1. Introduction

The hybrid origin of the Germanic languages has become a popular object of research in recent scholarship. Germanic by far and by large qualifies as an Indo-European language, sharing most of its morphology and lexicon with related languages elsewhere in Eurasia. With the help of the Comparative Method, these similarities have been isolated and unified into the partial reconstruction of the Indo-European parent language. However, during the course of the twentieth century, it has become increasingly evident that a significant part – by some estimated as much as one third of the Germanic lexicon (cf. Rifkin 2007: 55) – lacks a solid Indo-European background. The question therefore arose where these words come from.

No doubt, part of the non-Indo-European lexicon can be accounted for by internally Germanic derivational processes: any living language, after all, renews itself by coining new terms from already existing roots and suffixes. It is unlikely, however, that all of the purely Germanic words can be explained in this way. It has been argued that many isolated Germanic words could be a residue of an extinct European language – a language that was spoken in prehistoric Europe before the invasion of the Indo-Europeans (cf. Polome 1989). This is the Germanic Substrate Theory. It claims that, when the indigenous Europeans were subdued by Indo-European invaders, they abandoned their language in favor of the immigrant tongue.

The Substrate Theory entails that when the Indo-Europeans settled in the future core of the Germanic linguistic area, they imposed themselves and their language on an indigenous population with very different cultural and linguistic characteristics (cf. recently Rifkin 2007: 57). The original language(s) of this area ultimately disappeared, because its speakers shifted to a form of Indo-European speech, though not without leaving a trace. The language shift did not happen overnight, but was probably completed through a longer period of bilingualism,
perhaps lasting several generations. As a result, the Indo-European dialect, during its evolution into Proto-Germanic, may have become enriched with various indigenous terms for local plants and animals, geographical phenomena, and cultural practices.

Ever since the discovery of the Indo-European language family, there has been an inkling that Germanic was “blended with a very different idiom” (Jones 1798: 423), mainly because the phonology of Germanic is radically different from what is reconstructed for Proto-Indo-European. The Substrate Theory has nevertheless long been regarded a dubious field of research, not least because hardly any falsifiable results were obtained until the most recent decades. This changes with the development of a methodological innovation induced by the findings of among others Polomé (1986; 1989; 1990), Markey (1988), Hamp (1990), Huld (1990), Salmons (1992; 2004), Kuiper (1995), and Schrijver (1997). These linguists were able to show that non-Indo-European words frequently betray themselves by their non-Indo-European form, or by their irregular correspondence with alleged cognates in other Indo-European languages. In this way, it became possible to isolate non-Indo-European lexemes from the Germanic lexicon in spite of the fact that the source language had vanished well before the start of the historical record.

A showcase example of an important non-Indo-European feature is so-called \textit{a}-prefixation. This feature was discovered by Schrijver (1997), who observed a systematic, but non-Indo-European interchange of initial \textit{a}- versus \textit{zero} in a number of European lexical doublets, cf. 1) OHG \textit{amsala} ‘blackbird’ < \textit{*amsl-}: Lat. \textit{merula} < \textit{*mesal-}, 2) Gal.Lat. \textit{alauda} ‘lark’ < \textit{*alaw-}: OE \textit{læwerce} < \textit{*alaiw-} and 3) OHG \textit{aruz} ‘ore’ < \textit{*arud-}: Lat. \textit{raudus} < \textit{*raud-}. On the basis of alternations like these, Schrijver concluded that \textit{a}- was a prefix in the language from which these words were borrowed, i.e. a pre-Indo-European substrate, and that the prefixed forms received reduced root vocalism in comparison to the unprefixed forms.

The discovery of this feature was a major step forward in Indo-European linguistics, which a century after the discovery of the \textit{Ausnahmslosigkeit der Lautgesetze} had more or less reached its limits. Expectedly, many more examples of \textit{a}-prefixation will be found in the future. One possible candidate worth mentioning here is the Germanic and Balto-Slavic word for ‘sturgeon’, cf. ON \textit{styrja}, OHG \textit{sturio}, OE \textit{styrgia} < PGm. \textit{*sturjan-}, -\textit{jôn-} < \textit{*str-i-}: Ru. \textit{osëtr}, SCr. \textit{jësetra}, Lith. \textit{erškëtras}\textsuperscript{1} < \textit{*asetr / *esetr-}\textsuperscript{2} An etymological connection between these words has of old been suspected (cf. Miklosich 1886: 105; Fick/Falk 1905: 505; Vasmer 1964–1973: 3, 158–9), but the alternation of an initial vowel in Balto-Slavic and \textit{zero} in Germanic has never been given an explanation.

---

\textsuperscript{1} Lith. \textit{erškëtras} as well as OPruss. \textit{esketres} are formally more obscure. This can be due to contamination with \textit{erškëtis} ‘thorn’ (Fraenkel 18; Derksen 2008: 145). The Lith. variant \textit{asëtras} probably is a Belorussian loanword, so that we largely rely on the Slavic evidence for the original form.

\textsuperscript{2} Due to the reshuffling of initial \textit{*a} with \textit{*e} in the Baltic and Slavic dialects, i.e. the phenomenon referred to as “Rozwadowski’s change” (see Andersen 1996: §5.3.3/4; Derksen 2002; Kortlandt 2002–3), the material is largely inconclusive as to the anlaut. We may consequently reconstruct either \textit{*asetr-} or \textit{*esetr-}. 
Although Schrijver’s reduction rule does not seem to apply in this case, there is a high chance that the word originates from a non-Indo-European language. The Substrate Hypothesis is of great interest not only to linguists, but has also drawn the attention of archaeologists and geneticists who focus on the migrational history of Europe. Obviously, the lexical residue or substrate has the potential to shed light on the nature of the interaction between Indo-Europeans and non-Indo-Europeans in that period (cf. Rifkin 2007).

Concretely, the Germanic substrate has previously been associated with Basque, Semitic (Vennemann 1995), and Finno-Ugric (Wiik 2002), but the evidence for these conjectures has so far been non-compelling (Mees 2003: 21). To my mind, the most promising hypothesis regarding the Germanic substrate is the linkage with the introduction of agriculture in North-West Europe. The Neolithic Revolution gradually spread from the Fertile Crescent to Europe through Anatolia and the Balkans between the eighth and the fourth millennium BC, and is strongly associated with the Central European Linear Pottery Culture of ca. 5500–4500 BC. It was argued by Gimbutas (1989) on the basis of archaeological finds that “Old Europe” was culturally and therefore also linguistically homogeneous prior to the arrival of the Indo-Europeans. This is in agreement – or at least not in disagreement – with the available linguistic data, namely that traces of the non-Indo-European features such as a-prefixation can be found from Greek and Latin in the South to Celtic and Germanic in the North. The geographical distribution of this linguistic feature thus patterns with the gradual spread of agriculture.

