

## Milk and the Indo-Europeans

Romain Garnier, Laurent Sagart, Benoît Sagot

► **To cite this version:**

Romain Garnier, Laurent Sagart, Benoît Sagot. Milk and the Indo-Europeans. Martine Robeets; Alexander Savalyev Language Dispersal Beyond Farming, John Benjamins Publishing Company, pp.291-311, 2017, 978 90 272 1255 9. <10.1075/z.215.13gar>. <hal-01667476>

**HAL Id: hal-01667476**

**<https://hal.inria.fr/hal-01667476>**

Submitted on 31 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Milk and the Indo-Europeans

Romain Garnier, Laurent Sagart and Benoît Sagot

Université de Limoges and Institut Universitaire de France / Centre National de la Recherche Scientifique / Institut National de Recherche en Informatique et en Automatique

Recent evidence from archaeology and ancient DNA converge to indicate that the Yamnaya culture, often regarded as the bearer of the Proto-Indo-European language, underwent a strong population expansion in the late 4th and early 3rd millennia BCE. It suggests that the underlying reason for that expansion might be the then unique capacity to digest animal milk in adulthood. We examine the early Indo-European milk-related vocabulary to confirm the special role of animal milk in Indo-European expansions. We show that Proto-Indo-European did not have a specialized root for ‘to milk’ and argue that the IE root *\*h<sub>2</sub>melǵ-* ‘to milk’ is secondary and post-Anatolian. We take this innovation as an indication of the novelty of animal milking in early Indo-European society. Together with a detailed study of language-specific innovations in this semantic field, we conclude that the ability to digest milk played an important role in boosting Proto-Indo-European demography.

**Keywords:** Indo-European, etymology, DNA, archaeology, Yamnaya culture

### Introduction

The Indo-European hypothesis is well over two hundred years old. A strong consensus exists among linguists on the existence of an Indo-European proto-language. There is no disagreement on which languages are Indo-European and which are not. There is also a broad consensus that the first split in the family separated the Anatolian branch, whose main representative is Hittite, from the rest, which we refer to as “Core Indo-European” in this paper. There exists a healthy range of opinions on issues of reconstruction. Beyond these, areas under discussion relate to the location of the homeland, the time depth of the ancestral language and the subsistence of the original community.

Regarding these issues, mainly two theories are in presence: the Pontic Steppe theory, supported by a majority of scholars, and the Anatolian theory. The first identifies the ancestral group with Yamnaya culture, in the steppes north of the Black Sea around 4000 BCE or slightly later, and argues that Proto-Indo-European speakers were hunter-gatherers or pastoralists. Archaeologists working under the Pontic Steppes hypothesis (Mallory 1989; Anthony 2007) have presented detailed accounts of the interface between archaeology and linguistics. An unresolved issue is that the Pontic Steppes hypothesis has not so far provided a principled answer to the question of why the Indo-European languages have replaced the languages of the farmers over much of Europe and South Asia, despite the presumably more favorable demography of farmers. The Anatolian theory arose as an attempt to answer this question. Renfrew (1987) proposed that the first Indo-Europeans were western Eurasia's first farmers, who domesticated barley and wheats in Anatolia 10,000 years ago; and that the success of their languages was the direct result of the success of agriculture. According to him, the languages of the farmers are now spoken over large tracts of western and southern Eurasia because the demography of farmers is generally more favorable than that of hunter-gatherers or pastoralists. Accordingly, the Anatolian theory currently places Proto-Indo-European speakers in Anatolia ca. 6500 BCE. The Anatolian theory of Indo-European origins is one of the models for the more general Farming/Language Dispersal Theory.

More than the mismatch between the alleged early date of Proto-Indo-European and the reconstructability of wheeled transport vocabulary – which absolutely cannot be as old as agriculture – it is the absence at the highest node in the Indo-European tree of a clear, diversified Proto-Indo-European agricultural vocabulary (Uhlenbeck 1895, 1897; Kortlandt 2009), which reveals the basic problem of the Anatolian theory. Indo-European cereal-related vocabulary exists, but is either regional, semantically too vague to permit the inference of farming, or unrelated to agriculture. Thus Lat. *hordĕum*, -ī n. 'barley', often seen as a direct cognate of Germ. \**gerstō*- f. 'barley', is a late formation from *horridus* 'shaggy, bristly' > \**horrid-ium* n. 'ear of barley', regularly syncopated in \**hórdium* > Vulg. Lat. *hórdĕum*. The source is PIE \*ǵ<sup>h</sup>*ers-* 'to be bristly'; the Germanic word is perhaps independently derived from the same source: ears of barley are indeed strikingly bristly. The cognate set Lat. *Cĕrēs* f. 'goddess of vegetal growth', Hitt. *karaš* n. 'cereal plant', MoGerm. *Hirse* m. 'millet' (< Com. Germ. \**hersija(n)-*), does not allow the name of a specific cereal to be reconstructed: rather, it goes back to PIE \**kerh<sub>3</sub>-* 'satiated' (cf. Lith. *šerti* 'feed', Gr. κορέννυμι 'satiated') with 'nourishing substance, kernel' as intermediate notion.

Another well-known name for 'grain' was PIE \**ǵéu-o-* (NIL: 407–410), cf. Ved. *yáva-* m. 'barley, wheat, grain' (= YAv. *yauua-*), Hitt. *ewa-* (*ewan-*) 'name of a cereal', Gr. ζειαί f.PL. 'wheat', Lith. *jāvas* m. 'wheat', PL. *javai* 'wheat grains', OArm. *jov* 'sprout'. The possibility of a final laryngeal (PIE \**ǵéu(h<sub>1</sub>)-o-*) was assumed because

of a wrong etymological connection with Ved. *gáv-yū-ti-* F. ‘pasture’ which is unrelated according to Nikolaev (2014: 131). According to Ivanov (2003: 195 ff.), we are dealing with the PIE root *\*ǵieǵ-* ‘to bind, mix’ (*LIV*<sup>2</sup>: 314) reflected by AVed. *yaúti* ‘to unite, bind’ and Lith. *yaūti* ‘to mix’, on which an adjective *\*ǵieǵ-ó-* ‘mixed’ was built (on the same pattern as Gr. λευκός ‘white’). We may assume that the barytonesis is a marker of nominalization (PIE *\*ǵieǵ-o-* M.PL. ‘mixed grains’). The original meaning was probably \*‘mixed fodder for cattle’.

Words like ‘grain’, ‘awn’ in themselves do not necessarily indicate agriculture: knowledge of such notions is consistent with the collecting of wild cereals, as are words for grinding. Conspicuously lacking in the earliest Proto-Indo-European vocabulary are words for notions that unequivocally indicate agriculture: sowing, weeding, harvesting, fields, seeds for sowing, as well as stable names for domesticated cereals. The Austronesian family, the other model for the Farming/Language Theory (Bellwood 1985), has a much stronger claim of having arisen at least partly as a result of a shift to agriculture: Austronesian vocabulary reconstructable at the highest level includes all the notions (‘to sow broadcast’, ‘to weed’, ‘to harvest’, ‘field’, ‘seeds for sowing’) that are missing in Proto-Indo-European, plus the names of three domesticated cereals: foxtail millet, broomcorn millet and rice (Sagart et al. in press). Proto-Indo-European therefore cannot have been the language of a group of farmers, whether in Anatolia or elsewhere. Instead, Proto-Indo-European vocabulary at the highest level (i.e. including Anatolian) is animal-oriented, with stable names for bovines and ovines, animal fodder, and cattle-drawn carts, at least.

While we think the Anatolian theory is in all likelihood incorrect, we regard the idea that the formation of a language family normally implies demographic expansion as a precious insight of the Farming/Language theory. In this paper, we propose that a demographic mechanism explains part of the success of the Indo-European languages and the demise of the languages that preceded them, although the mechanism we have in mind is different from Renfrew’s. In the first part, we report on recent strands of research in archaeology, human genetics and the early history of dairying. These give additional support to the Pontic Steppe hypothesis by showing that the speakers of Proto-Indo-European were the first in Eurasia among whom the ability to drink milk into adulthood developed, and that this ability became dominant in western Eurasia as a result of Indo-European expansions.

In the second part, we examine the Indo-European dairy vocabulary, especially the verb ‘to milk’ and the noun ‘milk’, and describe historical changes in this vocabulary that testify to the rise of milking activities and the growing importance of animal milk in the early Indo-European diet. Second, we reproduce ancient textual evidence associating adult milk drinking with Indo-European, especially Indo-Iranian, speakers. Finally, we document the earliest evidence for adult milk drinking based on parallel expressions from the ritual Indo-Iranian literature.

