reason we can’t determine whether most of these words are loanwords or common inheritances: in both subfamilies breathy-voiced stops became voiced obstruents). There is *only one* pair of words in which a PGmc. voiceless stop corresponds to a Celtic voiced stop, namely PGmc. **brōk-* ‘leggings, breeches’ (> ON *brók*, OE *brōc*, OHG *bruoh*) = Celtic **brāk-* (borrowed into Latin as pl. *brācae*)—and that is the one case for which Germanic rather than Celtic origin has been claimed, on the reasonable grounds that the inflection of the Germanic word is archaic (see e.g. de Vries 1960:71).

There is at least one other probable Celtic loan in Germanic that exhibits the Celtic sound change **ē* > **ī*, and it is interesting for a different set of reasons. Consider the following words meaning ‘iron’:

Goth. *eisarn*, ON *ísarn* (poetic), *járn*, OE *īsern*, *īsen*, *īren*, OS, OHG *īsarn* < PGmc.

**īsarna*; the shorter ON and OE forms have undergone startling sound changes, but their descent from the PGmc. word is not doubtful;

OIr. *īarn*, Gaulish (?) *Īsarno-* (in a place-name recorded in the early middle ages; see

Feist 1939:131 with references) < Celtic **īsarnom*.

The Celtic and Germanic words match perfectly; moreover, this is the only word for ‘iron’ shared by first-order subgroups of IE. (By itself that doesn’t prove that iron was acquired relatively late by speakers of IE languages; but we know from archaeological evidence that it was, and the distribution of words for the metal is consistent with that fact.)

Maybe we can’t completely exclude the possibility that this is a common inheritance of Germanic and Celtic, but the cladistic tree is against it; the split between Celtic and Germanic (or their immediate parents) should have occurred well back in the 3rd millennium BCE at the latest, when no one in northern Europe had any iron. Borrowing

is overwhelmingly likely; borrowing from Celtic into Germanic is also overwhelmingly likely, both because the Celts were closer to sources of iron and because almost all the lexical borrowing between these two groups is from Celtic into Germanic, with very few counterexamples.

Moreover, if the word was originally Celtic, we can hazard an etymology (Cowgill 1986:68 fn. 10). Celtic *īsar-nom can reflect earlier *ēsar-nom, by the sound change discussed above; the latter can reflect *ēsh₂r-no-m, derived from PIE *ésh₂r ‘blood’ (Hittite *ēshar*, Skt. *ásrk*) by lengthening the root vowel (a derivational process called “vṛddhi”) and adding a well-known suffix. As Warren Cowgill pointed out to me some thirty years ago, there are at least two good reasons why iron might be called ‘blood-metal’; the fact that it rusts is one of them.

Bibliography.

- Cowgill, Warren. 1986. “Einleitung.” Cowgill, Warren, and Manfred Mayrhofer, *Indogermanische Grammatik, Band I* (Heidelberg: Winter) 9-71.
- Eska, Joseph. 1989. *Towards an interpretation of the Hispano-Celtic inscription of Botorrita*. Innsbruck: IBS.
- Feist, Sigmund. 1939. *Vergleichendes Wörterbuch der gotischen Sprache*. Leiden: Brill.
- Jasanoff, Jay. 1989. “Old Irish *bé* ‘woman’.” *Ériu* 40:135-41.
- . 2002. “The nom. sg. of Germanic *n*-stems.” Wedel, Alfred, and Hans-Jörg Busch (edd.), *Verba et litteræ: explorations in Germanic languages and German literature* (Newark, DE: Linguatext) 31-46.
- Luick, Karl. 1921. *Historische Grammatik der englischen Sprache*. Leipzig: Tauchnitz.
- Melchert, H. Craig. 1994. *Anatolian historical phonology*. Amsterdam: Rodopi.
- Ringe, Don. 2008. *From Proto-Indo-European to Proto-Germanic. A linguistic history of English, volume 1*. Revised ed. Oxford: OUP.
- Schmitt, Rüdiger. 1981. *Grammatik des Klassisch-Armenischen*. Innsbruck: IBS.
- Vries, Jan de. 1960. *Kelten und Germanen*. Bern: Francke.