It was further shown by Haak (2005) that the earliest European farmers were genetically distinguishable from modern Europeans, suggesting that agriculture was transmitted to European hunter-gatherers by immigrant communities. In this way, the linguistic, archaeological, and genetic evidence seems to pattern with the hypothesis that the Germanic substrate is related to the non-Indo-European layer of words in Greek (“Pelagian”), and represents the linguistic residue of the first European farmers (Kallio 2003; Schrijver (2007: 21).

The aim of this article is to adduce new evidence in support of the Agricultural Substrate Hypothesis. I will introduce a linguistic marker by which a small group of substrate words can be isolated from the rest of the Germanic lexicon. The resulting corpus of words forms the basis for further analysis. It turns out that some words belonging to this group contain morphological elements that also appear in the layer of Greek substrate words characterized by a-prefixation. This is an important argument for the claim that the Germanic substrate and “Pelagian” are indeed related, as has been argued by Schrijver (2007).
2. A new substrate marker: root noun inflection

Like in Indo-European, Germanic nouns usually consist of three elements: 1) a root, 2) a suffix, and 3) an ending. The word for ‘day’, for instance, is reconstructed as PGm. *dag-a-z in the nominative, and in this case, *dag- is the root, *-a- the suffix, and *-z the ending. Not all nouns conform to this formula, however. A considerable amount of nouns have a root and an ending, but lack a suffix. Referred to as root nouns, this type of noun is also part of the Indo-European heritage and occurs abundantly in Greek, Latin, and Sanskrit. In Germanic, however, root nouns are perceived as an archaic category. With only a few exceptions⁴, it seems to be either closed or moribund in most attested languages. The evidence nevertheless suggests that at an early stage in the history of Germanic, this class was not closed at all, but indeed open to ancient loanwords. This is substantiated by at least the following cases: 1) *arwīt- ‘pea’, 2) *gait- ‘goat’, 3) *hnit- ‘nit’, 4) *hnut- ‘nut’, 5) *edis- ‘lady’, and probably also 6) *wisund- ‘bison’, as I will try to demonstrate below. While the incorporation of ancient loanwords into the dying class of the root nouns may seem paradoxical at first, it is in fact rather logical when one assumes that the loanwords in the source language did not have a suffix that was recognizable as such to the speakers of Proto-Germanic. The structure of this language is after all likely to have been very different from the Indo-European languages. What I therefore wish to propose here is that root noun inflection can serve as a borrowing (substrate, contact) feature, although it is, in fact, not a feature of the donor language, but of the receiving language itself.

3. PGm. *arwīt- ‘pea’

A demonstrably non-Indo-European word that is inflected as a root noun in Germanic is the word for ‘pea’. In the literature, we find two different reconstructions, viz. *arwīt- and *arwait- (cf. Bugge 1899: 438; Falk/Torp 1909: 19; Pokorny IEW: 335). This disagreement is a result of the situation in Old High German, which has both arawīz and arowīz. This doublet is further mirrored by MHG ar(e)weiz, arwīz, erbiz (G Erbse), but the other West Germanic forms, OS erit, MLG erwete, MDu. a(e)wete, Du. ert are ambiguous and no longer offer compelling evidence for either *arwīt- or *arwait- (except, perhaps, for MDu. arweete, which seems to point the latter variant). Theoretically, the formal variation can easily be ascribed to the fact that the word was borrowed from an unknown language (see below), but it is possible, too, that arowīz arose due to popular etymology, i.e. as a result of contamination with (h)weizi ‘wheat’ < *hwaitja-. The form *arwīt-, at any rate, seems to be the reconstruicio difficilior; it cannot be explained by recent analogies, and must therefore be old.

⁴. In Old Norse, root noun inflection is secondary in most if not all words ending in -ǫng, cf. rǫng, pl. rengr ‘rib in a ship’, spǫng, pl. spongr ‘spangle, flake’, stǫng, pl. stangir ~ stenger ‘pole’, tǫng, pl. tangir ~ tengr ‘smith’s tongs’. Also note the loanword hǫnk, pl. hankar ~ hēnkr ‘coil’ << MLG hank (De Vries 1962: 281).
That the Germanic word for ‘pea’ was a root noun is demonstrated by the Old Norse plural form *ertr with -r < *-iz < *-es. It is in fact not attested in the singular in this language, which is not a matter of coincidence, of course, because peas obviously are encountered in large quantities more often than one at a time. The Modern Icelandic singular form ært, for instance, is a backformation from the plural ærtur (cf. hneta ‘nut’ from hnetur < ON hnøtr, hnetur). The genitive ærttra and the dative ærttrum further show that the ending -r was mistakenly taken together with the root, something that is only understandable if the singular was marginal or absent. In Faroese, the plural ærtur was reanalyzed as a singular form, which gave rise to the secondary plurals ærtar. A vestige of the originally r-less stem is found in the genitive plural form ærta. The original inflection was preserved in Old Swedish as ært, pl. æter, gen. pl. ærtta.

Etymologically, PGm. *arvīt- belongs to the cluster of Gr. ἐρέβινϑος, ὀρόβος ‘(chick)pea’, Lat. ervum and perhaps also Olr. orbaind pl. ‘kinds of grain’. The reconstruction of an Indo-European proto-form *h₁erewgʷ- (Mallory/Adams 2006: 167) or of a doublet *erogʷo- ~ erogʷo- (cf. Sturtevant 1911: 212) is beside the point, because the irregular vocalic and consonantal correspondences prove that we are confronted with a non-Indo-European item (Furnée 1972: 98, 231, 273; Brown 1985: 60; Schrijver 1991: 36; Witczak 1996: 175).

Likewise, Van Windekens’ proposal to reconstruct the word as *ereu- through a lost Indo-European substrate in which PIE *y became *b (1950: 399–400) seems futile. Given the formal irregularities and the fact that the word denotes a crop that reached Europe together with the agricultural revolution, not with the Indo-European expansion, we are clearly dealing with a Wanderwort that ultimately derives from a Near Eastern source. Remains of peas and chickpeas are found in the Fertile Crescent as early as the 8th millennium BC, and genetic evidence suggests that both legumes were domesticated in Southern Turkey and the Levant (Zohary/Hopf 2000: 101–111).

Of all correspondences, Gr. ἐρέβινϑος and PGm. *arvīt- seem to form the closest match. The Greek word contains the notoriously Pre-Greek suffix -νθος as in λαβύρινϑος ‘labyrinth’, ἀσάμινϑος ‘bath tub’, ὑάκινϑος ‘hyacinth’ etc. In Germanic, there are hardly any traces of this suffix, but it can no doubt be identified with the element *īt- in *arvīt-, as was already suspected by Kuiper (1956: 217ff). Pokorny’s statement that we are dealing with a “bloßes Suffix” (IEW: 335) in Germanic is mistaken. Just so is the idea that *arwait- is a compound of a form *arwa- corresponding to ὀρόβος and *aitōn- ‘oat’ (Kluge 1910: s.v. Erbse), because this derivation leaves the variant *arvīt- unexplained. I assume that the long *ī of PGm. *-īt- reflects a Pre-Germanic suffix *-īd- that developed out of *ind- by nasalization of the vowel. The fact that this suffix *-ind- shows
a different Lautstand than its counterpart *-ind- in ἐρέβινθος is not necessarily problematic; it can have many different reasons, such as, for instance, the phonetic peculiarities of a possible intermediate language or simply the irregular way in which sounds are substituted during the process of borrowing. In fact, even within Greek, the interchange of *d- with *-d- and *-in- with *-ī- is well attested. A case that certainly must be mentioned here is the word for ‘garlic head’, i.e. γέλγις, gen. -ῑδος, -ῑϑος and the closely linked ἄγλις, gen. -ῑδος ‘id’. Both variants show an interchange of δ with ϑ. Clear examples of prenasalization are γάλινϑοι, γάλιϑοι, γέλινϑοι, γέρινϑοι ‘peas’, τερέβινϑος, τέρμινϑος, τρέμιϑος ‘turpentine tree’ (Kuiper loc. cit.; Beekes 2010: 258, 1469–70). Thus we arrive at a Pre-Greek suffix *īd(h)- or *ind(h)-.