In conclusion, we argue that lactose tolerance provided the early Indo-Europeans with a demographic edge and possibly with an increase in physical stature, both leading to military advantage over preexisting farming communities that were economically successful but lacking in the political means to mount a coordinated resistance. Elite dominance of Indo-European speakers led to widespread language shift towards Indo-European dialects on the part of farmers, explaining the success of Indo-European languages over those of their European farming predecessors.

## 1. The archaeological and genetic background

Recent archaeological and genetic work has provided decisive evidence for the Pontic Steppe theory. Haak et al. (2015) showed that a massive migration of Yamnaya hunter-gatherers out of the Pontic steppes into the Corded Ware culture of NW Europe ca. 4500 years ago established a new population component there, distinct from both palaeolithic hunter-gatherers and early European farmers who had previously spread from Anatolia. Further, in a study of ancient DNA from 101 Bronze Age Europeans, Allentoft et al. (2015) showed that the highest levels of a gene allowing adults to digest lactose and consume raw milk are found in the burials of Yamnaya culture and its offshoots the Corded Ware and Afanasievo cultures. They state that by 3000 BCE Yamnaya culture had replaced Neolithic farmers from Hungary to the Urals: they regard the Corded Ware culture of northwestern Europe as possibly derived from Yamnaya, but also including Neolithic farmers. They date its establishment at 2800 BCE. Despite the differences in dates, both Haak et al. and Allentoft et al. link the westward Yamnaya migration with the spread of Indo-European languages in Europe; Allentoft et al. further argue that the spread of lactose tolerance in Europe is due to Indo-European expansions.

Different strands of recent work on dairying in Neolithic Europe provide useful background on the development of lactose tolerance in Europe. As recently as 7000 years ago all human populations were lactose-intolerant (Leonardi et al. 2012): adults lacked the enzyme lactase and could not digest the sugar lactose contained in milk. Lactose tolerance arose independently in several of the world's populations, both in Africa and Eurasia. As for Eurasia, the areas of maximum lactase persistence, as mapped by Leonardi et al., broadly coincide with the Corded Ware culture in NW Europe and with a zone centered on coastal Pakistan, extending into southeastern Iran and northwestern India. This is consistent with a link between lactose tolerance and the spread of Indo-European speakers.

The invention of cheese, a milk derivative poor in lactose, by early farmers in Northwest Anatolia ca. 8500 BP (Evershed et al. 2008) for the first time allowed

humans to turn animal milk into a stable source of food. This presumably contributed to the positive demography of early farming populations. As they spread over Europe, the farmers brought cheese-making with them (Salque et al. 2013). However, they were themselves largely lactose-intolerant (Burger et al. 2007; Allentoft et al. 2015): the capacity to directly drink animal milk results from a genetic mutation allowing the enzyme lactase to persist in adults, a mutation which only arose a few millennia later. We follow Burger (oral remarks cited in Owen 2010) in supposing that contact with cheese-making farmers revealed the lactase persistence gene in certain hunter-gatherer individuals from the Pontic steppes, and that this beneficial gene was subsequently strongly selected for. Presumably, the incidence of the gene rapidly increased in the Yamnaya population, fostering population growth; increased reliance on animal milk required more pasture lands; these became scarce in the homeland area, leading to migrations and territorial expansions – towards Afanasievo culture in the Minusinsk basin before 3000 BCE (perhaps ancestral to the Tocharians); towards northern Europe and towards the Andronovo culture (perhaps ancestral to Indo-Aryan) around the Sea of Aral in the early/mid-second millennium BCE.

In the next section we examine the linguistic and philological evidence on the place of milk in the early Indo-European diet.

## 2. Linguistic and philological evidence on the place of milk among Indo-Europeans

In this section, we examine the Proto-Indo-European word for ‘to milk’ (2.1), starting with Hittite (2.1.1), then moving to the Core IE root  $*h_2melǵ-$  ‘to milk’ (2.1.2). In an excursus in Section 2.1.3 we discuss the Indo-Iranian root  $*d^hauǵ-$ , both ‘to milk’ and ‘to give milk’. We next move on to the noun ‘milk’: we first examine languages where both ‘to milk’ and ‘milk’ are from  $*h_2melǵ-$  and those where only ‘to milk’ is from  $*h_2melǵ-$  (2.2). We discuss, and reject, the widely accepted equation between Gr. γάλα N. ‘milk’ and OArm. *kat’n* ‘id’, proposing a new etymology for Gr. γάλα (2.2.1). Our new etymology for Lat. *lac*, *lactis* N. ‘milk’ (2.2.2) tentatively places it under the root IE  $*h_2melǵ-$  ‘to milk’. We then show that the Core IE root  $*h_2melǵ-$  ‘to milk’ (2.2.3) is secondary, suggesting it originates in a Core IE compound  $*h_2m̥H-léǵ-$ , GEN.  $*h_2m̥H-lǵ-ós$  ‘he who collects ( $*leǵ-$ ) liquids/milk’. In Section 2.3, we scan Greek and Latin texts for evidence of milk-drinking among “barbarian” adults: Homer and Homeric scholia (2.3.1), Hesiodus (2.3.2), Hippocrates (2.3.3), Herodotus (2.3.4) and Pliny the Elder (2.3.5), showing that all such references point to speakers of Indo-European languages. In a conclusion to Section 2 (2.3) we note the involvement of milk with Indo-Iranian ritual, pointing

out that it prescribes milk drinking by adults. Finally, in our general conclusion we describe the demographic and biological mechanisms through which milk-drinking promoted the spread of Indo-European languages and the demise of the languages of early European farmers.

## 2.1 Indo-European words for ‘to milk’

### 2.1.1 Hittite

Hittite does not have a specialized verb ‘to milk’. Milking was practiced but the texts either use the Hittite root *lā-* ‘to let, make flow’ (< PIE *\*leh<sub>1</sub>-* ‘to let’), for instance GA *lāttat* ‘he let the milk flow = he milked’ (KBo III 8 III 30–31), or the locution GA *ḫamikta* ‘he pressed the milk, he milked’ (KBo III 8 III 12–13), where GA, the sumerogram for ‘milk’, is more probably an accusative of product or result than an accusative of direct object. The verb *ḫamikta* ‘he pressed’ is from the nasal-infixes present stem *ḫamink-* ‘to tie together, press together’ (< PIE *\*h<sub>2</sub>emǵh-* ‘to squeeze; narrow’). A third expression occurs in Hittite texts: *ḫūratiššan ḫamikta* ‘he squeezed the udder’ (KBo III 8 III 12–13). This makes it likely that like Anatolian, its primary branch, Proto-Indo-European lacked a specialized root for the verb ‘to milk’. However, this is only an *argumentum ex silentio*.

The Hittite name for ‘milk’ cannot be recovered due to generalized use of the sumerogram GA: consequently the Proto-Indo-European word for ‘(to) milk’ cannot be known either. Only a Proto-Indo-European root for ‘to suck mother’s milk’ is known: *\*d<sup>h</sup>eh<sub>1</sub>-* (*LIV*<sup>2</sup>: 138 ‘Muttermilch saugen’). The same root (with *\*-i-* extension) is well attested in Anatolian (cf. Hitt. *tēdan* ‘teat’ < *\*d<sup>h</sup>éh<sub>1</sub>-i-tom*).<sup>1</sup> The archaic reduplicated neuter stem PIE *\*d<sup>h</sup>éd<sup>h</sup>h<sub>1</sub>-i* ‘mother’s milk’ (Ved. *dādhi*, *dadhnás* N. ‘thick sour milk’) underwent a sporadic shift to a generic name for ‘milk’, as is clear from OPr. *dadan* N. ‘milk’.

