Studia Linguistica Universitatis Iagellonicae Cracoviensis 133 (2016): 75–96
doi:10.4467/20834624SL.16.006.5152

FILIP DE DECKER
Universiteit Gent
filipdedecker9@gmail.com

AN ETYMOLOGICAL CASE STUDY ON THE <PG> AND <PG?> VOCABULARY IN ROBERT BEEKES'S NEW ETYMOLOGICAL DICTIONARY OF GREEK: M

Keywords: substrate, inherited lexicon, sound laws, Indo-European, Greek, etymological dictionaries

Abstract

This article presents an etymological case study on Pre-Greek (PG): it analyzes about 20 words starting with the letter M that have been catalogued as <PG> or <PG?> in the new *Etymological dictionary of Greek* (EDG), but for which alternative explanations are equally possible or more likely (discussing all instances would be tantamount to rewriting the dictionary). The article briefly discusses the EDG (for an in-depth appraisal the reader is referred to part one of the article) and then analyzes the individual words. This analysis is performed by giving an overview of the most important earlier suggestions and contrasting it with the arguments used to catalogue the word as PG. In the process, several issues of Indo-European phonology (such as the phoneme inventory and sound laws) will be discussed.

1. Observations on the Leiden etymological dictionaries and the EDG.

In part one of this article (in which we discussed the pre-Greek lemmata of the letter N in the EDG), we pointed out that it cannot be denied that Greek borrowed words from many non-Indo-European languages (there is a consensus that more than half of the words in the Greek lexicon are of non-Indo-European origin). It is therefore logic that the EDG often argues for non-IE origin of words, but the problems with the EDG are the following: first, it assumes that Greek borrowed most of its words

from one and the same language, without taking into account the time depth problem (i.e. when was the word attested in, in which author can the word(s) be found) and the unlikeliness of all words coming from one language, given the multitude of languages that were spoken in Antiquity (according to the Ancient sources); second, when a borrowing and an inherited etymology are equally possible, the EDG assumed the word was borrowed without explicitly stating why, but we think that it might be better to prefer an inherited etymology when the evidence allows it; third, it uses the concept "Pre-Greek" to explain away words that have a possible Indo-European etymology that includes elements that are in contradiction with the Leiden school [see Verhasselt (2009a, 2009b, 2011); Meissner (2014); De Decker (2015)].

2. Individual etymologies[1]

1. *máka:r* 'happy, blessed' (Beekes 2010: 893). Older etymological dictionaries linked the word with *makrós* 'long' and assumed a semantic evolution from 'long, great' into 'happy' [Curtius (1879: 161); Prellwitz (1905: 278–279)]. This evolution cannot be ruled out, but is not evident (Boisacq 1938: 601–602), especially since the root *mak* means 'meagre, long' and from 'meagre' to 'happy' is difficult. Brugmann argued that *máka:r* was an original neuter noun 'blessedness', which was then reinterpreted as adjective and received masculine and feminine forms [Brugmann (1905: 434), supported by Benveniste (1935: 18); Boisacq (1938: 601–602); Schwyzer (1939: 519); Chantraine (1968: 659)]. Frisk (1970: 162–163) stated that this explanation was formally sound, but not supported by the texts. Beekes only stated that the texts did not confirm Brugmann's hypothesis and argued that the isolated formation and the variation between long and short *a* in the second syllable pointed at Pre-Greek origin. This argumentation is not convincing. First, as Frisk stated (and Beekes left out), the distinction between long and short *a* is formally expected: if we start from a stem in a short *a*, the case forms outside the masculine singular have a short vowel *a* whereas the masculine singular has a long *a* because of Szemerényi's Law:[2] **makars* with a nominative singular ending *s* would regularly become *máka:r*. The case forms attested in Homer follow this schema. Second, there might be another example of a neuter noun in *ar* that was reinterpreted as an adjective or noun: Pedersen (1893: 244), explained the noun *dámar* 'wife' as an original neutre *r*/*n* noun which was reinterpreted as a feminine, but there are no examples in our texts of a neuter noun *dámar*. If one accepts the analysis for *mártu:s* 'witness' (cf. infra), this would be another example of an originally neuter noun that

[1] We decided to transcribe the Greek. In doing so, we used the accents ´ (acutus), ` (gravis) and ˆ (circumflexus). We use the sign : to indicate vowel length. Vowel length is not indicated when a vowel is written with a ˆ, because vowels with a circumflex are always long. A sign H refers to any laryngeal, a C to any consonant, a P to any plosive, an R to any resonant and a V to any vowel.

[2] This law states that at word end a sequence VRs (with V being any vowel and R being any resonant) became V:R (see Szemerényi 1996: 116; Weiss 2009: 47; Kümmel forthcoming). This was already noted in the 19th century, as Szemerényi stated himself.

became masculine in Greek. Third, it is true that there are no cognates in the other Indo-European languages, but that does not mean that the word was borrowed from "Pre-Greek". In any case, there are other neuter nouns ending in *ar*, such as *néktar* 'drink of the gods' (although this word was debated as well, cf. part one). Brugmann's explanation has the advantage that it explains the adjective and its inflection as an inner-Greek development, but the problem is that the original noun is not attested (anymore). This should not be a problem in itself, because the adjective *aske:thé:s* 'unharmed' is derived from a noun **skêthos* 'harm' which is also unattested (albeit it is attested in other Indo-European languages).

2. *makednós* 'tall, slim' (Beekes 2010: 894); *Makedó:n* (nominative), *Makedónos* (genitive) 'Macedonian'. As this word means 'tall', a connection with *makrós* seems logical [Prellwitz (1905: 279); Frisk (1970: 163)]. Fick (1901a: 242) interpreted the name as 'people living on the high planes'. Building on Fick's explanation, Frisk explained the adjective *makednós* as having the zero grade *dn* while the personal name has the full grade *don-*. Already Krahe (1928: 159) doubted the Greek origin of *makédo:n* and Chantraine (1968: 660) and Beekes (2010: 894) followed him in this. Chantraine suspected that Krahe might have been right in doubting the Greek origin of the word *Makedó:n*. Beekes admitted that the meaning might point at a link with *makrós* but argued for PG on two grounds: firstly, because an analysis of *make-dn-os* would have been impossible for Indo-European and secondly, because there was a variant *Makéte:s* 'Macedonian'. The existence of a form with a *d* as in *Makedo:n* and with a *t* in *Makéte:s* was in Beekes's opinion indicative for PG origin, as a variation *t*/*d* was a feature of PG. The probative value of the form in *-éte:s* is in our opinion limited, as we might be dealing with a suffix *éte:s* in *mak-éte:s* as in *oik-éte:s* 'living in one's house, house slave' (this example was even adduced by Beekes himself). We are also unsure why an ablaut pattern *don*/*dn* would be impossible. The use of the full grade in the noun declension and the zero grade in the adjectival derivation is paralleled in the word for 'father', where we have the nominative *paté:r* and the genitive *patéros* besides *patrós* and an adjective *pátrios* 'fatherly'.