Turning back to the Germanic situation, it seems safe to assume that the athematic inflection of PGM. *arwīt- indicates that the -ind(h)- suffix did not end in a vowel in the source language. The original athematicity of the suffix is in fact not only demonstrated by the Germanic form, but also by athematic substrate words in Greek such as μέρμις, -ῑϑος ‘string’, which occurs beside the “nasalized” (Beekes 2010: 932) variant μέρμινϑα. An even better example of athematic -nuθ- is offered by ἔλμις, gen. ἔλμινϑος f. ‘helminth, intestinal worm’. The Pre-Greek origin of this word is ascertained by the irregular variants gen. ἔλμιγγος, nom.pl. λίμινϑες (Frisk 1,501), which make attempts at Indo-European etymologies, such as the connection with Skt. kṝ̥mi-, Lith. kirmìs, OIr. cruim < PIE *kʷrm-i- and εἰλέω ‘to turn’, a priori futile (cf. Beekes 2010: 414–5). Also note the already mentioned γέλγις ~ ἄγλις.

Parenthetically, the doublet ἄγλις ~ γέλγις represents a key case itself as it combines the ind(h)- suffix with the equally un-Indo-European a-prefix9 in one single word. It can therefore theoretically be claimed on the basis of this item that the two elements belonged to one and the same language, i.e. Pre-Greek. The ablaut of ἄγλις and γέλγις has been reconstructed as *a-gl- vs reduplicated *gel-gl- (Frisk 1,295) or *ge-gl- (Beekes 2010: 265). In my view, the alternations can better be settled with the help of Schrijver’s rule, which revolves around the idea that prefixed roots have vowel reduction in the root. Assuming that γέλγις developed out of *γέγλις by metathesis, I arrive at a triconsonantal root *g-g-l- (on which now see Kroonen 2012). This root was again combined with the *-īd(h)- suffix, thus *a-ggl-īd(h)- and *gegl-īd(h)-.

7. I think that it is wrong to speak of “prenasalization” (Beekes passim), because this implies that the nasal was added, while it is more probable that it was lost.
8. Can there be a link with Du. elft ‘larva’ < *alb-it-, OHG alba ‘id.’ and Nw. dial. alme ‘id.’? The latter two forms are usually derived from PIE *h₂elbh-o- ‘white’ (Pokorny IEW: 30–31), but the m of Nw. alme must then be explained from the gen.pl. *alma < *alfna.
4. PGm. *gait- ‘goat’

The PGm. root noun *gait-, cf. ON geit (pl. geitr), OE gāt, OHG geiz, has only one extra-Germanic correspondence, i.e. Lat. haedus ‘kid, young goat’. This word is not exactly irregular, but it does seem to contain the conspicuously non-Indo-European phone *a; the hypothetical PIE reconstruction **gʰe̞h₂i̞d- at any rate seems unattractive in view of the bisyllabic root structure, and *gʰh₂e̞id- is not much better.¹⁰ The non-Indo-European origin is further indicated by its limited geographical distribution and the semantics pointing to a culture familiar with goat-keeping. In spite of common opinion, the Proto-Indo-Europeans probably did not have a word for ‘goat’. This follows from the fact that all “Indo-European terms for ‘goat’ for the most part have areal distributions” (Gamkrelidze/Ivanov 1995: 500). Mallory and Adams (2006: 104) reconstruct five goat words for PIE, viz. *dig-, *bʰu̞g-, *h₂e̞i̞g-, *h₂e̞g- and *kapr-¹¹, but all of these reconstructions, except perhaps *bʰu̞g-, turn out to be post-PIE regionalisms with roots that defy the rules of PIE phonology.

The root *dig- is primarily based on OE ticcen ‘kid’ < *tikkīna-¹², the Hesychius gloss δίζα · αἰξ Λάκωνες and Arm. tik ‘animal skin’ (cf. Pokorny IEW: 222). This is a rather poor foundation, because of the marginal attestation in Greek and the non-compelling semantic link with Armenian. A real problem with this etymon is that roots with two voiced stops were not allowed in PIE, which precludes the possibility that *dig- was inherited. Of course, OHG ziga ‘goat’ could be adduced to prove that the original root was *digh-, and that the geminate of ticcen developed out of *gʰ-n- under Kluge’s law.¹³ This only makes things worse, however, because the Germanic forms can then no longer be reconciled with Armenian tik and – for what it is worth – the Hesychius gloss. In the end, the only way in which an etymological relation between *dig- and *digh- can be maintained is to assume that we are dealing with an ancient yet post-PIE loanword.

Gr. κάπρος ‘boar’, Lat. caper ‘billy goat’, ON hafr, OE hæfer ‘id.’ < *kapr-is even less likely to have existed in PIE, because it is formally close to but ultimately incompatible with Ofr. gabor, W gafir ‘id.’ < *gabro-. Even if one accepts a as a phoneme in the Indo-European proto-language (in spite of the fact that it predominantly occurs in words with a local distribution), an Indo-European origin is doubtful at best, because there is no way to unite *k with *g or *p with *b. Again, this un-Indo-European interchange points to a contact word.

Skt. aja-, Lith. ožys ‘goat’ < *h₂e̞g̱-(i)jo- can hardly be an Indo-European word either (cf. Brown 1985: 175), even though it often features as such in the handbooks. The word itself has a perfectly Indo-European structure, but the non-Indo-European origin still follows from its irregular correlation with the root noun Gr. aiz ‘goat’, Arm. aye ‘id.’, Alb. edh ‘kid’; given the considerable

¹¹. Notation mine.
¹². For the suffix *-īna-, cf. Go. gaitīne ‘little goat’.
¹³. E.g. *tigō, *tikkaz < *d(e)įg̱-on, *dįg̱-n-ōs.
formal and semantic resemblances, I find it unsatisfactory, at any rate, to detach the two words from each other. An Indo-European etymology is especially unattractive because Gr. *αἴξ ‘goat’, Arm. *ayc cannot be reconstructed as PIE *ʰ₂eig-. this would have become **hayc with an initial h. The alternative *ʰ₂oig- does not work either in view of the vocalism. As a result, we are forced to reconstruct the word as *ai̯g-, a form that violates PIE phonology in two different ways: 1) it starts with a vowel and 2) it contains **α.