### 2.1.2 The Core IE root *\*h<sub>2</sub>melǵ-* ‘to milk’

This root is widespread among Indo-European languages outside of Anatolian: Lat. *mulgeō*, *ere* ‘to milk’ (< PIE iterative stem *\*h<sub>2</sub>molǵ-éǵ-e/o-*), reflected by Rom. *mulge*, It. *mungere*, OFr. *moudre* (< Vulg. Lat. *\*mulgērē*); Gr. *ἀμέλω* ‘to milk’ (< PIE root present *\*h<sub>2</sub>mélǵ-e/o-*), whence MoGr. *ἀλέω* (Vulg. *ἀρμέγω*); Lith. *milžti* (present stem *mélžu*) ‘to milk’; OCS *mlěšti* ‘id.’; Com. Germ. *\*mel(u)kan<sup>an</sup>* (OE *melcan*, MoGerm. *melken* and *melchen*). Albanian agrees with the reflex of a *e*-grade present stem as well (Alb. *mjell* ‘to milk’). Common Celtic is unique among Indo-European

1. According to Kloekhorst (2008: 877), the lenition is triggered by the preceding accented diphthong.

languages of Europe in reflecting a zero-grade thematic root present: Com. Celt. \**mlig-e/o-* ‘to milk’ (< PIE \**h<sub>2</sub>mlǵ-é/ó-*), whence OIr. *bligim* ‘id.’ and Gallo-Rom. \**blig-āre* ‘id.’ (< Gaul. \**blig-*) reflected by OFr. *blechier* ‘to milk’, mostly famous for its designation of a French cheese: *Roblochon* (or *Re-*), which is made from milk of a second milking (cf. OFr. *re-blechier* ‘to milk a second time’). Root \**h<sub>2</sub>melǵ-* ‘to milk’ is also found in the very far east of the Indo-European domain: Toch. B *malower* n. ‘milk’ and Toch. A *malke* ‘id.’. These Tocharian nominal stems are not likely to be directly inherited from the *Ursprache*: rather, they point to an unattested verb Com. Toch. \**mälk-* ‘to milk’ (< Core IE \**h<sub>2</sub>mlǵ-*).

It is noteworthy that there is no evidence at all for root \**h<sub>2</sub>melǵ-* ‘to milk’ in Indo-Iranian, not even in the modern dialects. The Vedic Narten present *márj-mi* ‘to rub’, sometimes presented as related to \**h<sub>2</sub>melǵ-* (e.g. Mayrhofer *EWAia* II: 325), must in fact relate to a distinct root, namely \**h<sub>2</sub>merǵ-* ‘to wipe clean, cleanse, purify, remove completely’.<sup>2</sup> The two roots have largely non-overlapping semantics, although the derived meaning ‘to pluck’ in Gr. *ἀμέργω* ‘pluck’ (always applied to plant products) could be construed as similar to the action of milking.<sup>3</sup> The fact that Vedic *márj-mi* ‘to rub’ and other Indo-Iranian forms under the Vedic root *MRJ-* ‘to wipe, brush’ seem to regularly reflect \**h<sub>2</sub>melǵ-* ‘to milk’ is the result of the merger of PIE \**l* and \**r* in Indo-Iranian: Ir. \**marj-* can reflect both \**h<sub>2</sub>melǵ-* and \**h<sub>2</sub>merǵ-*. In addition, the initial laryngeal \**h<sub>2</sub>* in \**h<sub>2</sub>merǵ-* is problematic: there is no reflex of it in Vedic or in Avestan – Mayrhofer’s reluctance to assume an Indo-Iranian etymon \*(*H*)*marj-* is understandable (*ibid.*).<sup>4</sup> The proposed \**h<sub>2</sub>-* relies exclusively on initial *ǎ-* in Gr. *ἀμέργω*: but the alternation with initial *ó-* in the related form *ὀμόργνυμι* ‘to dry’ (< \*‘to rub, wipe out’) is not consistent with \**h<sub>2</sub>-*.<sup>5</sup> It is more probable that initial *ǎ-* in *ἀμέργω* is the fruit of contamination from the phonetically and semantically similar, but etymology distinct verb *ἀμέρδω* ‘to deprive, take away’.<sup>6</sup> The by-form *ὀμόργνυμι* ‘to dry’ itself is analyzable as an old preverbed zero-grade stem \**h<sub>2</sub>o-mrǵ-néu-*, with a dialectal reflex of \**r*. As a result, \**h<sub>2</sub>merǵ-* should be emended to \**merǵ-*, without a laryngeal initial, removing it further away from \**h<sub>2</sub>melǵ-* ‘to milk’.

2. Cf. Late Av. *ni-marəzišta-* ‘best cleanser’ (of Ahura Mazda).

3. E.g. in *καρπὸν ἀμέργουσιν πεποτημέναι* ‘they pluck the fruit on their wings’, of bees (AP 1. 882).

4. As pointed out by a reviewer, the lengthened reduplication of the Vedic perfect *māmri-* may be considered an argument for \**Hmarj-* but it must be admitted that it is not very strong.

5. *Mid.*: ‘to dry oneself’ (most often tears), # *δάκρυ ὀμορξαμένην* ‘drying her tears’ (λ 530).

6. PIE \**h<sub>2</sub>merd-* ‘to harm, mistreat’ (*LIV*<sup>2</sup>: 280, s.v. \**h<sub>2</sub>merd-* ‘ein Leid antun, mißhandeln’). Note the confusion between *ἀμέρσᾶς* and \**ἀμέρξᾶς* in AP 7.657.7, pointed out in *BDAG*: 2015, 107, s.v. *ἀμέρδω*.



This in fact suggests a likely explanation for the lack of Indo-Iranian reflexes of  $*h_2melĝ-$  ‘to milk’: homonymic clash with a verb ‘to rub, wipe etc.’ reflecting  $*merĝ-$  may have caused Indo-Iranian speakers to replace  $*h_2melĝ-$  ‘to milk’ with an innovated form, in this case  $*d^hauĝ^h-$  (discussed in Section 2.1.3. below). Homophony between a verb ‘to rub’ and a verb ‘to milk’ would have been particularly undesirable, since rubbing a cow’s udders during milking is painful to the animal, causing it to balk, as is well known to those who practise milking. A homophonic clash of  $*h_2melĝ-$  ‘to milk’ and  $*h_2merĝ-$  ‘to rub, etc.’ occurs only in Indo-Iranian because only Indo-Iranian does lose the distinction between  $*l$  and  $*r$ . In Section 2.2.5., we will propose a new etymology for  $*h_2melĝ-$  ‘to milk’.

To sum up, Vedic *mārj-mi* and other forms under the Vedic root *MRJ-* ‘to wipe, brush’ may be connected to Gr. *ἀμέργω* ‘pluck’ (earlier *\*μέργω*): both are from a root  $*merĝ-$  without a laryngeal, and without any significant connection to  $*h_2melĝ-$  ‘to milk’.

### 2.1.3 Excursus: Indo-Iranian $*d^hauĝ^h-$ ‘to milk; to give milk (of a cow)’

It is generally assumed that the IIr. root  $*d^hauĝ^h-$  ‘to milk; to give milk’ directly reflects PIE  $*d^heug^h-$  ‘to be efficient’ (Mayrhofer *EWAia* I: 747–8), making it a very ancient root and raising the possibility that Proto-Indo-European may have had another verb ‘to milk’ competing with  $*h_2melĝ-$ . Indeed, the Vedic verb exhibits a very archaic conjugation pattern, associating an athematic root active present in PIE 3SG.  $*-ti$ , 3PL.  $*-énti$ : *dōgdhi* 3SG., *duh-ánti* 3PL. ‘to milk (a cow), extract (soma)’ (< PIE  $*d^héug^h-ti$ ,  $*d^heug^h-énti$ ) with a middle present in PIE 3SG.  $*-óĵ$ , 3PL.  $*-róĵ$ : *duh-é* 3SG., *duh-ré* 3PL. (< PIE  $*d^huĝ^h-óĵ$ ,  $*d^huĝ^h-róĵ$ ). This supports the Indo-Iranian verb’s Proto-Indo-European antiquity and is consistent with a link to the PIE root  $*d^heug^h-$ , at least on a phonological plane.

At the same time, in the languages (outside of Indo-Iranian) where it is attested, the root  $*d^heug^h-$  is unrelated to milk: Gr. *τεύχω*, ‘to do, make, prepare, build’, Com. Germ.  $*dugan^{an}$  (intr.) ‘to be fit, avail’ ~  $*daug-$  (*o*-grade) ‘id.’ (Go. *daug* 3SG.prf-prs. ‘id.’, G. *taugen* ‘id.’). In addition, there are no expressions using the PIE root  $*d^heug^h-$  and meaning ‘to produce milk’, whether in Greek, Germanic or Indo-Iranian. Moreover, a semantic shift from ‘to produce’ to ‘to milk’ strikes us as unmotivated. These points seem to argue that the IIr. root  $*d^hauĝ^h-$  ‘to milk; to give milk’ acquired its connections to milk no earlier than Indo-Iranian, and not as a result of a straightforward semantic shift.