3. *malthakós* 'weak, tender, soft, mild' (Beekes 2010: 897–898).
4. *málthe:* 'mix of wax and pitch' (Beekes 2010: 898). Hesykhios has a gloss *málthe: trupheré: málthe:* means 'delicate'.

The adjective *malthakós* is traditionally linked with the Germanic words for 'mild' (*mild* in English, German and Dutch; *mildeis* in Gothic and *mildr* in Old Norse) and with Sanskrit *mardhati* 'to neglect', and can be reconstructed as **meld*^{h} [Kluge et al. (1957: 479); Chantraine (1968: 662); Van Veen, Van der Sijs (1997: 563); Zehnder (2001b); DWDS *s. u. mild*; Mayrhofer (1996: 328–329) was more skeptical]. As *malthakós* and *malakós* 'soft' show similar suffixes and have the same meaning, it is likely that they influenced each other [Chantraine (1968: 662); Frisk (1970: 167)]. Solmsen (1909: 55–56; quoted in Frisk 1970: 167) and Chantraine (1968: 662) argued that the feminine noun *málthe:* was in origin a feminine adjective form from **malthós* and reconstructed *malthakós* as **mlthn̥kos*. Beekes argued that *malakós* 'soft' and *malthakós*

did not influence one another, but did not state why he thought so. He assumed that *malthakós* was PG, because it could not be derived from a zero grade of **meld*^{h} as this would have given ***blathakós*. In addition, there was no Indo-European suffix **ṇko* (as suggested by Solmsen), and as Beekes (2010: 898), who discussed *the well-known PG suffix ako* and assumed that most words in *ako* were of substrate origin, and posited that *malthakós* was of PG origin as well. He denied the link between *málthe:* and *malthakós*, but doubted the meaning 'delicate' that was given by Hesykhios and considered *málthe:* to be PG, because of its technical meaning. We, on the other hand, see no reason to doubt the link between *málthe:* and *malthakós*. As wax is soft and mild, this poses no serious semantic problems. It is true that the lexicon by Hesykhios is not always trustworthy, but in order to decide when it can(not) be trusted, an in-depth analysis is needed and one cannot just doubt words because they do not fit in into a certain theory. One could explain *malthakós* as a derivation from *málthe:* influenced by *malakós*. As the *-akos* in *malakós* was inherited, Greek had words ending in *akos* that were not PG and consequently, not every word with this suffix can be catalogued as PG. As the meanings of both words were very close, mutual influence between them cannot be denied.

5. *mallós* 'flock of wool' (Beekes 2010: 899). Fick and Prellwitz connected this word with Lithuanian *mìlas* 'coarse homespun wool', reconstructed **mal-yos* and linked it with *malakós* [Fick (1872: 176); Prellwitz (1899: 285); Pokorny (1959: 721)]. Semantically, this is not convincing.[3] Greppin rejected the link made by Fick, compared the Greek word to Armenian *mal* 'wether, castrated ram' and reconstructed **malyos* (as Fick had done). He argued that the basic meaning was not 'castrated sheep', but that it received that meaning through the resemblance with the verb *malem* 'to crush, castrate' (Greppin 1981: 72). Greppin's explanation was doubted by Hamp (1982) and Clackson (1994: 232) because of the semantics.[4] Clackson explained the Armenian word as a borrowing from Arabic *māl* 'possession' with a semantic evolution from 'possession' into 'sheep' in Armenian. The semantics of Greppin's explanation are not problematic,[5] and certainly less problematic than Clackson's suggestion (also because it offers an etymology that does not involve a borrowing). If the initial meaning was 'sheep', it is possible that Greek narrowed the meaning into 'wool'. Beekes argued for PG because a reconstruction **mh$_2$l* would be improbable and because the cluster *ll* could have been a geminate from PG. If Greppin is right in his equation (and we see no reason to doubt it), this would be an Helleno-Armenian isogloss.[6] As such, a reconstruction (Proto-)Helleno-Armenian **malyos* would be possible and there is thus no need for a PG etymology. Assuming an Helleno-Armenian etymon would

[3] As was already noted by Boisacq (1938: 606), Greppin (1981: 70). Frisk (1970: 168) was less critical, while Chantraine (1968: 663) only stated that the etymology was unknown.

[4] This word was not discussed in Martirosyan (2008).

[5] As Beekes (2010: 899) pointed out as well.

[6] For evidence in favour of a close relationship between Greek and Armenian, see Solta (1960) and evidence against it can be found in Clackson (1994), but we leave out a detailed discussion on the issue of the relationships of Greek and the other Indo-European languages.

also solve the *a* problem, because at that stage the Helleno-Armenian proto-language could have had an *a* phoneme already (even if one started from the assumption that PIE did not have it).

6. *mártuːs* 'witness' (Beekes 2010: 908–909).
7. *mérimna* 'care, concern' (Beekes 2010: 932).