All things considered, there are no indications that *ai̯g(ī)- and *ai̯g- were of Indo-European origin. I would rather suggest that they represent two independent reflexes of a *Wanderwort that sounded like *ai̯dī or *ai̯: the word entered Proto-Greek, Proto-Albanian, and Proto-Armenian as *ai̯g-, whereas in Balto-Slavic and Indo-Iranian it turned up as *ai̯g(ī)-. I think that this formal split resulted from dissimilar treatment of the palatal element, which was adopted as *-i̯g- and *-i̯g- respectively. The independent adoption of the word thus seems to reflect an ancient linguistic and geographic separation between a Balkan group consisting of Greek, Albanian, and Armenian, and the more central satem-dialects underlying Balto-Slavic and Indo-Iranian to the North of the Black Sea.

Another item with a questionable Indo-European background is OCS koza, Alb. kedh ‘goat’. In spite of the formal and semantic similarities, this word cannot be reconciled with OE hēcen, MDu. hoekijn ‘kid’ < *hōk-īna- as is traditionally assumed (thus Pokorny IEW 517–518), because the former presupposes a root *kag̑h- (Winter’s law), the latter *kāg̑i-. But even without the Germanic correspondence, the Slavic word can positively be identified as non-Indo-European, because roots combining a plain and a voiced aspirate stop were forbidden in the proto-language.

It was suggested by Möller (1911: 128) that the source of OCS koza is Semitic *gadi̯- ‘goat’, cf. Arab. gadīn, Hebr. g̑ō̯tī, Berb. agāyd. The same etymon, however, was more appropriately linked to **g̑a̯i̯d- by Illich-Svitych (1964: 3), who assumed borrowing at the Proto-Indo-European level. This view was embraced by Levin (1995: 115–9) and the occidentalist Bernal (2006, III: 101–2), the latter arguing that the languages in which it occurs, i.e. Germanic and Latin, were too distant from South-West Asia to have adopted it individually after the disintegration of the Proto-Indo-European unity. To Vennemann, however, exactly this dialectal isolation meant the confirmation of his view that Northern Europe was colonized from the Western Mediterranean by Semit(id)ic sea-farers (e.g. 1995: 90; 2003a: 250–2, 615). Both hypotheses are actually defendable, but Vennemann’s scenario implies that cattle breeding spread to North West Europe through sea-bound colonization along the Atlantic coast, which is contradicted by the archaeological evidence.

On the whole, it seems better to abandon the idea that Semitic served as the direct source language for **g̑a̯i̯d. Archaeological and genetic evidence suggests that the goat was first domesticated in the Zagros Mountains as early as ten thousand years ago (cf. Naderi 2008) and the ultimate origin of the word may therefore actually lie well beyond the Proto-Semitic horizon. I am therefore inclined to follow Diakonov’s suggestion (1985: 132) that both Italo-Germanic
Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis

*ghaid- and PSem. *gadj- originate from a third source, which ultimately may be connected to the aforementioned synonyms *a'id- and *kāg(ʰ)-. In this context, it is important to consider similar words for ‘goat’ and ‘billy goat’ in the Caucasian languages, e.g. Adyge āča ‘he-goat’, Dargwa (Akushi) ḫeža ~ (Chirag) ḫač:ə (Witzel 2004: §2.3). Most plausibly, the ‘goat’ word filtered through the old continuum of agricultural and cattle breeding cultures that had expanded into Europe from the East in the millennia preceding the arrival of the Indo-Europeans. It must, at any rate, have been adopted by the Germanic and Italic sub-groups after their settlement into or around these communities.

5. PGm. *hnit- ‘nit’

Another Germanic root noun with an unclear Indo-European etymology is the word for ‘nit’, cf. ON gnit, pl. gnitr, OE hnitu, OHG niz. The word occurs only in the European dialects, and the proto-forms evinced by these languages are seriously at odds with each other. The Germanic form points to PIE *knid-, a root that possibly also underlies Ru. gnīda, Latv. gnīda < *knid-eh₂, although the development of *kn- to *gn- is slightly unexpected. Similar, but not identical, is the base of Gr. κονίς, -ίδος and Alb. thëñi < *konid-. Theoretically, it is possible to assume an ablauting root *kon- ~ *kn- here, but the suffix *-id has no real Indo-European foundation. A non-Indo-European origin is further implied by the more irregular cognates, Arm. anic ‘louse’ < *h₂nid-, in which a laryngeal suddenly pops up, and MIr. sned ‘nit’ < *snid-eh₂ with an initial s. Even more strange are Lith. glīnda and Lat. lēns, lendis < *gle/ind-eh₂, although these forms may have developed from *gne/indeh₂ by dissimilation of the first n (Derksen 2008: 169; De Vaan 2008: 334).

All data taken together, it is striking that the word for ‘nit’ in almost every single language displays at least one significant irregularity. These irregularities have often been attributed to folk etymology or taboo mechanisms (cf. Derksen l.c.; Beekes 2010: 747), but they can equally well be interpreted as indications that the word was adopted from a non-Indo-European substrate. Key forms, in this respect, are Lith. glīnda and Lat. lēns. Apart from their irregular onset, these variants are especially noticeable because they contain a suffix *-ind-, whereas all the other languages have *-id-. Since this alternation can positively be linked to the agricultural substrate in Greek (see the discussion on *arwūt-), it is possible that the word for ‘nit’, too, belonged to this layer. Its source would then have to be something resembling *c~ỹ(o)n-ître.14

14. It is not extremely likely that the original root contained an l, because its occurrence seems to be phonetically conditioned: whenever the nasalization of the suffix was retained, as in Latin and Lithuanian, the first nasal was liable to dissimilation.
6. PGm. *hnut- ‘nut’