Based on an old suggestion of Szemerényi, we attempt a new solution to this conundrum. Almost sixty years ago, Szemerényi (1958: 171, fn. 3) suggested that the IIr. root  $*d^hauĝ^h-$  ‘to milk’ is a back-formation from the Indo-Iranian name for ‘daughter’ (IIr.  $*d^huĝ^{(h)}-H-tár-$ ), which he thought had originally meant ‘suckling child’ or the like. Szemerényi’s proposal has against it the fact that a back-formation

in Indo-Iranian times from ‘daughter’ could not have possessed the archaic conjugation pattern of IIr.  $*d^h\acute{aug}^h$ -. His hypothesis has met with a great deal of resistance among scholars. Yet it can be adapted as follows. We assume an unattested action noun PIE  $*d^h\acute{e}ug$ - $h_2$ - $e/os$ - N. (<  $*d^h\acute{e}(h_1)$ - $u$ - $g$ - $h_2$ - $e/os$ ) ‘action of sucking mother’s milk’, ultimately based on the PIE root  $*d^h\acute{e}h_1$ - ‘to suck mother’s milk’, whose  $u$ -stem PIE  $*d^h\acute{e}(h_1)$ - $u$ - ADJ. ‘female, breastfeeding’ had a velar enlargement  $*d^h\acute{e}(h_1)$ - $u$ - $g$ - with a concrete meaning ‘teat (*vel sim.*)’. This secondary derivative served as the basis for an amphidynamic abstract noun PIE  $*d^h\acute{e}(h_1)$ - $u$ - $g$ - $h_2$  (GEN.SG.  $*d^h$ - $u$ - $g$ - $\acute{e}h_2$ - $s$ ) ‘femininity’. From this hypothetical form the Proto-Indo-European name for ‘daughter’, containing an athematic variant of the “characterizing” suffix  $*-ter-o$ - (Pinault 2007) can be derived:  $*d^h(h_1)$ - $u$ - $g$ - $h_2$ - $tér$ -. Semantically a daughter would then be a ‘suckling [female] child’, or, perhaps more convincingly, a person giving suckle, assuming the term first designated daughters of child-bearing age. Because PIE  $*gh_2$  and  $*g^h$  merge as  $*g^h$  in Indo-Iranian – and nowhere else – the secondary derivative PIE  $*d^h\acute{e}ug$ - $h_2$ - $e/os$ - N. would have resulted in IIr.  $*d^h\acute{aug}^{(h)}$ - $H$ - $as$ ,  $*d^h\acute{auj}^{(h)}$ - $H$ - $as$ - N. ‘sucking’ (whence also ‘milking’). This term is in fact attested as Ved.  $dóh$ - $as$ - ‘milking’. There is another possibility: a thematic secondary derivative PIE  $*d^h\acute{oug}$ - $h_2$ - $o$ - M. ‘id.’ reflected by Ved.  $dógham$  ‘milking’ (hap. leg.) and by Pašto *lway* ‘id.’ (< Com. Ir.  $*daug$ - $a$ -).<sup>7</sup> As a result of the phonological merger, to Indo-Iranian speakers,  $*d^h\acute{aug}^{(h)}$ - $H$ - $as$  N. ‘sucking’ or  $*d^h\acute{aug}^{(h)}$   $H$ - $a$ - M. ‘id.’ would have seemed to contain the homophonous – but unrelated – primary IIr. root  $*d^h\acute{aug}^h$ - ‘to be efficient, produce’. This would have resulted in the appearance of a hybrid verb, combining the archaic conjugation pattern of root  $*d^h\acute{e}ug^h$ - and the milk-related semantics of the action noun PIE  $*d^h\acute{e}ug$ - $h_2$ - $e/os$ - (or its thematic by-form  $*d^h\acute{oug}$ - $h_2$ - $o$ -).

## 2.2 Indo-European words for ‘milk’ derived from ‘to milk’

We have argued that the Core IE root  $*h_2melǵ$ - ‘to milk’ is an innovative form, since Hittite has no specialized root for ‘to milk’. Two sets of languages may be distinguished with respect to this root: (1) those where both the verb ‘to milk’ and the noun ‘milk’ are from  $*h_2melǵ$ - (Table 1), and (2) those where only the verb ‘to milk’ is from  $*h_2melǵ$ - (Table 2). The situation in Tocharian is more complex: the nouns for ‘milk’ in the two dialects: Toch. A *malke* ‘milk’, B *malkwer* ‘id.’, have different Common Tocharian etymologies: *malke* is from Com. Toch.  $*mælk$ - $æy$  (< IE  $*h_2molǵ$ - $óǵ$ -), a secondary derivative built on the (isolated) IE action noun

7. The Pašto *lwaš* ‘to milk’, from Com. Ir.  $*dauš$ - $a$ - ‘id.’, may rather reflect IIr.  $*dauk$ - ‘to milk’, from IE  $*deuk$ - ‘to draw’ (e.g. Ossetic *doc*-, Waxi *dic*-). According to Cheung (2007: 66f.), the reconstruction  $*dauš$ - is not secure, since most verbs in question can also be explained from Com. Ir.  $*daušya$ - < IIr.  $*daučja$ - which is required for Ossetic anyway.

\* $h_2$ *molǵ-í-* F. ‘milking’, whereas Toch. B *malk-wer* is a secondary derivative built on a verbal stem Com. Toch. \**mälk-* ‘to milk’ (< IE \* $h_2$ *mǵ-*).

The fact that the languages where the verb is derived from \* $h_2$ *melǵ-* are a subset of those where the noun from \* $h_2$ *melǵ-* argues in favor of the hypothesis that the nouns are derived from the verb. However, there is clear evidence that we are not dealing with a single innovation: in each language where both the noun and the verb reflect \* $h_2$ *melǵ-*, the noun has the same vocalic grade as the verb: therefore, terms for ‘milk’ must have been derived independently in the daughter languages of Core Indo-European.

**Table 1.** Languages where both ‘to milk’ and the name for ‘milk’ are from \* $h_2$ *melǵ-*

Com. Germ.	* <i>melk-a-</i> ‘to milk’	* <i>melk-</i> F. ‘milk’ (root noun)
Com. Celt.	* <i>mliġ-e/o-</i> ‘to milk’	* <i>mliχtos</i> M. ‘milk’
Com. It.	* <i>molġ-éj-e/o-</i> ‘to milk’	* <i>mlóktō-</i> M. ‘milking’ * <i>mlaktá</i> F. ‘milk flow’ (2.2.4)

**Table 2.** Languages where only the verb ‘to milk’ is from \* $h_2$ *melǵ-*

Greek	<i>ἀμέλω</i> ‘to milk’	<i>γάλα, γάλακτος</i> N. ‘milk’
Lithuanian	<i>mélžu</i> ‘to milk’	<i>pienas</i> M. ‘milk’ (Latv. <i>piēns</i> )
Albanian	<i>mjel</i> ‘to milk’	<i>dhallë / dhalltë</i> F. ‘buttermilk’
Com. Slav.	* <i>melz-ti</i> ‘to milk’	(* <i>melkò</i> ‘milk’ < Com. Germ.) <sup>8</sup>

Because the nouns for ‘milk’ in Table 1 are derived from the verb ‘to milk’, their original referent must have been ‘animal milk’ rather than ‘mother’s milk’. Other innovative Indo-European words for ‘milk’ are not derived from the verb ‘to milk’. Lith. *pienas* M. ‘milk’ and Latv. *piēns* ‘id.’ reflect an IE masculine stem \**póǵH-no-* ‘thick fluid, \*mother’s milk’ (the acute intonation of Lith. *pienas* for expected \*\**piēnas* according to Saussure’s effect is analogical to the Lith. verb *pýti* ‘to have milk’).

From the underlying IE root \**peǵH-/peǵH-* ‘to be thick’ (*LIV*<sup>2</sup>: 464 ‘anschwellen’), a neuter stem \**peǵH-mǵ* ‘thick fluid’ was also built. This is reflected by OAv. *paēman* N. ‘mother’s milk’, MidPers. *pēm* ‘milk’. This term was also borrowed by Fin. *piimä* ‘sour milk’. The Iir. etymon \**páǵH-as-* N. (Ved. *páyas-* ‘Lebenskraft’, OAv. *paiiah-* ‘milk’) reflects IE \**peǵH-e/os-* N. ‘thick fluid’. On the zero grade of *peǵH-/peǵH-* ‘to be thick’, an adjective \**piH-iú-* ‘thick’ was built, whence the abstract noun \**piH-iú-h<sub>2</sub>* F. ‘thickness’. That word became the starting point for a secondary derivative \**piH-iu-h<sub>2</sub>-s-ó-* ‘thick fluid’ (cf. Ved. *piyúṣa-* M.N. ‘colostrum, the milk of

8. A borrowing from Germanic (Derksen 2008: 307). Ru. *молозиво* N. ‘colostrum, beestings’ is inherited.

a cow during the first seven days after calving, biestings, any thick fluid'). According to Garnier (2016b: 1.8), the primary root is PIE  $^{*}(s)peh_1-$  'to swell, get fat, fatten, thrive', with an acrostatic neuter PIE  $^{*}(s)póh_1-i$  'fat'. The adjective PIE  $^{*}(s)pəh_1-i-tó-$  'full of fat, having corpulence' underwent the regular metathesis of laryngeals: PIE  $^{*}(s)pih_1-tó-$  'fattened, fat' and PIE  $^{*}pih_1-nó-$  'swollen', reinterpreted as participles of a secondary root IE  $^{*}(s)pejh_1-$  'to be fat, be thick'.