Since Fick, *mártuːs* has been linked with Sanskrit *smárati* 'remember' from the root **smer* 'remember' (Fick 1890: 338; Boisacq 1938: 612; Hofmann 1950: 191; Frisk 1970: 170).[7] Frisk (1970: 170) suggested the following evolution for *mártuːs*: first, the root was put in the zero grade and extended by an abstract suffix *tu* and would have meant 'remembrance, testimony'. In a second stage, an adjectival *ro* was added and in a third stage, *martus* and *márturos* became crossed, leading to a new nominative **márturs* which became **mártuːr* and eventually dissimilated into *mártuːs* (which explains the apparent exception to Szemerényi's Law) (Schwyzer 1939: 260, with a list of similar dissimilations; Frisk 1970: 170; a list of dissimilations can be found in Grammont 1948).[8] In this scenario, *mártuːs* would originally have meant 'testimony, remembrance' and only later 'he who remembered, witness'. A similar evolution from an abstract noun into a concrete one or adjective can be seen in *mákaːr* (cf. supra). Chantraine (1968: 669) rejected this hypothesis, because an action noun *martu-* remained hypothetical. Beekes assumed a PG origin, because a form **smrtu* in the zero grade[9] should have given ***bratu* and because the suffixes *tu* and *r* were of non-Indo-European origin. We doubt this, because the evolution form *mr* into *bra* is only certain for word initial position, but in **smr̥tu* the cluster *mr* stands in word internal position. We also disagree with the assessment that the suffixes *tu* and *r* were of non-Indo-European origin. The root **smer* can also be seen in *mérimna* (Chantraine 1968: 687; Mayrhofer 1996: 781). Beekes doubted the Indo-European heritage of this word, because he and Furnée (1972: 246) considered the suffix *mna* to be PG. We believe that there are neither semantic nor formal reasons excluding an Indo-European etymology for both words.

8. *máthuia* 'jaw'.
9. *masáomai* 'I chew' (Beekes 2010: 909).
10. *mástaks* 'mouth' (Beekes 2010: 911).
11. *mástiks,* genitive *mástigos* 'whip' (Beekes 2010: 911–912).
12. *móthos* 'battle din' (Beekes 2010: 961).
13. *Moûsa* 'Muse' (Beekes 2010: 972–973).

We believe that the first three words are related and that the last three are related as well. We discuss them together, because their etymologies pose the same problems,

---

[7] Prellwitz (1905: 282–283) mentioned both *mar* and *smer* as possible etymologies. Surprisingly enough, Greek *mártuːs* was not mentioned in Mayrhofer (1996: 781, contrary to *mérmina*).

[8] They did not address the absence of Szemerényi's Law.

[9] Beekes did not distinguish between vocalic and consonantic resonants. There is a case to make for such an approach, especially since different languages seem to follow different vocalization rules, but after some hesitations, we decided to make the distinction after all.

namely the issue of the aspirates and the treatment of a laryngeal preceded by a consonant and followed by a yod.

The words *máthuia*, *masáomai* and *mástaks* are related to Latin *mandere* 'chew' and can be linked to either Sanskrit *MATH* 'rob, take quickly'[10] or *MANTH*^{i} 'move heavily, move quickly' (Hofmann 1950: 191; Walde, Hofmann 1954: 26; Zehnder 2001d: 442; Meiser 2005).[11] The former continues PIE **mat^{h}*, while the latter continues **me/ont^{h}*:[12] The Greek words could continue a zero grade from the root **me/ont^{h}* or the full grade from **mat^{h}*, but Latin *mandere* cannot be reconstructed from the zero grade of a root with **e/o*. *Mástiks* and *móthos* can be linked to Sanskrit *mánthati* 'agitate', OCS *mętetъ* 'causes confusion' and ON *mǫndull* 'Drehholz'.[13] In case of *mástiks*, the word is built on the zero-grade (with Greek *a* being the reflex of a sonantic *n*) and the meaning would be that a whip is a tool to drive and agitate animals. The word *móthos* is a bit more problematic: it is either a formation on the zero grade with Aeolic treatment of the vocalic *n* (which would then be an Aeolism of the epic language, *móthos* first being attested in Homer) (Kuiper 1934: 104), or it is built on a nasalless form of the root **me/ont^{h}* which is attested in Indic as well.[14] We believe that *Moûsa* can be linked as well, but will discuss the word at the end. Because he ruled out that PIE **tH* became *th* in Greek and because he did not accept voiceless aspirates for PIE, Beekes (2010: 909) rejected the connection between the Greek words and the other cognates, and considered the Greek words to be PG. He assumed that the suffix *ig* in *mástiks* was an additional indication for PG origin. Frisk and Chantraine also rejected the etymologies, because they thought that the Indo-European **t^{h}* was rendered by *t* in Greek (Chantraine 1968: 669, 708; Frisk 1970: 248–249; see especially Frisk 1936). We agree with Beekes that laryngeals did not aspirate in Greek (Beekes 1969: 179–181, 2010: 909; Elbourne 2000),[15] but – contrary to Beekes – believe that PIE did in fact have a fourth category of plosives, namely the voiceless aspirates (Rasmussen 1987, 1989; Elbourne 1998, 2000, 2001, 2011, 2012; De Decker 2011, forthcoming a, forthcoming b). Their existence is no longer

[10] The Indic roots are quoted in capital letters, because that is the way they are printed in Mayrhofer's etymological dictionary.

[11] For the difference between Sanskrit *MATH*^{i} and *MANTH*^{i} see Narten (1960); Hackstein (1995: 29–30), discussing the Tocharian evidence; Mayrhofer (1996: 311–312) who pointed out that both roots were confused only in later texts and not in the RigVeda; Zehnder (2001c, 2001d). Fick (1890: 283) only mentioned the root "quirlen" and not "kauen".

[12] We explain later on why we reconstruct the forms with **t^{h}* and not **th$_{2}$*.

[13] For the listing of the cognates, see Fick (1890: 283, 'without the Greek words); Prellwitz (1905: 297), Boisacq (1938: 642–643), Pokorny (1959: 732–733); Mayrhofer (1996: 311–312). Latin *mamphur* 'Stück aus einer Drehbank' (only attested in Paulus ex Festo) and *mentula* 'dick, penis' have been linked as well, but they pose some problems and we will leave them out of the discussion.

[14] According to Whitney (1885: 117), the Atharva Veda has a form *máthati* 'he agitates', but it is possible that this nasalless form is the result of inner-Indic evolutions (see above).