A non-Indo-European origin must also be assumed for the Germanic root noun *hnut- ‘nut’, e.g. ON hnut, pl. hnotr, hnetr, OE hnutu, pl. hnyte, OHG nuzz, pl. nuzzi. The word only has correspondences in West Indo-European, viz. Lat. nux, -cis f. ‘nut (tree)’ < *knuk-, OIr. cnú, gen. cnó, W cneuen, pl. cnau ‘nut’ < *knu̯H-s, gen. *knu̯H-os. This distribution lends little credibility to the reconstruction of an Indo-European root *knu- (thus Pokorny IEW: 558–9). Celtic and Germanic isoglosses are a priori suspect, because there was no independent Germano-Celtic proto-language; similarities between the two branches can only indicate one of the following three relations: 1) Indo-European archaisms that were coincidentally preserved in Germanic and Celtic only; 2) borrowing from Celtic into Germanic or vice versa, and 3) shared contact with a third language. In the case of ‘nut’, it is implausible that it is an Indo-European archaism, because Pre-Gm. *knud-, Proto-Italic *knuk-, and Pre-Celt. *knu̯H- cannot be unified into a single proto-form. It is theoretically possible to assume secondary suffixation of a root *knu- in Germanic and Italic, but an athematic *d suffix cannot be demonstrated for Germanic, at least.15 It therefore seems more attractive to reconstruct a substrate word. Theoretically, the final Pre-Germanic *d, the Latin *k, and possibly the *H of Pre-Celtic *knu̯H- can be reconciled by reconstructing a glottal stop for the source word, i.e. *knuʔ (Kroonen 2009: 221–2). This glottal stop may have been substituted by the different languages in different ways, resulting in a *d in Pre-Gm., a *k in Pre-Italic, and perhaps a laryngeal in Pre-Celtic. The alternative is to compare the Gm. *t to the one of *arwūt- and to assume that it reflects the substrate element *-ūd(h)-, a variant of *-ūd(h)- (cf. Gr. κολοκύνϑη ‘round gourd’, ὀδόλυνϑοι ‘chick-peas’, ὀλυνϑὸς, ὀλυνϑός ‘wild fig’ (Beekes 2010: 738, 1046, 1074), but this does not explain the Italic form *knuk-. Whatever the case may be, the fact that the word emerges as a root noun in Germanic (as well as Latin) is likely to be the result of the vowelless coda of the word in the source language.

7. PGm. *edis- ~ *disi- ‘lady’

Another root noun with a potentially non-Indo-European origin is *edis- or *idis- as continued by OE idis, OHG itis, OS idis ‘woman, lady (of high standing), matron’. The inflection as a root noun is supported by the ending-less dative form itis in Old High German (Braune §240) and idis in the Old Saxon Heliand, which derive from PGm. *edis or *idis.16

---

15. In Old English, there was a productive *t suffix, cf. OE frēot ‘freedom’, sweofot ‘sleep’, béowet ‘service’, but this suffix was thematic, started in a vocalic element *e, *a, or *u, and was used only to create abstract nouns from verbs and adjectives.

16. More circumstantial evidence in favor of a root noun has been adduced in the form of the Middle Dutch feminizing suffix -nede as in e.g. swasenede ‘sister-in-law’, graefnede ‘countess’, enede ‘wife’, which Hendrik Kern in a letter to Matthias de Vries derived from an s-less form of *idis- added to n-stems such as *grēff(j)an- ‘boss, count’. Although Kern later retracted his idea, he could not prevent Franck from adopting
Several attempts have been made to give this word an etymology, none of them being generally accepted. In spite of its very limited distribution, the word has been projected back into PIE as a hysterodynamic *s-stem *h₁edʰ-ēs, pl. *h₁edʰ-ēs-es. This was done by Nedoma and Eichner (2000: 33), who tentatively supposed an etymological correlation with OHG etar ‘pale in a fence’. Even more recently, Bammesberger (2007) proposed a link with Skt. édhas- ‘firewood’ < *h₁eídʰ-os- and other formations to the root *h₁eídʰ- ‘to burn’, assuming that the hysterodynamic *s-stem, which he reconstructed as *h₂idʰ-ēs, pl. *-ēs-es, was originally applied in metaphorical sense, like Modern German Flamme. These semantic interpretations do not seem self-evident, however.

I would like to stress that there is no compelling reason to think that the word is Indo-European. On the contrary, the non-Indo-European character of the word seems to be betrayed by the internally Germanic correspondence with ON dís f. ‘woman, goddess’ < *dīsi-. Given the close formal and semantic match it would be highly unsatisfactory to deny the etymological link between the two words. But this is precisely what sometimes has been done, because from the Indo-European perspective, the correlation between *idis- and *dīsi- is not fully regular.17 This incongruity is exactly the point; however, as it may very well indicate that the word was not inherited from the parent language, but rather adopted from a local pre-historic language in Northern Europe.

Not only is the etymological separation of *edis- and *dīsi- artificial on the linguistic side, there are furthermore strong cultural arguments for connecting the two words. In the First Merseburg Charm, one of the few glimpses into West Germanic paganism, the idisi act as valkyries that actively interfere with the fate of the combatants during battle. This is in conformity with Old Norse mythology, in which the dísir form a general category of matrons encompassing the valkyries, norns, and fylgjur (cf. Simek 2004: 84–5). Further proof in favor of these two words being the same was furnished by Jacob Grimm, who in his Deutsche Mythologie (II: 373) pointed at the parallelism of the Eddaic phrase dís Skjǫldunga ‘lady of the Shielding clan’ with idis Scildinga ‘id.’ in Beowulf.18

In view of this telling linguistic and cultural-historic evidence, it seems rather audacious to me to deny an etymological link between *edis- and *dīsi-. I am unaware, however, of any Proto-Germanic or Proto-Indo-European morphological process according to which *edis- and *dīsi- could be unified into a single form or paradigm. One could theoretically resort to something as intricate

---

17. Bammesberger derived *dīsi- from the PIE root *Pēi- ‘to shine’, for instance.
18. It was recently inferred by Nedoma that the syntagm of dis Skjǫldunga was a semantically empty title and therefore cannot be old. I think that this reasoning must be reversed: exactly the fact that dis Skjǫldunga evolved into an epithet proves its old age. Also see the discussion between Frederic Armory and Pat Belanoff (1990).
as an ablauting *is-stem\(^{19}\), e.g. PIE *h₁édʰ-iōs, gen. *h₁dʰ-is-ós, acc. *h₁dʰ-iés-m > PGM. *edjōz, *dizaz, *djesun. But even a paradigm like this could probably not account for the long *ī of *dīsi-, although, admittedly, I am not entirely sure what the exact outcome of PIE *h₁dh-ies-m would be. It therefore seems worthwhile to consider the possibility that both Germanic formations were adopted from a local language that had two different forms, viz. *edis and *dīs. But even a paradigm like this could probably not account for the long *ī of *dīsi-, although, admittedly, I am not entirely sure what the exact outcome of PIE *h₁dh-ies-m would be. It therefore seems worthwhile to consider the possibility that both Germanic formations were adopted from a local language that had two different forms, viz. *edis and *dīs. It is, of course, tempting to compare the alternation in this hypothetical language with the process of a-prefixation that is found in other non-Indo-European words. One could speculate, for instance, that the *e of *edis is a regional, more fronted vowel variant of *a, e.g. *æ or *e. The alternation between short and long i in *edis and *dīs could be attributed to the stress: it was claimed by Schrijver that the prefix attracted the accent, resulting in a reduced vocalism in the second syllable, cf. *arud : *raud ‘ore’. With this principle, the long i of *dīs can be understood as reflecting stressed i in the donor language.