### 2.2.1 Other innovative forms for 'milk': Gr. γάλα N. and OArm. kat'n.

An etymological link is widely assumed between Gr. γάλα N. 'milk' and OArm. kat'n 'id.' (Dial. Arm. kaxc'). A link cannot be taken for granted. According to Martirosyan (2010: 345–6), the Armenian forms reflect a proto-Arm. paradigm NOM.SG. \*kac' (< \*kalc' < PIE  $^{*}g_lk-t-s$ ), acc.SG. \*katt'n (< PIE  $^{*}g_lk-t-m$ ), levelled to \*kac', \*kat'n; this in OArm. kat'n, used as both nominative and accusative while the dialects exhibit the symmetrical levelling \*kalc', \*katt'n with analogically re-introduced velar *l*, and extension of the old nominative \*kalc' (Dial. ModArm. kaxc') throughout the whole paradigm. This theory relates the Armenian forms to an IE etymon  $^{*}g_lk-t-s$  and appears to provide a viable link to Gr. γάλα, γάλακτος. However, it stumbles upon three obstacles. First, the animate gender of the PIE etymon  $^{*}g_lk-t-s$ , acc.  $^{*}g_lk-t-m$ , does not match the neuter gender of Gr. γάλα. Second, from a Greek point of view, the unexpected disyllabism of the stem γάλακτ- is hardly compatible with an original stem  $^{*}g_lk-t-$ . Third, as recently demonstrated by Kümmel (2017: 445f.), the inner-Armenian connection of kat'n with kit' 'milking; harvest' and kowt' 'harvest' is no longer compatible with a reconstruction  $^{*}g_lk-t-$ . An alternative and perhaps preferable explanation is to posit an etymological link between Gr. γάλα and Alb. dhallë / dhalltë f. 'butter milk', reflecting a Proto-Alb. \*dzalā- 'id.' (whence also the Rom. loanword zară 'id.'), where \*d regularly reflects a PIE palatal  $^{*}g$ , not a pure velar  $^{*}g$ . On the basis of the Homeric formula γάλα λευκόν 'white milk' (Δ 434, E 902), we propose an origin of the Greek and Albanian forms in a color adjective Gr. γάλαξ, -ακος  $^{*}$ 'white'; this form is actually attested, with the meaning 'a kind of a shell, prob. *Mactra lactea*' (Aristot. HA 528a 23). *Mactra lactea* is white in color. This adjective, reflecting a PIE stem  $^{*}g_lh_2-n-k-$  'bright, white' from PIE  $^{*}g_lh_2-$  'to shine', could have resulted in a substantivized neuter Gr. γάλα. The dental stem of GEN.SG. γάλακ-τος would be secondary. As a typological parallel, we may mention MoAr. laban M. 'milk, whey' with root LBN- 'to be white'.

### 2.2.2 Other innovative forms for ‘milk’: Lat. *lac*, *lactis* N.

Contrary to a tenacious legend,<sup>9</sup> this Latin word has nothing to do with Gr. γάλα, γάλακτος N. ‘milk’. Garnier (2016: 306–7) proposes that Lat. *lac*, *lact-is* N. ‘milk’ is a back-formation.<sup>10</sup> The stem *\*lact-* would be from an unattested verb *\*amb-lactāre* ‘to milk’<sup>11</sup> resulting from *\*ambī-blactāre* ‘to milk with both hands’ through haplology; *\*amb-lactāre* itself underwent depreverbatation to *lactāre* ‘to milk’, and a stem *\*lact-* ‘milk’ was extracted through back-formation. The underlying Proto-Indo-European root must have been IE *\*h<sub>2</sub>melǵ-* ‘to milk’ (cf. Gr. ἀμέλω, Lat. *mulgĕō*). We may envision an action noun IE *\*h<sub>2</sub>mólǵ-to-* M. ‘the milking’, regularly metathesizing to Common Italic *\*mlók-to* M.; this then further affected with Italic collective suffix *-ā* of concrete meaning (< IE *\*-eh<sub>2</sub>*), giving ‘milk flow’; affixation of *-ā* in turn required change to zero degree still in Italic. One would have expected *\*molk-tá* (< *\*młk-tá*) but due to analogy with the strong stem *\*mlók-to*, resyllabification resulted in *\*mlāk-tá* F. ‘milk flow’, coexisting with *\*mlók-to* at Common Italic level.<sup>12</sup> In turn, *\*mlāk-tá* regularly evolved to unattested Lat. *\*blactā-*, out of which *\*ambī-blactāre* ‘to milk with both hands’ was formed.

### 2.2.3 A new etymology for IE *\*h<sub>2</sub>melǵ-* ‘to milk’

The IE root *\*h<sub>2</sub>melǵ-* ‘to milk’ is phonologically too complex to be primary. Its meaning is both highly specialized and remarkably stable across languages, despite widespread attestation, suggestive of a relatively recent formation. Benveniste (1935: 157) assumed a primary root *\*h<sub>2</sub>em-* ‘to collect liquid’ (cf. Gr. ἄμη F. ‘bucket’) with nominal enlargement PIE †*h<sub>2</sub>m-el-g-* in his notation. As a parallel to the proposed *-el-g-* enlargement, he cited Ved. *s<sup>u</sup>vargá-* ADJ. ‘heavenly’ which he took to be from PIE †*su-él-g-*. Kümmel (*LIV*<sup>2</sup>: 265) reconstructs the same root with a final laryngeal: PIE *\*h<sub>2</sub>emH-* ‘to pour’. Reflexes are Com. Celt. *\*ande=am-je/o-* ‘to pour [water] upon’ (cf. OIr. *and.aim* ‘to wash’) and the doublet Com. Celt. *\*ad=am-je/o-* ‘id.’ (cf. v.-irl. *ad.aim*), supported by Matasović (2009: 31). The fact that *\*h<sub>2</sub>emH-* was

9. Szemerényi (1991: 1117) and Leumann (1977: 187) assume for Lat. *lac*, *lactis* a PIE etymon *\*ǵlakt-* N. ‘milk’ supposedly also explaining the Greek forms. This etymology is maintained by Weiss (2011: 147, fn. 82).

10. Archaic nominative *lact* in Varro (*Men.* 26), deemed incorrect by Julius Caesar according to Pompeius Grammaticus (*GLK* 5: 199). Vulgar form *lactĕ* in Plautus (*Bacch.* 6), prefiguring the Romance evolution (cf. It. *latte*).

11. The by-form *lactĕ* (pl.) would be a back-formation from a vulgar doublet *\*lactiāre* ‘to milk’.

12. Such a resyllabation may be paralleled by OHG *nusta* F. ‘Verbindung’ (< Com. Germ. *\*nustō<sup>n</sup>*), analogous to the strong stem Com. Germ. *\*násta<sup>z</sup>* M. ‘binding’ (< PIE *\*Hnód<sup>h</sup>-to-*), instead of phonetically expected *unstō<sup>n</sup>* according to Griepentrog (1995: 457).

the verb used for pouring or collecting milk is clear from textual evidence: Hom. ἀμάομαι ‘to draw milk, collect’ (mid. ἀμάομαι) is said of curdled milk in ι 247:

αὐτίκα δ’ ἤμισυ μὲν θρέψας λευκοῖο γάλακτος  
 πλεκτοῖς ἐν ταλάροισιν ἀμησάμενος κατέθηκεν,  
 ἤμισυ δ’ αὐτ’ ἔστησεν ἐν ἄγγεσιν, ὄφρα οἱ εἴη  
 πίνειν αἰνυμένῳ καὶ οἱ ποτιδόρπιον εἴη.

He curdled half the white milk and *collected* it in wicker strainers, but the other half he poured into bowls so that he might drink it for his supper.