[15] We discussed all the examples in De Decker (2011 and forthcoming a), and showed that none of the examples in favour quoted in Peters (1993a, 1993b); Meiser (2005) and Nikolaev (2010: 66–67) was absolutely convincing. Beekes (2010: 909) also stated that there is not enough evidence to assume aspiratory force of laryngeals in Greek.

generally accepted, after de Saussure had shown that certain Indic voiceless aspirates could be explained by the combination of a plain plosive and a laryngeal.[16] Nevertheless, for a (relatively small) number of words their presence is needed.[17] We also believe that the Greek evidence excludes a laryngeal. Latin *mandere* and Sanskrit *MATH*^{i} could theoretically continue both PIE **math*$_2$ as **mat*^{h}, while Sanskrit *MANTH*^{i}, the Germanic and Slavic cognates could continue both PIE **me/onth*$_2$ as **mo/ent*^{h},[18] but this is not the case for the Greek words. If we start from the forms with a laryngeal, we can theoretically explain the aspiration in *máthuia* and *móthos*,[19] but we cannot arrive at *mástaks*, *mástiks, masáomai* or *Moûsa*. If one starts from **math*$_2$, the forms *mástaks*, *mástiks* and *masáomai* cannot be explained, because the *transponat* **math*$_2$*taks* would have given Greek ***mátaks* and **math*$_2$*tiks* would have yielded ***matatiks*. The form *masáomai* is also difficult to explain starting from a root **math*$_2$*-y-* because that would have given **matai-*. The same applies to *Moûsa*: in laryngealistic terms, this would be **month*$_2$*-yh*$_2$, but that would have given ***montaya*. There is a (supposed) sound law that states that a laryngeal disappeared between a consonant and a yod in word internal position (the so-called Lex Pinault or Pinault's Law).[20] If this rule were correct, *masáomai* and *Moûsa* would be regular outcomes from **math*$_2$*-ye/o* and **month*$_2$*-yh*$_2$ respectively, but there are some doubts about the validity of this sound law for Greek (Lindeman 2004: 126–129; Piwowarczyk 2008, forthcoming; Verhasselt forthcoming, §3 treats the Greek material). First of all, there are counterexamples such as *aróo*: 'I plough' from **h*$_2$*erh*$_3$*-ye/o* and *(w)eméo* 'I vomit' from **wemh*$_1$*-ye/o*, forms which Pinault explained as *thematische*

[16] In 1892, De Saussure, quoted in Bally, Gautier (1922: 603), argued in a short article in *BSL* that *certain cas* of the Indic voiceless aspirates went back to a combination of a plain voiceless plosive and what we would now call a laryngeal. In his *Mémoire sur le système primitive des voyelles* (dating from 1879) he had already suggested that the *th* in e.g. *grathnati* and *granthitas* was possibly the reflex of the *i* elsewhere in the verbal flection (Bally, Gautier 1922: 228). See Mayrhofer (1981a) for a detailed analysis of de Saussure's reconstructions.
It is important to note that de Saussure never said that all cases of Sanskrit *th* could be explained this way. The summary in BSL mentions *certains cas*, but since we only have a summary of what he actually said, we will never know how he actually envisaged the Indo-European consonant system. De Saussure's explanation was expanded by Pedersen (1893: 269–273, 1926: 48, 63–64); Kuryłowicz (1927: 202–204, 1928: 55–56, 1935: 46–52). Cuny (1912) showed that laryngeals could also aspirate voiced plosives in Indo-Iranian. For a detailed treatment of the laryngeal effects in Indo-Iranian, see Mayrhofer (1981b, 2005).

[17] Of the grammars on Indo-European, only Szemerényi (1996) accepted the existence of voiceless aspirates; Fortson (2004) considered them to be secondary and Clackson (2007) and Meier-Brügger (2010) stated that there were too few instances to reconstruct a separate category. For a reconstruction of phonemic voiceless aspirates, see Elbourne (1998, 2000, 2001, 2011, 2012) and Rasmussen (1987, 1989).

[18] For the laryngealistic reconstructions, see Mayrhofer (1996: 298–299, 311–312), Zehnder (2001c, 2001d). The form **math*$_2$ was suggested to include the Greek personal name *Proma:theús* but the long *a* in that name might be a case of secondary ablaut a/ā with the Greek *math* from *manthán*o*:* 'I learn'.

[19] Pedersen (1926: 52–54) already alluded to the fact that the Greek aspirate might be due to a laryngeal.

[20] This had first been noticed by Wackernagel (1896: 81) for Indic. For PIE, see Pinault (1982), Ringe (2006: 15), Byrd (2015: 208–240) (admitting that there are still unexplained counterexamples).

Umbildungen of originally athematic verbs based on the aorist forms *é:rosa* 'I ploughed' and *é:mesa* 'I vomited'. This would presuppose that all instances were analogically levelled out, which cannot be proved nor disproved. Secondly, while there are several good examples that seem to confirm this sound law for Greek, they can be explained differently (Piwowarczyk forthcoming). The first example is the noun *aosse:té:r* 'helper' from **sm̥-sok^{w}h$_2$y-* (literally 'together-follower') (Pinault 1982: 271–272).[21] This word is related to Latin *socius* 'ally' and Sanskrit *sákhā-* and Avestan *haxā* 'friend'. The indications for the laryngeal come from Indo-Iranian, namely the aspirate[22] and the absence of Brugmann's Law.[23] If the reconstruction as **sok^{w}h$_2$-i-* is correct, this would be an important example for the Law. Piwowarczyk, referring to Harðarsson, explained this as a secondary thematicization or a backformation on the aorist (Piwowarczyk forthcoming, referring to Harðarsson 1998: 328). In addition, it is possible that the laryngeal suffix **h$_2$-o* was only added in Indo-Iranian. In Latin and Greek, sequence **(sm̥)sok^{w}y-* without laryngeal would have given *áoss-* and *socius* as well and if *aosse:té:r* is related to Greek *hépomai* 'I follow' (from **sek^{w}-o-mai*), the question remains why *hépomai* has no laryngeal while would have had a laryngeal. The form *aosse:té:r* can be explained as a thematicization of **sok^{w}-y* as is the case for Latin *socius*.[24] A second example is the comparative *meídzo:n* 'bigger' from *mégas* (**meǵh$_2$s*). The expected comparative form would be **meǵh$_2$-yos-* and this would normally have given ***megaío:n*. The loss of laryngeal is not necessarily a result of the rule. As the positive was *mégas* and the superlative *mégistos* 'biggest', it is possible that the stem *meg* was reintroduced to have a comparative and superlative **megyo:n – mégistos* besides **kretyo:n – krátistos* 'better, best' and **elakh-yo:n -elákhistos* 'fewer, fewest'.[25] A third example is the verb *teíro:* 'I annoy' (Pinault 1982: 270). This is generally reconstructed as **terh$_1$-yoh$_2$* and would confirm the rule, but Greek *térnon* 'thorn' shows that the root also existed without a laryngeal and the connection with English *thorn* indicates that the laryngealless form might have already existed in PIE.[26] A fourth example is the verb *eíro:* 'I speak, declare' from **werh$_1$-ie/o* (Pinault 1982: 270). This present is rare and might well be a later creation based on the future *eréo:* 'I will say' (Chantraine 1948: 267, 1968: 325–326; Frisk 1960: 470; Kümmel 2001: 689–690; Piwowarczyk forthcoming). As such, we believe

---

[21] This was already noted by Peters (1980: 80–81).