Possibly, a fronted pronunciation of the phone *a in the pre-Germanic language is corroborated by PGM. *þeura- ‘bull’ as represented by ON ḣjórr, Du. dial. duur. This word has been reconstructed as PIE *tauro- on the basis of e.g. Lat. taurus, Gr. tauros, Lith. tauras, OCS τυρ, Alb. ter\(^{20}\), but in view of the irregular formal relationship with OIr. tarb < *taruo-\(^{21}\), Go. stiur, ON st jórr, OE stēor, OHG stior m. ‘bull’ < *steuro- and *þeura- itself, there is little point in projecting this word back into the Indo-European proto-language. This would, in fact, only result in an unfortunate increase of the corpus with problematic a-vocalism and movable s. It is far more plausible that the word ultimately shares its origin with Proto-Semitic *ṭawr, cf. Akk. šūru, Arab. twr and Hebr. taʿur ‘steer’ (Schmidt 1890: 7; Möller 1907: 214; De Vries 1962: 614; Levin 1995: 403–5; Vennemann 1995: 88–89; Mallory/Adams 1997: 135). The dialectal divergence in Europe further proves that the word was not borrowed at the Proto-Indo-European level, but only after the fragmentation into the different daughter languages. The exact source language from which Pre-Germanic borrowed *þeura-, whether a Semitic or a third intermediary language, appears to have been characterized by the brightening of *a to *e at least under certain circumstances.

In earlier scholarship, it has been claimed that the *e of *þeura- was introduced secondarily under the influence of *steura- (Pokorny IEW: 1080–5), which is formally compatible with Av. staora- ‘Großvieh’. This solution does not help much, however, because neither *þeura- nor *steura- can in my view be separated from PSem. *tawr (and Etr. thevrúmines ‘Minotaur’). The alternation between *þ- and *st- is usually accounted for by invoking a movable s. The point is that the exact origin of this movable s is considered to be obscure. I find it likely, in this particular case, that the different onsets of *þeura- and

---

\(^{19}\) The PIE comparatives were inflected as hysterodynamic is-stems, cf. Gr. ṭhiovo, acc. ṭhio < *sueh jd-iōs, *ios-m (with analogical transfer to the n-stems).

\(^{20}\) The e-vocalism is due to Umlaut that took place in the Proto-Albanian paradigm *tar, pl. *tar̭ (Demiraj 1997: 46). It does not imply *eu, as is claimed by Mallory/Adams (2006: 136).

\(^{21}\) Fi. tarvas is not comparable to the Celtic form, because it metathesized *-vr- to *-rv-, cf. karva ‘hair’ from Lith. gaũras ‘id.’
Non-Indo-European root nouns in Germanic:
evidence in support of the Agricultural Substrate Hypothesis

*steur-* ultimately represent two different outcomes of the sound that corresponds to *τ* in Proto-Semitic. The differentiation may have taken place in various ways. It is conceivable, for instance, that a form *peur* was borrowed twice by Germanic, first as *steur-* before Grimm’s Law when the language did not yet have dental fricatives, and later as *peur-* when it did. Another likely scenario is that the different onsets arose by repeated sound substitution when the Wanderwort passed from one language into another (Schmidt 1890: 7; Uhlenbeck 1896: 136). Finally, the “movable s” can be ascribed to a dialectally divergent evolution of the sound corresponding to *τ* in the source language; the proto-forms with *τ*- and *st-* would then ultimately have been borrowed from different dialects. This is probably supported by the parallelism of PGm. *steura-* with Av. *staora-*, which could have been adopted from two opposite sides of the alleged linguistic continuum.

At any rate, the situation is reminiscent of the PIE word for ‘star’, cf. Hitt. FileStream, Gr. *ἀστήρ*, Go. *stairno*, Lat. *stella* etc. < *h₂ster-*. In the past, attention has been drawn to the formally and semantically similar Proto-Semitic form *aṭtar- ‘(star) goddess, Venus’ as evidenced by Akk. *ištar*, Hebr. *ʾaššōret*, SArab. *ṭtr*. Bomhard (1986) rejected this equation, arguing that the Indo-European word must be derived from the root *h₂e₁h₁-s- ‘to burn’ (cf. Mallory/Adams 2006: 129; Pinault 2007). But it was argued by Hamp (2000), that a form *h₂H₁s-ster-* probably would have given Lat. **astella, with vocalization of the second laryngeal, an objection that is generally bypassed by assuming simplification of *h₂h₁ster-* to *h₂ster-*, however. The most important argument against an Indo-European origin then consists of the irregularities displayed by Sanskrit, viz. nom. pl. *tāraḥ* < *h₂tēr-es*, inst. pl. *sṭfbhiḥ* < *h₂str-bh₁is*. Instead of circumventing these irregularities by assuming a movable s, it is perhaps more efficient to assume that the word itself passed into the Indo-European dialect continuum from an external source, and that the formal problems result from the difficulties in pronouncing the PSem. cluster *ṭṭt-*. Gamkrelidze and Ivanov (1995: 592–2) argued that, in the process of borrowing, the in PSem. *ayin* could have been substituted by *h₂a* and the *t* by either *s* or *zero* in (disintegrating) Proto-Indo-European (cf. Vennemann 2003a: 357–8; 2003b: 888). The substitution of *ṭṭ* by *(s)t* would then be comparable to the development of *τ* in the words for ‘bull’.

---

22. I further have doubts about the adduced semantic parallel of Alb. (h)yll ‘star’, which Mallory and Adams (2006: 129) take to be derived from the root *h₂μs-* ‘to burn’. Possible, too, is the connection with PIE *h₂eus-* as in *h₂eus-ōs*, *h₂μs-s-ōs* ‘dawn’. The proto-form would then be *h₂μs-lo-.*

23. Technically, the *s* would be an infix in *h₂ster-*, because it comes after the laryngeal.

24. It was claimed by Diakonov (1985: 123) that the meaning ‘Venus’ developed secondarily in Semitic, and that the word could therefore not have been borrowed into (post-)PIE, where ‘star’ is the only connotation. But this argument can just as well be reversed: the fact that the meaning ‘star goddess’ developed in Semitic must mean that it was a native word in this family.

25. In a footnote (1995: 772), Gamkrelidze and Ivanov contradict themselves when they accept Bomhard’s internal Proto-Indo-European derivation from the root *h₂H₁s- ‘to burn’.
8. PGm. *wisund- ‘bison’

Another case that must be considered here is OHG *wisunt, -ant, ON visundr, OE we(o)send m. ‘bison’. The word is typically inflected as an a-stem in Old English, but there are indications in Old High German that it originally was a consonant stem. First, there is the OHG plural form *wisunti, which with its ending -i points to an i-stem. Since practically all masculine consonant stems shifted to the i-stems in Old High German (Braune 1891: 180), this is, too, likely to have happened in the case of *wisunt.26 A second indication that ‘bison’ was originally inflected as a consonant stem consists of the Verner variation that is implied by the proper name Wirunt. Of old, it has been claimed that this name is etymologically identical to the appellative, which is not at all implausible in view of the name of the Gothic flag-bearer Ovísaundo mentioned by Procopius. The underlying form *wizund- implies an accentual alternation within the paradigm. It follows that the athematic reconstruction as given by Fick, Falk, and Torp (1909: 413) and Schaffner (2001: 631–634) is better than the a-stem that was reconstructed by e.g. De Vries (1971: 61).