The secondary derivative Gr. ἄμης, -ητος m. ‘milk cake’ (Aristoph., *Ploutos* 999) is perhaps from an unattested masculine or neuter o-stem \*ἄμος ‘milk left to curdle in a bucket, curdled milk’. The Greek word ἄμη f. ‘bucket’ is considered a back-formation (Dieu 2016: 112). We propose that IE \**h*<sub>2</sub>*melǵ-* ‘to milk’ is a secondary root based on a compound \**h*<sub>2</sub>*m̥H-léǵ-*, GEN. \**h*<sub>2</sub>*m̥H-lǵ-ós* ‘one who collects (\**leǵ-*) liquids/milk’. We assume this compound dates back to a period preceding the formation of Core Indo-European since its derivatives meaning ‘to milk’ and ‘milk’ are widespread in the daughter branches of Core Indo-European, including Tocharian. Phonologically, the old GEN.SG. \**h*<sub>2</sub>*m̥H-lǵ-ós* resulted in \**h*<sub>2</sub>*m̥.lǵ-ós* with hiatus, whence resyllabation as \**h*<sub>2</sub>*m̥lǵ-ós*. For a parallel, cf. the resyllabation in the Proto-Indo-European name for ‘wind’: PIE \**h*<sub>2</sub>*ueh<sub>1</sub>-nt-ó-* > \**h*<sub>2</sub>*ue.nt-ó-* > \**h*<sub>2</sub>*uent-ó-* m. ‘wind’ (cf. Go. *winds*, Lat. *uentus*), a derivative of appurtenance (‘the fast one’) built on the PIE *nt*-stem \**h*<sub>2</sub>*uh<sub>1</sub>-ónt-*, \*-*nt-és* ‘running’ (Garnier 2014: 63). Finally, through back-formation IE \**h*<sub>2</sub>*m̥lǵ-* ‘milker’ would have triggered the creation of the secondary root \**h*<sub>2</sub>*melǵ-* ‘to milk’, out of which several words for ‘milk’, described above, would later be derived: pre-Core IE \**h*<sub>2</sub>*m̥lǵ-* ‘milker’ → Core IE \**h*<sub>2</sub>*melǵ-* ‘to milk’ → Post-Core IE names for ‘milk’.

## 2.3 Greek and Latin textual evidence for milk-drinking among Indo-European “barbarians”

### 2.3.1 *Homer and Homeric scholia*

Homer’s *Iliad* already alludes to milk-drinking among a legendary people of pastoral nomads referred to as “the lordly Hippemolgi”; Herodotus mentions the (Indo-Iranian) Scythians, who drink mare’s milk. Let us start with the very beginning: Homer’s *Iliad*, in which dairy culture was first depicted.

αὐτὸς δὲ πάλιν τρέπεν ὄσσε φαεινῶ  
 νόσφιν ἐφ' ἰπποπόλων Θρηϊκῶν καθορώμενος αἶαν  
 Μυσῶν<sup>13</sup> τ' ἀγγεμάχων καὶ ἀγαυῶν Ἰππημολγῶν  
 γλακτοφάγων Ἀβίων τε δικαιοτάτων ἀνθρώπων (N 4–6)

Now Zeus turned away his bright eyes, and looked afar, upon the land of the Thracian horsemen, and of the Mysians that fight in close combat, and of the lordly Hippemolgi who are cheese-eaters, and of the Abii, the most righteous of men.

Those words became enigmatic to the ancients themselves; this very passage was widely commented<sup>14</sup> by antique scholiasts, who identified the Abii either with the Scythians or the (equally Indo-Iranian) Sarmatians:

1. γλακτοφάγων Ἀβίων τε δικαιοτάτων ἀνθρώπων· †λακτινες† ἔθνος, οἱ γαλακτοπόται. Τινὲς τούτους Σαρμάτας φασιν. (Il. xiii.5) “dairy (?) people, who are milk-drinkers. Some also call them *Sarmatians*.”
2. Ἀβίων· πάντων Σκυθῶν ὑποκυψάντων Ἀλεξάνδρῳ μόνους Ἀβίους φασὶν οὐχ ὑπεῖξαι “The *Abii*: amongst all Scythians who have bowed to Alexander the Great, it is said that only the *Abii* didn’t surrender.”
3. οὐς δικαιοτάτους φησὶ διὰ τὸ ἀνεπίμικτον “(Homer) says they are the most righteous among men for their people is unmixed.”
4. Ἀβίων· τῶν νομάδων Σκυθῶν “*Abii*: the nomad Scythians.”
5. τινὲς δὲ τούτους Σαρμάτας φασὶν “some others call them *Sarmatians*”

Modern scholars<sup>15</sup> identify the Abii either with the legendary Hyperboreans, or with the Gabii mentioned by Aeschylus in a fragment of *Prometheus Unbound* (fr. 186). Aristarchus himself endeavoured without success to distinguish between epithet and ethnonym in Homer’s *Iliad* (N 5–6). The word ἰππημολγός could be understood as an epitheton meaning ‘mare-milkers’ (cf. Gr. ἵππος m.f. ‘horse, mare’), associated to Hom. ἀγαυός ‘noble’ and to metrically syncopated # γλακτο-φάγος ‘cheese-eaters’ (here standing for †# γαλακτο-φάγος). Even the word ἄβιος could be understood as an epitheton: ‘without (fixed) subsistence’, whence ‘nomad’ (Gr. βίος m. ‘life’ means also ‘means of life, resources, sustenance’).<sup>16</sup> The legendary Abii who have occasioned so much discussion may eventually be nothing else but

13. Those are the Mysians living on the shore of the Danuvius, not the Mysians from Asia.

14. Text of the scholia by Erbse (1974: 392–396).

15. For instance, Deforge (1986: 194–195), Janko (1992: 42–43), Reece (2001).

16. See for instance βίον πορίζειν τινί “to furnish s.o. the means of sustenance” (Aristoph. *Ve*. 706).

a ghost-word;<sup>17</sup> but the fact that a dairy nomad tribe living on milk and “without fixed subsistence” (that is to say without agriculture) is referred to in this passage is beyond doubt.

### 2.3.2 *Hesiodus*

Γλακτοφάγων ἐς γαῖαν ἀπίννας οἰκί' ἐχόντων (fr. 54)<sup>18</sup> [probably Scythians]

To the land of Cheese-Eaters, whose houses are chariots.

### 2.3.3 *Hippocrates*

In his famous treatise *Airs, Waters, Places*, 18, Hippocrates depicts the Scythians' milk-based diet: *Αὐτοὶ δ' ἐσθίουσι κρέα ἐφθὰ καὶ πίνουσι γάλα ἵππων. Καὶ ἵππάκην τρώγουσι· τοῦτο δ' ἐστὶ τυρὸς ἵππων.* “They themselves eat boiled meats and drink mares' milk.<sup>19</sup> They have a sweet-meat *hippake*,<sup>20</sup> which is a cheese from the milk of mares.”

### 2.3.4 *Herodotus: The Massagetae and the Scythians*

#### 2.3.4.1 *The Massagetae*

γαλακτοπόται δ' εἰσὶ (Hdt. i.216) “the Massagetae are milk-drinkers”.

#### 2.3.4.2 *The Scythians (Hdt. iv.2.1–2)*

(1) Τοὺς δὲ δούλους οἱ Σκύθαι πάντας τυφλοῦσι τοῦ γάλακτος εἵνεκεν τοῦ πίνουσι ποιεῦντες ὧδε. Ἐπεὰν φυσητήρας λάβωσι ὀστεῖνους αὐλοῖσι προσημφερεστάτους, τούτους ἐσθέντες ἐς τῶν θηλέων ἵππων τὰ ἄρθρα φυσῶσι τοῖσι στόμασι, ἄλλοι δὲ ἄλλων φυσώντων ἀμέλγουσι. Φασὶ δὲ τοῦδε εἵνεκα τοῦτο ποιέειν· τὰς φλέβας τε πίμπλασθαι φυσωμένας τῆς ἵππου καὶ τὸ οὔθαρ κατίεσθαι. (2) Ἐπεὰν δὲ ἀμέλξωσι τὸ γάλα, ἐσχέαντες ἐς ξύλινα ἀγγήια κοῖλα καὶ περιστίζαντες κατὰ τὰ ἀγγήια τοὺς τυφλοὺς δονέουσι τὸ γάλα, καὶ τὸ μὲν αὐτοῦ ἐπιστάμενον ἀπαρύσαντες ἠγεῦνται εἶναι τιμώτερον, τὸ δ' ὑπιστάμενον ἦσσαν τοῦ ἐτέρου.

17. In this case, we may amend the *textus traditus* as follows: Μυσῶν τ' ἀγχεμάχων καὶ ἀγαυῶν \*ἵππημολγῶν # γλακτοφάγων \*ἀβίων τε δικαιοσάτων \*τ' ἀνθρώπων (N 4–6) “and upon the land of the illustrious Mysians that fight in close combat, mare-milkers, cheese-eaters, without (fixed) subsistence, the most righteous of men.”