[22] As we stated above, an Indo-Iranian voiceless aspirate can – in most cases – be explained as the result of a plain plosive and a laryngeal.

[23] This law states that an Indo-European **o* becomes *ā* in Indo-Iranian in an open syllable. That this lengthening did not happen in this word, means that the verb did not end in **k^{w}* followed by a vowel but in **k^{w}* and a laryngeal (as the laryngeal counts as a consonant). There are nevertheless several examples to this sound law and there are several publications on the problem, but we cannot address the issue here. It was first stated in Brugmann (1876: 380, note 9). The most in-depth analysis is Volkart (1994) (but the literature on the topic is enormous and the issue cannot be addressed here).

[24] As is argued by Verhasselt (forthcoming, example 32 in his article).

[25] Piwowarczyk (forthcoming) also assumed that the stem *meg* was used to form the comparative and superlative.

[26] See Verhasselt (forthcoming) under his example 30.

that Pinault's Law is not an Indo-European sound law,[27] and consequently, a reconstruction with a laryngeal cannot account for Greek forms *masáomai* (and *Moûsa*). As the forms cannot be reconstructed with a laryngeal, the Greek aspirates need to be accounted for in a different way. We believe that a reconstruction **me/ot^h* and *me/ont^h* (with an Indo-European voiceless aspirate) can solve the problem. If we accept that laryngeals had aspiratory force in Indo-Iranian but not in Greek, and that Greek and Indo-Iranian also preserved the inherited voiceless aspirates, the difference in consonantism between Greek *platús* 'flat' and Indic *pṛthus* 'flat' from **plth$_2$us* is explained, as is the difference in consonantism between Greek *máthuia* and *platús*. If one accepts the aspiratory force of laryngeals in Greek, *platús* is a difficult counterexample. If one does not accept aspiratory force and denies the existence of voiceless aspirates, the forms *móthos* and *máthuia* are not easily explained.[28] The form *oîstha* 'you know' seems to be a strong example for aspiratory force of laryngeals in Greek, if one accepts the reconstruction **th$_2$e* for the 2nd person singular ending (but there is nothing that argues against an ending **t^{h}a*). It is possible, however, that the imperative *ísthi* 'know' spread its aspiration to *oîstha* (Frisk 1936: 41–43; Ruijgh 1978: 302). A similar example for such an 'aspiration extension' can be found in the 3rd person singular imperative *anó:khtho:* 'let him order' and the 2nd plural *áno:khthe* 'you (pl.) order!' which have their cluster *khth* from the 2nd singular imperative *áno:khthi* 'order!' from the verb *áno:ga* 'I order'.

The last form that needs to be explained is *Moûsa*. Several suggestions have been made for this word. Brugmann interpreted the word as a compound of a root **men* 'think' and a suffix **tya*: the Muse would then be 'the thinking one, the inspiring one' (Brugmann 1894: 253–256, building on a suggestion by Theodor Benfey). Wackernagel (1895) argued that the suffix *tya* was not attested and suggested to link the word *moûsa* with Latin *mons* 'mountain', namely *mont-ya*: the Muse would then be 'goddess of the mountains'. The last suggestion was that by Ehrlich (1907), who argued that the Muse was the goddess that agitated and inspired knowledge and reconstructed **mont^{h}ya*. Wackernagel's ingenious suggestion has the problem that the root *mont* is not attested in Greek (Chantraine 1968: 716). We therefore prefer to link *Moûsa* to **month* rather than to **mont*. *Moûsa* would then be another word linked to the root **me/ont^h* (a laryngealistic reconstruction **month$_2$yh$_2$* for *Moûsa* is problematic, because Pinault's Rule did not apply in Greek). Beekes (2010: 972–973) argued that the word could be reconstructed as **monthya* and be linked with *manthán:o* 'I learn' or was PG because it did not have to be of Indo-European origin. We believe that there is no need to doubt the inherited nature of *Moûsa*

[27] Lindeman (2004: 126–129) and Piwowarczyk (2008: 37, forthcoming) pointed out that the rule only applied in younger languages and in Indo-Iranian, a language where the anaptyctic vowel between laryngeal and consonant was an *i*. They wondered if this could not have triggered the deletion. See most recently Verhasselt (forthcoming).

[28] Zubaty (1892) and Elbourne (2012) argued that PIE **t^h* lost its aspiration when it was preceded by a resonant or *s*. Paul Elbourne (2012) informed us that he therefore denied the link between *móthos* and *mánthati* but this is not necessary. If there was already a nasalless form in PIE, *móthos* might have come from that form. This would not contradict his sound law (but we think that the examples of this sound law can be explained differently).

(or any of the other words) and accepting voiceless aspirates allows us to link these words with other Indo-European cognates and there is no need to assume that the Greek words were PG.

14. *mágeiros* 'cook'.
15. *mákhaira* 'large knife', later also 'dagger' (Beekes 2010: 915).
16. *mákhomai* 'I fight' (Beekes 2010: 916).

It is not certain whether these three words are related, but we discuss them together as some dictionaries have linked them.