PGm. *wisund- is generally taken to be of Indo-European origin. It has been connected to e.g. Skt. viśāna- f. ‘horn’ (Pettersson 1921: 39), to Lat. virum n. ‘stench, slime’, the original meaning allegedly having been “horned one” or “smelly one” respectively (e.g. Mallory/Adams 1997: 136) and to PIE *uēs- ‘to consume’ (Stiles 2004). The reconstruction that one finds in the literature is an nt-stem, i.e. *uis-ont-. This stem was assumed by Schaffner (2001: 633) to be the continuation of a fully ablauting paradigm *uēis-ont, *uis-nt-ōs by the generalization of the zero grade of the root. Those who prefer an Indo-European origin may find it interesting, however, that the root vowel of ON visundr in fact was long. In spite of the dictionaries (e.g. Hægstad 1930; De Vries 1962; Zoëga 1910), which cite the form with ĕ, Molcel. visundur clearly shows that the Old Norse vocalism was *ī. This means that, if one ignores the probability that visundr was borrowed from Low German (thus De Vries 1962: 669), the Germanic forms can be accounted for by deriving them from an ablauting paradigm *wīsand, *wizundaz < *uēis-ont, *uis-nt-ōs.

Yet however elegant the reconstruction *uēis-ont, *uis-nt-ōs may seem from the Germanic perspective, it loses its validity as soon as the Balto-Slavic evidence is taken into consideration, viz. OPruss. wissambs27, Lith. stumbras, Latv. sūbrs, sūbrs, OCS, zōbr. The formal differences of these forms with the Germanic word have been reason to deny any etymological link between Balto-Slavic and Germanic, but the similarities, especially those with OPruss. wissambs’, are too considerable to accept such a separation. The divergence

26. Such a scenario is further supported by the dative plural forms wisuntun, wisintun. With the ending -un, they fit into the paradigm of either the a-stems or the consonant stems, but not the i-stems, where the ending is -im. The inflectional type of *wisunt is, in other words, comparable to the one of the original root noun fuoz ‘foot’, which in the nom. pl. appears as fuozzi and in the dative as both fuozim and fuuzzum, fuozun (cf. Braune 1891: §229).

27. It is found in this form in the Elbing Vocabulary. I will refrain from the question of whether or not this form should be emended to *wissambris, as was done by Trautmann (1910).
between the two branches nevertheless remains problematic, because there is no single proto-form that can explain the entire set of forms. In fact, when one takes a closer look at the other Balto-Slavic correspondences, it becomes clear that the word is not only lacking a solid Indo-European form, but that even within this branch, several irreconcilable proto-forms must be reconstructed in order to account for the material: OCS, *zōbrъ goes back to PBSl. *źambras/źambris. It can be unified with the second element of OPruss. *wissambs’, which seems to go back to *vízamb(r)as28, but not with the Baltic correspondences, where the expected outcome would have been Lith. **žambras, Latv. **zuobs. The attested forms, Lith. stūmbras, Latv. sumbrs rather presuppose PBSl. *s(t)umbras or rather to *s(t)ambras with the Eastern Aukstatian development -aN- > *-uN-.

Obviously, the four different proto-forms *źambras, *vízamb(r)as and *s(t)ambras cannot be reconciled with each other, let alone with PGm. *wisund-. In spite of this, some have projected the form *źambras back into PIE as *ǵomh-ro- (cf. Mallory/Adams 1997: 136), and derived it from the PIE root *gombh- ‘jag, tooth’ (cf. recently Schütz 1997: 97). OPruss. *wissambs’ has even been explained as a contamination of this *źambras with the Germanic word (Petersson 1921: 40). Both approaches seem futile, however, because they still do not offer an explanation for the East Baltic words. It is of course possible to assume that the initial st- of Lith. stūmbras is analogical after stémbti ‘to be stubborn, buck’ (Buga 1912: 45), or that the dissimilarities are unproblematic because “such phonetic transformations of cultically significant words are frequent among the names of animals” (Gamkrelidze/Ivanov 1995: 440). One might even assume that the correspondence Lith. st ~ Sl. z is indicative of the Indo-European reflex of the alleged Nostratic phoneme *č (Ivanov 1975). But, to my mind, it is more economical in this particular case, to abandon the idea that the word was inherited from Proto-Indo-European, especially in view of Caucasian correspondences such as Ossetic, Adygh, Georgian dombaj, Abkhaz a-dawp-ēy ‘aurochs’ (cf. Ivanov 1975; Gamkrelidze/Ivanov 1995: 440; Aabev 1996: 206).29 It rather seems that the word penetrated into Germanic and Balto-Slavic separately, when they had already settled in Western and Central Europe. Given the homogeneity of the Germanic material as opposed to the dialectal variation in Balto-Slavic, this must have happened before the disintegration of Proto-Germanic, but after the splitting-up of Proto-Balto-Slavic, i.e. somewhere in the middle of the first millennium BC.

The form of the non-Indo-European word can be approximated by approaching it top-down. Comparison of the different forms shows that the original coda can safely be reconstructed as *-ombr for Balto-Slavic. This element was borrowed into Slavic as *ombr > *-oBR, because PBSl. *a had already changed into PSl. *o at that time, but as *umbr in Baltic, because the only vowels occurring before -mb- were *a and *u in this language. The same argument goes for Germanic: since this language did not have a short o, only *u and *a, the original

---

28. For the *ss from *z cf. assaran ‘lake’ = Lith. ėzeras.
29. Derksen (2008: 549): “A connection with the root *ǵomh- of PSl. *zōbs, Latv. zuobs ‘tooth’ cannot be ruled out, but it is possible that we are dealing with a migratory term, cf. Osset. dombaj ‘bison’.”
vocalism can have been *o, too. The Germanic coda, however, clearly points to *-nd.

For the onset there are a number of possibilities, but the most plausible options are *ts- or *ʧ- for Baltic and Germanic, and *dz- or *ʤ- for Slavic. In spite of Lith. stumbras, *st- is out of the question, because this cluster would have remained in both Germanic and Slavic. It seems more apt to start from a cluster in which the dental element preceded the sibilant, viz. *ts-. This is in accordance with the evidence that PBSl. *ʒ < *gʰ ʷeːr- in early Slavic, as it still is before ʰ in e.g. Mac. dzver, Ukr. (d)zvir ‘animal’ (< *gʰ ʷeːr- (Kortlandt 1980: 250; 2008: 5). I therefore assume that PSl. *zbrъ/β entered Slavic as *dzombr or *ʤombr. The Baltic forms can be explained by starting from the voiceless variant of the same form, viz. *tsombr or *ʧombr. Since this *ts emerges in metathesized form in Lith. stumbras and not as š or č, it seems plausible to assume that it entered the language at a time when there was no *ts or *ʧ with which to substitute it. This was the case in the considerable period between the simplification of *tš (PBSL *ʒ < PIE *k) to š and the late affrication of *tʃ to č (cf. ėntis ‘face’, gen. ėnçio).