18. Rzach's edition (1902: 145).

19. Note OPr. *aswinan* [*dadan*] N. ‘mare's milk’.

20. This Greek word could be a calque of an Iranian word, \**aspa-kā* or the like.



“(1) Now the Scythians put out the eyes of all their slaves because of the milk which they drink; and they do as follows: they take blow-pipes of bone just like flutes, and these they insert into the vagina of the mare and blow with their mouths, and others milk while they blow: and they say that they do this because the veins of the mare are thus filled, being blown out, and so the udder is let down. (2) When they have drawn the milk they pour it into wooden vessels hollowed out, and they set the blind slaves in order about the vessels and agitate the milk. Then that which comes to the top they skim off, considering it the more valuable part, whereas they esteem that which settles down to be less good than the other. For this reason the Scythians put out the eyes of all whom they catch.”

### 2.3.5 *Pliny the Elder*

*Mīrum barbarās gentēs, quā lacte uīuant, ignōrāre aut spernere tot saeculīs cāseī dōtem, densantēs id alioquī in acōrem iucundum.* (HN xi.96.3),

“It is a remarkable circumstance, that the barbarous nations which subsist on milk have been for many ages ignorant of the merits of cheese, or else have totally disregarded it; and yet they understand how to thicken milk and form therefrom an acrid kind of liquid with a pleasant flavour.”

## 2.4 Concluding remarks

The Post-Anatolian innovation points to the creation of a “secondary” root  $*h_2melǵ-$  ‘to collect liquid (in a bucket), to milk’. A major part of the Indo-European languages (including Tocharian) used this specialized root to build new names for ‘milk’: Com. Germ.  $*mel(u)k-a^n$  N. ‘milk’, Com. Celt.  $*mliχtos$  M. ‘id.’, Toch. A *malke*, B *markwer* ‘id.’, and (maybe) Lat. *lac*, *lact-is* N. ‘id.’. The highly innovative Balkanic area, although using  $*h_2melǵ-$  as a verbal root (Alb. *mjell* ‘to milk’, Gr. ἀμέλω ‘id.’), shows a lexical renewal exemplified by Gr. γάλα N. ‘milk’ (< PIE  $*ǵ̑h_2-n̥-k-$  ‘white’) and Alb. *dhallë* / *dhalltë* F. ‘buttermilk’, which could reflect Proto-Alb.  $*dzalā-$  (< PIE  $*ǵ̑h_2-éh_2$  F. ‘whiteness’).<sup>21</sup> Such “modern” designations point to an innovative name for ‘milk’ as consumed by both infants and adults.

21. As already mentioned, there is evidence for a similar lexical renewal in Semitic, where the root ‘to milk’ is  $\sqrt{HLB-}$ , whereas several languages created a new name for ‘milk’ based on  $\sqrt{LBN-}$  ‘to be white’.

The Indo-Iranian data are particularly complex: we may admit that the Core IE root *\*h<sub>2</sub>melǵ-* ‘to collect liquid (in a bucket), to milk’ was lost, because of its homophonic confusion with the unrelated root PIE *\*merǵ-* (cf. Ved.  $\sqrt{MRJ}$ - ‘to wipe, brush’). Besides, the Indo-Iranian tribes seem to have been quite significantly living on milk, at the time when they were still nomads: the first mention ever of milk-drinking by lactose-tolerant adults appears to be in the old Indo-Iranian formula *\*sáumas iás gáuā* ‘soma-juice<sup>22</sup> mixed with (cow’s) milk’, reflected by Late Av. *\*haomō.yō gauua* ‘soma-juice mixed with milk’ (Yt 3.18 ff.).<sup>23</sup> References to milk in Early Vedic texts are ubiquitous<sup>24</sup> and the posterior Ayurvedic literature emphatically states that milk can be consumed by all healthy individuals.<sup>25</sup>

Common Indo-Iranian has an etymologically totally obscure generic word for ‘milk’: Ved. *kṣírā-* n. ‘milk’ (also Classical Skr.), and Com. Ir. *\*xšīra-* ‘id.’, reflected by MoPers. *šīr* ‘id.’. The Iranian languages also have a protean word for ‘milk’, not reflected in Indo-Aryan: OAv. *xšuuīpta-* ‘milk’, Pašto *šaudə* ‘id.’, Khot. *švīda* ‘id.’ from Com. Ir. *\*xšwīfta-*, of unknown origin (not from Com. Ir. *\*xšwīd-*, pace Mayrhofer *EWAia* I: 453).

### 3. Conclusion

Some of the findings in Section 2 are directly interpretable in terms of the genetic and archaeological findings on lactase persistence described in Section 1. First, we have shown that after the separation of the Anatolian branch, and before the breakup of Core Indo-European (dated to ca. 2800 BCE by Chang et al. 2015), a specialized root for ‘milker’ came into existence, out of which a specialized verb ‘to milk’ *\*h<sub>2</sub>melǵ-* was formed. We take the appearance of this term as signalling the new status of animal milking as a well-identified social activity in early Indo-European society. While Proto-Indo-European must have had a word for human milk – not recoverable due to the specificities of Hittite script, we have shown that

22. Whatever soma-juice may have been, it certainly referred to a strong intoxicating liquor – definitely not a beverage for suckling infants.

23. *Lectio* supported by de Vaan (2003: 370) for the *textus traditus*, which reads here †*haomaiō*.

24. If we may say so, the whole *Rig-Veda* is crawling with mentions of milk and sperm.

25. See for instance the gnomic stanza: *kṣīraṃ sarveṣāṃ dehināṃ cānuṣete kṣīraṃ pibanti ca na roga eti || kṣīrāt paraṃ nānyadhāsti vṛṣyaṃ kṣīrāt paraṃ nāsti ca jīvanīyam ||90||* (*Ka.Ka.* 7.90) “Milk is beneficial for healthy individuals; by drinking milk one does not get diseases (*roga-*); hence there is no better aphrodisiac (*vṛṣya-*) than milk; there is no better life-prolonger (*jīvanīyam*) than milk.” Note also: *pravaraṃ jīvanīyānāṃ kṣīraṃ uktaṃ rasāyanam ||218||* (*Cāraka Saṃhitā Sūtrasthāna* 27.218) “Milk is said to be a life-elixir per excellence.”

new Indo-European words for ‘milk’ were formed independently from the verb *\*h<sub>2</sub>melǵ-* ‘to milk’ in the Germanic, Celtic, Italic, Slavic and Tocharian branches: these terms must have designated animal milk for human consumption. This probably indicates a widespread social need for distinct specialized names for the two notions, animal milk and mother’s milk.

By themselves, these linguistic data are silent on whether adult speakers drank animal milk or whether the milk was used to make cheese, or both; though the greater prominence of terms for ‘milk’ compared to those for ‘cheese’ does suggest that milk was directly consumed by adults. Unequivocal linguistic evidence of milk consumption by Indo-European-speaking adults can be found at a later date, in prescribed ritual drinking of *soma* mixed with milk by Proto-Indo-Iranian adults: lactose-persistence in the Proto-Indo-Iranian population (slightly after 2000 BCE according to Chang et al. 2015) must therefore have reached very high levels. Adult milk drinking by Indo-Iranian peoples is further confirmed by descriptions by Roman and Greek authors. Despite being linguistically Indo-European, Roman and Greek authors considered adult milk-drinking a barbarian custom, perhaps because Roman and Greek populations included a large pre-Indo-European farmer component.

To conclude, we suggest that the ability to drink milk in adulthood played an important role first in boosting Proto-Indo-European demography. A larger population in turn required more milk: the need for more pasture lands is probably one strong motivation behind Indo-European territorial expansion. In confrontations with preexisting farming populations, increased population numbers allowed Indo-European groups to prevail militarily over small, or even not-so-small, farming communities which had until then been secure. As a result Indo-European speakers were able to establish themselves durably as a ruling elite over sedentary farming communities speaking non-Indo-European languages. Like horseback riding, teenage and adult milk consumption may also have amplified the military might of Indo-European raider groups by conferring higher bodily stature to Indo-European individuals with the lactase persistence phenotype (Okada 2004).