Three suggestions have been made for *mákhaira*. The first stated that it was a derivation from *mákhomai* 'fight', the second considered it a borrowing from Semitic and the third linked it with *mágeiros* (Chantraine 1968: 673; Frisk 1970: 187, without further observations; Beekes 2010: 915). Beekes accepted the connection between *mágeiros* and *mákhaira*, and interpreted *mákhaira* as PG, because it had a voiceless aspirate and *mágeiros* a voiced stop. In addition, the suffix *eiros* in *mágeiros* pointed at PG as well. *Mákhaira* has been interpreted as a Semitic loanword from *mᵉkērā* 'sword',[29] but against a Semitic borrowing speaks the fact that the word that would have been the basis for the borrowing was uncommon in Hebrew and Phoenician (Lewy 1895: 177–178).[30] This makes it more likely that the word was borrowed from Greek into Semitic.[31] An additional problem is the meaning of the Semitic word: this means 'sword', but *mákhaira* is attested in Homer with the meaning 'knife, dagger used in a sacrifice' but is not used to refer to a sword (Seiler, Capelle 1889: 371; O'Sullivan 1993a).[32] This brings us to the third suggestion, namely the link between *mákhaira* and *mákhomai*. At first, the link between *mákhaira* and *mákhomai* seems self-evident, as *mákhaira* means 'knife, dagger' and *mákhomai* 'fight'. In that case, the former would be a derivation with suffix *ya* on an *r* extension of *mákhomai* (i.e. **makh – ṛ – ya*) (Boisacq 1938: 616; Frisk 1970: 188; Peters 1980: 181).[33] Some have doubted this derivation, because *mákhaira* is <u>not</u> used to refer to a sword in Homer (cf. supra) (Seiler, Capelle 1889: 371; O'Sullivan 1993a, cf. supra). Prellwitz (1905: 284) and Boisacq (1938: 616) referred to the labour-class *kheiromákhai* in Miletos who represented the 'hand-labourers' and had nothing to do with fighting. In addition, they pointed out that the Greek medic in the *Iliad* was called Makhaon. This proved in their opinion that *mákhomai* did not only mean 'fight' but also 'handle, treat with one's hands' and they therefore concluded that the link between *mákhomai* and *mákhaira* posed no problems. This analysis

[29] Lewy (1895: 177–178), with doubts because the suspected source was rare in Phoenician and Hebrew. The borrowing hypothesis was reiterated by Stella (1967: 121–122).

[30] See note 29.

[31] Sayce (1928: 162) noted that the words were identical, but did not say which language borrowed from which. Frisk (1970: 187) stated that Gordon had argued for a borrowing by Semitic coming from Greek. Rosół (2012: 192–193) rejected the borrowing by Greek because the meanings did not match.

[32] Lewy himself had already noted this.

[33] Schwyzer (1939: 475) only discussed the suffixation, but not the etymology.

is not convincing, however. First, *mákhomai* means 'fight' in Homer and never 'make' or 'treat with hands' or something similar (Ebeling 1885: 1023–1025; Seiler, Capelle 1889: 372; O'Sullivan 1993b). Second, Makhaon might be a medic but he was also a soldier and came to Troy with a contingent of soldiers (Seiler, Capelle 1889: 371, referred to *Iliad* 2,729–733). Third, *kheiromákhai* are attested much later. There is, however, a reason why *mákhaira* might have been derived from *mákhomai* and not from *mágeiros*. As a sacrificial knife is used to kill animals, it is more likely that such a word is derived from a verb from fighting, whereas a word for 'cook' is more likely to be formed from a verb that means 'prepare, handle (food)'. This brings us to the word *mákhomai*, which has no certain etymology either, cf. O'Sullivan (1993b: 45).[34] Five suggestions have been made. Fick (1901b: 320) linked *mákhaira*, *mákhomai*, *Makháo:n* with *me:khané:* 'means, trick' (Doric *ma:khaná:*). Wiedemann (1904: 62–63) followed this suggestion, added Germanic **mag* 'be able' to the equation and reconstructed **māgh*. Wiedemann's suggestion was expanded by Trümpy (for Greek) and adopted by Hofmann and LIV[2] (with a short vowel) (Hofmann 1950: 201; Trümpy 1950: 126–128; Zehnder 2001a: 422, with reference to Trümpy). The second suggestion was to link it with an alleged personal name *Amadzó:n* which would be the Aeolic adaptation of an Iranian name *hamazan* 'warrior' (Hofmann 1950: 192–193; Pokorny 1959: 697). This suggestion is very unlikely (Frisk 1970: 188, *ebenso geistreich wie unsicher*). A third suggestion is to connect it with Vedic *makha-* 'fighter' and to reconstruct the Greek and Vedic words as **mak*^{h}.[35] Grassmann (1873: 971) started from an original meaning 'hit with a (sacrificial) knife, slaughter, fight' and linked Greek *mákhomai*, Vedic *makhá* and Latin *mactāre* 'to slaughter'.[36] He argued that the other meanings of the word, 'hero' and 'enemy' could both be derived from 'fight', but Macdonell (1893: 272) and Monier Williams (1899: 772) translated the word as 'joyful, vigorously'.[37] After careful consideration, Renou (1966: 141) argued that the original meaning was 'fight' after all.[38] The fourth suggestion is that by Malzahn, Peters (2008: 266–267, without mentioning the Dutch word), who link it with Tocharian *mäke* 'run' and reconstructed **meg*^{h}*H* with a metathesis in Greek. They linked *mákhlos* 'lascivious' as well (cf. infra) and started from the meaning 'run': *mákhomai* originally meant 'run' and evolved via 'run aggressively' into 'fight', while *mákhlos* originally meant 'running' and this evolved into 'lascivious' (for this evolution there would be a parallel in German *läufig* 'sexually in heat [of females]' and Dutch *loops* 'sexually heated [usually of female animals]').

---

[34] See also the doubts in Chantraine (1968: 673–674) and Frisk (1970: 187–188).

[35] This had been suggested by Kuhn (1855: 19–21); Grassmann (1873: 971); Curtius (1879: 327); Renou (1966: 141); Dunkel (1979: 259); Mayrhofer (1996: 288, with doubts). **mak*^{h} is our suggestion and not that by Mayrhofer.

[36] The link with the Latin word was preserved in Wiedemann (1904: 62) and Lewis, Short *s. u.* but the other etymological dictionaries denied the link (Walde, Hofmann (1954: 5) and Hofmann (1935: 8) linked with Greek *másso:* 'I knead' while Ernout, Meillet (1967: 376) and De Vaan (2008: 357) denied it had any cognates at all).

[37] This meaning is also mentioned in Mayrhofer (1996: 288).

[38] Chantraine (1968: 673–674) stated that the meaning was uncertain, but referred to Renou (1966: 141) who suggested that the original meaning was 'fighter'.