The derivation of OCS zbrъ and Lith. stumbras from Pre-Balto-Slavic *dzombr or *tsombr leaves us with the first element of OPruss. wissambs’ and PGm. *wisund-z. I have already stated that I do not like the idea that wissambs’ is due to contamination of *źambras with PGm. *wisund-z. This is unlikely for several different reasons, the most important being that the element *wis- has a strange reminiscence in Ru. iz’ubr’ ‘red deer, Cervus elaphus xanthopygus’. This form ostensibly continues a PBSL variant *ižambris if not simply *vižambris through some kind of dialectal development. Either way, it represents the missing link between OCS zbrъ and OPruss. wissambs’, which means that the element *wi- must be old. I therefore reconstruct the additional variants *widzombr(r) for Old Prussian and *witsond for Germanic.

The variation of *tsombr, *dzombr, on the one hand, and *widzombr, *witsond, on the other, does not hark back to any known Indo-European process, and cannot be explained away by secondary developments in Baltic, Slavic, and Germanic either. It is therefore likely to originate from an external source. I surmise that it is a reflection of a morphological or derivational process in an extinct North-European language. We could speculate, for instance, that this language had a prefix that could be added to the element *dzombr or *tsombr under certain circumstances. It seems an impossible task to establish the original function of this hypothetical prefix *wi- in the now disappeared source language, because it can have been so many things such as a definite article or some kind of classifier.

The *wi- prefix has a potential parallel, however, in the notoriously non-Indo-European word for ‘wild boar’, which emerges in three similar but irreconcilable shapes, viz. cf. Lat. aper < *apr, OE eofor, OHG ebur < *epr and OCS

30. Alternatively, Caucasian dombaj could point to *domb-.
31. The assumption that the t is secondary due to contamination with taũras ‘bison’ or steũbti ‘to bump’ (cf. Fraenkel 932) is ad hoc.
veprь, Latv. vepris < *wepr.\(^32\) The initial \(w\) in Balto-Slavic has been explained as a hiatus-filler (Pedersen 1905: 311–2) or as analogical after Lith. \(ve\̄\text{rīs} \ ‘calf’\) (Kent 1926: 185). More recently, Gamkrelidze and Ivanov (1995: 434–5) have attempted to merge the word with the *kapros-cluster by reconstructing it as PIE *\(q̄\)wep-, assuming that the combination of the (idiomatic) post-velar *\(q̄\)h with a labial glide explains the anomalous correspondences of root vocalism in various dialects. Alternatively, the interchange of initial *\(w\) with zero can tentatively be compared with the one of *dzombr and *widzombr.\(^33\)

A final issue consists of the different Auslauts in Balto-Slavic and Germanic, the former consisting of the labial cluster *-*mb, the latter of the purely dental sequence *-*nd. This difference has been an important reason to separate the two words from each other. But this objection is really beside the point if we are dealing with a non-Indo-European word, of course. Theoretically, the two variants *(wi-)*tomb(r) and *wi-tsond can be reconciled by reconstructing a labio-dental cluster, so as to arrive at *(wi-)*tsonmd. There possibilities are legion, however, and I have chosen to refrain from any speculative reconstructions here. Suffice it to say that there are similar alternations in Greek words of non-Indo-European origin, cf. κορίαμβλον ~ κορίανδρον ‘coriander’, σάμβαλον ~ σάνδαλον ‘sandal’ and perhaps also σίδη ~ σίβδη ~ ξίμβαι ‘pomegranate’ (cf. Beekes 2010: 754, 1305, 1329).

9. Summary and Outlook

The words treated above reveal that, in Germanic, the Indo-European class of the root nouns was open to loanwords or substrate words. This can hardly be anything else but a reflection of the form of these words in the donor language; apparently, these words did not have a vocalic coda when they were adopted. The resulting collection of non-Indo-European root nouns is informative of the nature of the Germanic substrate. The fact that “Neolithic” words, such as *gait- ‘goat’ and *arwīt- ‘pea’ are overrepresented contradicts the idea that the Indo-Europeans were a deeply agricultural people (contra Renfrew 1987, 2001; Lehmann 2002).

In fact, at least a part of the data rather supports the Agricultural Substrate Hypothesis, which revolves around the idea that the Indo-Europeans, after their emigration from the homeland, settled among Neolithic cultures in Europe, and adopted agricultural terms from a continuum of possibly related non-Indo-European languages. An important linguistic argument for the linguistic unity of the European agricultural substrate is provided by the same *arwīt- ‘pea’, which shares the element *-īt- with the Pre-Greek substrate suffix *-īd-. The element further co-occurs with the equally non-Indo-European *a*-prefix in

\(^32\) Also note that again the Northern languages have *e whereas Latin has *a. It is fully parallel, in other words, to PGM. *þeura- vs Lat. taurus ‘bull’.

\(^33\) And what to think of Gaul. uisumarus ‘clover’ vs. OIr. seamar f. ‘id.’ < *semmar- and Icel. smæra f. ‘id.’ < *smēr-/*smair-?

Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis
the Greek doublet ἄγλις ~ γέλγις ‘garlic’ < *a-ggl-id-, *gegl-id-. This may be an additional indication that the Germanic and Greek substrate were indeed related, as has been claimed by Kallio (2003) and Schrijver (2007).

The linguistic evidence is in corroboration with some interpretations of the genetic prehistory of Europe in relation to the arrival of agriculture. Recent studies have shown that the Neolithization of Europe was not a purely cultural process, but at least partly resulted from “demic diffusion”. DNA samples obtained from Linear Pottery Culture sites confirm that Mesolithic hunter-gatherers and the earliest Neolithic farmers had different origins, and that the latter group “shares an affinity with modern-day Near East and Anatolia populations” (Haas et al. 2010). The spread of agriculture into Europe has further been linked to the microsatellite variance of the male DNA marker haplogroup R1b1b2 (Balaresque et al. 2010). The haplogroup is very common along the Atlantic coast, and has therefore been associated with the European post-glacial hunter-gatherer population. However, since the microsatellite variance within this haplogroup is maximal in West Anatolia, this has been interpreted by Balaresque et al. as the result of a severe founder effect. The distribution of haplogroup R1b1b2 has thus become geographically and linguistically compatible with the Agricultural Substrate Hypothesis that is evident for Greek as well as Germanic.

Bibliography


Beekes, R. S. P. 2010: Etymological Dictionary of Greek. With the assistance of Lucien van Beek. Leiden.


Brown, R. A. 1985: Evidence for Pre-Greek speech on Crete from Greek alphabetic sources. Amsterdam.
Non-Indo-European root nouns in Germanic: evidence in support of the Agricultural Substrate Hypothesis


Derkson, R. H. 2002: “Rozwadowski’s change” in Baltic. – Baltu Filologija 11/1: 5–12.


Petersson, H. 1921: Studien über die indogermanische Heteroklisie. Lund.


Witzel, M. 2004: Central Asian roots and acculturation in South Asia: linguistic and archaeological evidence from Western Central Asia, the Hindukush and Northwestern South Asia for early Indo-Aryan language and religion. – T. Osada (ed.), *Linguistics, Archaeology and the Human Past*. Kyoto. 87–211.