## Abbreviations

ADJ.	adjective	mid.	middle
Alb.	Albanian	MoAr.	Modern Arabic
AP	Anthology Palatine	MoGerm.	Modern German
AVed.	Atharva-Vedic	MoPers.	Modern Persian
Fin.	Finnish	MPers.	Middle Persian
GEN.	genitive	n.	neuter
Com. Celt.	Common Celtic	OArm.	Old Armenian
Com. Germ.	Common Germanic	OAv.	Old Avestan
Com. Ir.	Common Iranian	OCS	Old Church Slavonic
Com. It.	Common Italic	OE	Old English
Com. Slav.	Common Slavic	OFr.	Old French
Gallo-Rom.	Gallo-Romance	OIr.	Old Irish
Gaul.	Gaulish	OPr.	Old Prussian
Go.	Gothic	PL.	plural
Gr.	Greek	PIE	Proto-Indo-European
Hitt.	Hittite	Proto-Alb.	Proto-Albanian
Hom.	Homeric	Rom.	Romanian
IE	Indo-European	sg.	singular
Iir.	Indo-Iranian	Skr.	Sanskrit
It.	Italian	Toch.	Tocharian
Khot.	Khotanese	Ved.	Vedic Sanskrit
Late Av.	Late Avestan	Vulg.	Vulgar
Lith.	Lithuanian	YAv.	Younger Avestan
m.	masculine		

## References

- Allentoft, Morten E. *et al.* 2015. Population genomics of Bronze Age Eurasia. *Nature* 522: 167–172. doi:10.1038/nature14507
- Anthony, David. 2007. *The Horse, the Wheel and Language*. Princeton NJ: Princeton University Press.
- Bellwood, Peter. 1985. *Prehistory of the Indo-Malaysian Archipelago*. Sydney: Academic.
- Benveniste, Émile. 1935. *Origines de la formation des noms en indo-européen*. Paris: Maisonneuve.
- Burger, Joachim *et al.* 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proceedings of the National Academy of Sciences* 104(10): 3736–3741. doi:10.1073/pnas.0607187104
- Chang, Will, Cathcart, Chundra, Hall, David & Garrett, Andrew. 2015. Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91(1): 194–244. doi:10.1353/lan.2015.0005

- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb* [Leiden Indo-European Etymological Dictionary Series, Vol. 2]. Leiden: Brill.
- Deforge, Bernard. 1986. *Eschyle, poète cosmique*. Paris: Les Belles Lettres.
- Derksen, Rick. 2008. *Etymological Dictionary of the Slavic Inherited Lexicon* [Leiden Indo-European Etymological Dictionary Series 4], Alexander Lubotsky (ed.). Leiden: Brill.
- Dieu, Éric. 2016. *L'accentuation des noms en \*-ā (\*-eh<sub>2</sub>) en grec ancien et dans les langues indo-européennes*. Innsbruck: IBS 156.
- Erbse, Harmut. 1974. *Scholia Græca in Homeri Iliadem (scholia uetera). Recensuit Hartmut Erbse. Volumen tertium scholia ad libros K-Ξ continens*. Berolini: apud Walter de Gruyter et socios, MCMLXXIV.
- Evershed, Richard P. et al. 2008. Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* 455: 528–531. doi:10.1038/nature07180
- Garnier, Romain. 2014. Sur l'étymologie du latin *uirgō* 'vierge'. *Studia Etymologica Cracoviensia* 19: 59–70.
- Garnier, Romain. 2016. *La dérivation inverse en latin*. Innsbruck: IBS 157.
- Garnier, Romain. 2016b. Secondary roots in PIE. Handout, 15. *Fachtagung der Indogermanischen Gesellschaft*. Vienna, September 2016.
- GLK: *Grammatici Latini ex recensione Henrici Keil* 7: 1873–1923. Leipzig: Teubner.
- Griepentrog, Wolfgang. 1995. *Die Wurzelnomina der Germanischen und ihre Vorgeschichte*. Innsbruck: IBS 82.
- Haak, Wolfgang et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522: 207–211. doi:10.1038/nature14317
- Ivanov, Vyacheslav. V. 2003. On the origins of Tocharian terms for GRAIN. In *Language in Time and Space: A Festschrift for Werner Winter on the Occasion of his 80th Birthday*, Brigitte L. M. Bauer & Georges-Jean Pinault (eds), 189–210. Berlin: Mouton de Gruyter. doi:10.1515/9783110897722.189
- Janko, Richard. 1992. *The Iliad: A Commentary*, IV: 13–16. Cambridge: CUP.
- Kloekhorst, Alwin. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon* [Leiden Indo-European Etymological Dictionary Series 5], Alexander Lubotsky (ed.). Leiden: Brill.
- Kortlandt, Frederik. 2009. C.C. Uhlenbeck on Indo-European, Uralic and Caucasian. *Historische Sprachforschung* 122: 39–47.
- Kümmel, Martin. 2017. Even more traces of the accent in Armenian? The development of tenues after sonorants. In *Usque ad radices. Indo-European Studies in Honour of Birgit Anette Olsen* [Copenhagen Studies in Indo-European 8], Bjarne Simmelkjær et al. (eds), 439–452. Copenhagen: Museum Tusulanum.
- Leonardi, Michela, Gerbault, Pascale, Thomas, Mark G. & Burger, Joachim. 2012. The evolution of lactase persistence in Europe. A synthesis of archaeological and genetic evidence. *International Dairy Journal* 22: 88–97. doi:10.1016/j.idairyj.2011.10.010
- Leumann, Manu. 1977. *Lateinische Laut- und Formen-Lehre*. München: C.H. Beck'sche Verlagsbuchhandlung.
- LIV<sup>2</sup>: Rix, Helmut et al. 1998[2001]. *Lexikon der Indogermanischen Verben: Die Wurzeln und Ihre Primärstammbildungen*. Unter Leitung von H. Rix, bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp, Brigitte Schirmer. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Mallory, James Patrick. 1989. In *Search of the Indo-Europeans: Language, Archaeology and Myth*. London: Thames & Hudson.

- Martirosyan, Hrach K. 2010. *Etymological Dictionary of the Armenian Inherited Lexicon* [Leiden Indo-European Etymological Dictionary Series 10], Alexander Lubotsky (ed.). Leiden: Brill.
- Matasović, Ranko. 2009. *Etymological Dictionary of Proto-Celtic* [Leiden Indo-European Etymological Dictionary Series 9], Alexander Lubotsky (ed.). Leiden: Brill.
- Mayrhofer, Manfred. 1986–2001. *EWAia (I to III): Etymologisches Wörterbuch des Altindoarischen*, III Bände. Heidelberg: Universitätsverlag Winter.
- Nikolaev, Alexander. 2014. Greek εἰαμενή, Vedic yávasa-. *Münchener Studien zur Sprachwissenschaft* 68(1): 127–139.
- NIL: Wodtko, Dagmar S., Irslinger, Britta & Schneider, Carolin. 2008. *Nomina im indogermanischen Lexikon*. Heidelberg: Winter.
- Okada, Tomoo. 2004. Effect of cow milk consumption on longitudinal height gain in children. *American Journal of Clinical Nutrition* 80(4): 1088–1089.
- Owen, James. 2010. Stone Age milk use began 2,000 years earlier. *National Geographic News*, October 28.
- Pinault, Georges-Jean. 2007. A star is born: a ‘new’ PIE \*-ter- suffix. In *Verba Docenti, Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues, and Friends*, Alan J. Nussbaum (ed.), 271–279. Ann Arbor MI: Beech Stave Press.
- Reece, Steve. 2001. The ἄθιοι and the γᾶθιοι: An Aeschylean solution to a Homeric problem. *AJPh*, 122(4): 465–470.
- Renfrew, Colin. 1987. *Archaeology and Language*. London: Jonathan Cape.
- Rzach, Aloisius. 1902. *Hesiodi carmina. Recensuit Aloisius Rzach*. Lipsiae: In aedibus B. G. Teubneri.
- Sagart, Laurent, Hsu, Tze-fu, Tsai, Yuan-Ching & Hsing, Yue-İe. In press. Austronesian and Chinese words for the millets. *Language Dynamics and Change*.
- Salque, Melanie *et al.* 2013. Earliest evidence for cheese making in the sixth millennium BC in northern Europe. *Nature* 493: 522–525. doi:10.1038/nature11698
- Szemerényi, Oswald. 1958. Greek γάλα and the Indo-European term for ‘milk’ with an excursus on Avesta xšvid-. *Zeitschrift für Vergleichende Sprachforschung* 75: 170–190.
- Szemerényi, Oswald. 1991. *Scripta minora: selected essays in Indo-European, Greek, and Latin*, Vol. 3. Innsbruck: IBS 53.
- Uhlenbeck, Christian Cornelius. 1895. Waar werd de Indogermaansche stamtaal gesproken? *Tijdschrift voor Nederlandsche Taal- en Letterkunde* 14: 69–74.
- Uhlenbeck, Christian Cornelius. 1897. De voorgeschiedenis der Indogermaansche volken. *Taal en Letteren* 7: 1–25.
- de Vaan, Michiel. 2003. *The Avestan Vowels*. Amsterdam: Rodopi.
- Weiss, Michael. 2011[2009]. *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor MI: Beech Stave Press.