In that case, the link with Sanskrit has to be given up. What argues against this equation, is that Greek would have preserved two derivations from the root but that none of them maintained the original meaning. The fifth suggestion is that by Beekes (2010: 916) who called the word "probably PG", because it was isolated and because in the field of fighting inherited terms were unlikely. The last suggestion is *non liquet*, which is in our opinion only the last resort when there are really no other options. If the meaning of Sanskrit *makhá* was indeed 'fight', there is in our opinion nothing that argues against a link between the Greek and Sanskrit word, as they correspond perfectly in form and it would be a violation of Ockham's Razor to state that the Sanskrit word was borrowed from a Dravidian language,[39] and that the Greek word was borrowed from another non-Indo-European language.[40] Linking the Latin word is more problematic: phonologically, a form **mak^h* could account for the Latin *mac-tāre* (as a Latin sequence *mac* can only come from a root with an *a* in it), but semantically, it is more problematic, because one would then have to start from a meaning 'hit with a (sacrificial) sword', which would have been preserved in Greek *mákhaira* and Latin *mactāre* and would have evolved into 'fight' in *mákhomai* and *makhá*. This cannot be ruled out, but it seems more cautious to link the Greek and the Sanskrit word; to posit a Graeco-Aryan isogloss **mak^h-* 'fight' (rather than to link *mákhomai* with the Germanic root *mag*[41]) and to assume that *mákhaira* was a secondary derivation from *mákhomai*.

17. *mákhlos* 'lascivious (of a woman), horny' (Beekes 2010: 915–916). Prellwitz (1905: 284, with doubts) linked this word to Sanskrit *makha*,[42] but this is semantically unlikely (Chantraine 1968: 673; Frisk 1970: 187; Beekes 2010: 916). Furnée (1972: 209, 211) compared the word with the god Bakkhos and Beekes therefore suggested this word was PG because of the variation *m*/*b*. As we argued elsewhere, allowing such widespread variations is problematic, because it is not falsifiable and enables one to link almost any set of words (De Decker 2015).[43] Malzahn, Peters (2008: 267) argued that the word was related with Tocharian *mäke* 'run' and compared German *läufig* '(sexually) in heat (of females)', but as we argued above, the problem is that the alleged root **meg^{h}H* 'run' would then only have survived in Greek in two words with changed meanings and with metathesis. As the word is attested in Armenian *mahaz* 'lascivious' (as noted by Beekes himself) and is close in meaning to the Greek word, it could very well present another Helleno-Armenian isogloss.[44] If the words were independently borrowed from the same language, it would mean that the PG

[39] This was suggested in Mayrhofer (1996: 288).

[40] We refer to Dunkel (1979: 259): "no convincing argument has ever been made against the connection of *makhá* and *mákhe:*, only alternative suggestions".

[41] Neither Kluge et. al. (1957: 484) nor the online German lexicon DWDS included the verb *mákhomai* among the cognates of the German *mögen* 'be allowed, be able'.

[42] The link was recently reiterated by Malzahn, Peters (2008: 267, without mentioning Prellwitz).

[43] Similar arguments were made in Verhasselt (2009a, 2009b, 2011), and in Meissner (2014).

[44] It was not addressed in Clackson (1994) nor in Martirosyan (2008) and the Armenian word was not mentioned in Prellwitz (1905: 284) nor in Frisk (1970: 187).

language was spoken in an area vast enough to have influenced both Greek and Armenian (either when they were still together or individually). The question is if assuming an Helleno-Armenian isogloss would not be the more economic option.

18. *me:khané:* 'tool, ruse'.
19. *mánganon* 'ruse' (Beekes 2010: 949–950).

The question is if both Greek words can be considered related to Gothic, German and Dutch *mag* 'he is able', OCS *mogǫ* 'be able' and possibly also Vedic *maghá* 'gift', in which case they would go back to the root *$*magʰ$* 'powerful' (Von Miklosich 1886: 199; Osthoff 1891: 216–217; Chantraine 1968: 700; Frisk 1970: 235).[45] Beekes rejected the connection with the Germanic and Slavic forms because of the non-existence of PIE *$*a$* and pointed to a suggestion by Van Beek, who compared this word with *mánganon* 'ruse' (Van Beek *apud* Beekes 2010: 949–950). As *mánganon* had a voiced stop and "pre-nasalization" and *me:khané:* had a voiceless stop, this word was considered to be PG. Neither argument is convincing. First, there are words that prove the existence of a phoneme *$*a$* for PIE. An important example is in our opinion the word for 'blind', which is *caecus* in Latin and is related to Sanskrit *kekaras* 'squinting'. The Sanskrit word rules out reconstructions *$*kh₂eik$* or *$*keh₂ik$* as the former would have given **$*khekaras$* and the latter **$*kaikaras$*. Second, the consonant variation *ng*/*kh* can be explained as the result of an internal Greek sound law. If we start from a stage of Proto-Greek with the voiced aspirates still present, we could assume that from the form *magh*-, a derivation with a nasal infix *n* and suffix *an* (a similar derivation occurred in *túmpanon* 'kettledrum' derived from *túpto:* 'I beat') was made, namely *$*mánghanon$*. In that form, the voiced aspirate was preceded by a nasal and also preceded by the accented syllable. Under these conditions, Greek rendered the voiced aspirate by a voiced stop. This is known as Miller's Law (Miller 1977a: 151, 1977b: 37–38). As such, *mánganon* is an expected outcome and is not an indication of Pre-Greekness. The long vowel in *me:khané* needs an explanation as well. Frisk, following Schwyzer, suggested that besides the *s* stem *mêkhos* 'means, remedy' with lengthened grade, there was also an *s* noun *$*mâkhar, *mákhanos$* from which a feminine and oxytone noun *me:khané:* was derived with the lengthened grade (although the lengthened grade from *mêkhos* could have contributed as well) (Schwyzer 1939: 459; Chantraine 1968: 700; Frisk 1970: 235). The Greek forms could also be explained from the root *$*māgh$* with a long vowel. In that case, the short vowel of *mánganon* could be explained by Osthoff's Law.[46] As the variations between the Greek words can be explained by internal Greek sound laws and there are cognates in other Indo-European languages, we believe that there is no need to catalogue this word as PG.

---

[45] Frisk noted that Bopp and Pott had already made this equation. The connection with Greek goes back to Fick (1901b: 320) and Wiedemann (1904: 62–63), and was adopted in Pokorny (1959: 659), Mayrhofer (1996: 289) and Zehnder (2001a).

[46] This sound law, which is not of Indo-European date as it does not operate in Indo-Iranian and Tocharian, states that a long vowel is shortened in Latin and Greek when it is followed by a resonant and a consonant.

20. *molobrós* (uncertain meaning) (Beekes 2010: 963). This word is used as epithet for Odysseus when he was still disguised as a beggar and therefore must have had a negative or derogatory meaning. Earlier attempts to explain this word are phonologically impossible. One example is that by Fick (1904: 97), who linked it with *blábe:* 'damage' from a root **mleb* which would have had *molob* in the full grade. This is impossible (Frisk 1970: 250–251). This word is not only attested in Homer but also appears in Mycenaean as *mo ro qo ro* and therefore the word must have had a labiovelar. Chantraine considered this word a compound of *molo* and **g*^{w}*ro* and translated it as 'animal qui dévore les jeunes pousses' (Chantraine 1968: 709, 1972: 203–205). The first part of the Greek word is not attested in the meaning 'flea', however, but has an equivalent in Indic *mala* 'dirt, shit' and the second is a form of the root **g*^{w}*erh*$_3$ 'devour' in the zero grade (Normier 1980: 276; Neumann 1992: 75–80; De Leeuw 1993). As this is a compound, the form was subject to the so-called *neognós* rule, which states that in compounds or reduplicated forms, a laryngeal is lost when it is preceded by a sonorant and followed by a vowel:[47] thus **molo-g*^{w}*rh*$_3$*-os* became **molog*^{w}*ros*, leading to the attested Greek form. Beekes considered this word PG, because he did not accept the loss of laryngeals in compounds (Beekes 2010: 963 *I do not accept the loss of laryngeals in compounds*). This is strange, because elsewhere in his dictionary and his publications he mentioned the *neognós* rule. As this word can be explained from an Indo-European perspective and has a meaning that makes perfect sense in the context ('Dreckfresser'),[48] we see no need to assume PG origin.

21. *mogéo:* 'I am in distress, suffer' (Beekes 2010: 960–961).
22. *mógos* 'pain' (Beekes 2010: 960).
23. *mókhthos* 'difficulty, distress' (Beekes 2010: 973).
24. *mókhlos* 'handle, long or strong rod' (Beekes 2010: 973).

The first three words are clearly linked and the basic word is *mogéo:* (Chantraine 1968: 707–708). Attempts have been made to find an Indo-European etymology. Schulze

[47] The Greek word *neognós* 'newly born' is a compound **neo-ǵnh*$_1$ *-os* of *néos* 'new' and *ǵenh*$_1$ 'originate'. In that compound, the laryngeal is lost. Beekes is considered the inventor of that rule, see Beekes (1969: 241–245, pointing out that Hirt might be the first one to state that laryngeal loss in compounds could occur; 1982: 114; 1988: 60–61, pointing out that there were only a few examples). In Beekes (2010: 1079) he mentioned the rule.
Mayrhofer (1986: 129) assumed that the rule applied to **h*$_1$ alone, but this example proves that all laryngeals were subject to this rule. Weiss (2009: 113) stated that the rule operated in "long" words such as reduplications and compounds. See also Byrd (2015: 26).
There are, nevertheless, several exceptions to this rule, especially – but not exclusively – in reduplicated presents, such as Greek *titró: sko:* 'I wound' from **terh*$_3$. The reduplicated form **titrh*$_3$*sḱoh*$_2$ should have given **titrsḱoh*$_2$ by the *neognós* rule and also by the Schmidt-Hackstein rule (this rule, based on Schmidt 1973 and Hackstein 2002 states that a cluster CHCC became CCC in PIE). This last form should have given Greek **titrásko:* but it is very likely that this form was reformed (or that the laryngeal was reintroduced) after the aorist *étro: sa* 'I wounded'. Byrd (2015: 85–125) argued that the Schmidt-Hackstein rule only applied in cluster PHCC (i.e. only when the consonant preceding the laryngeal was a plosive and not just any consonant), but even if this is correct, the presents are still an exception to the *neognós* rule. Analogical reintroduction seems the only possible solution.

[48] This is the translation suggested by Neumann and De Leeuw.

(1887: 270) tried to connect these words with Latin *mōlēs* 'burden, heavy weight to carry', Solmsen (1888: 85–86) suggested to link *mógos* with Lithuanian *smagùs* 'heavy to carry, heavy to drag' and Latvian *smags* 'heavy'. Meier-Brügger (1993) suggested that the Greek words contained the *o* grade of the adjective *mégas* 'big' and compared *mélas* 'black' and *mólos* 'dirt'. The suggested cognates are not certain: the link with Latin *mōlēs* has been doubted by Walde, Hofmann (1954: 102), and neither Ernout, Meillet (1967: 410) nor De Vaan (2008: 386) even mentioned the link,[49] and according to Fraenkel the Baltic words should be linked to Lithuanian *smagiù* 'hit, throw something heavy' (quoted in Frisk 1970: 262). The connection between *mókhlos* and Latin *mōlēs* can, however, not be excluded a priori, if one reconstructs *mogslo-* with the suffix *slo* that is attested elsewhere in Latin as well (as can be seen in*pālus* 'pale, stake' from **pagslo-*) (Schulze 1887: 270; Walde, Hofmann 1954: 243; Ernout, Meillet 1967: 478; Frisk 1970: 262). The Greek form *mókhlos* 'handle' could be included assuming that a handle is a tool to perform (heavy) labour: it could be an original *mogslo-* with a suffix *slo*. The form would then have lost the interconsonantic sigma and have aspirated the other consonants, which occurred in *érkhomai* 'I go' from Proto-Greek **erskomai*. This is better than Chantraine's (1933: 240) explanation that the suffix was *lo* and that the aspiration was expressive. The main problem with Schulze's equation is the difference in declension type in Greek and Latin. Solmsen's (1888) explanation assumes a link between 'heavy' and 'difficult' which is acceptable, but Meier-Brügger's (1993) suggestion is more problematic. He assumed an evolution from 'big' into 'heavy' into 'difficult' and explained the verb's original meaning as 'groß machen' which became then 'unter großer Anstrengung tun'. A last remark involves the form. If **me/oǵh*$_2$ were the basis, would one not have expected Greek **mogáo:*? If the Baltic cognates are not related, the word has no etymology (but even with Latin and Baltic cognates, a PIE origin would not have been entirely certain). Beekes argued that the links of *mogéo:* with other Indo-European languages were hardly credible. He also argued that the form *móklos* (attested in Anakreon, living in Asia Minorin the 6th century BC) with a plain voiceless plosive instead of an aspirated one in *mókhlos* and the variation *khth* in *mókhthos* and *g* in *mogéo:* proved that all the words were PG. We do not believe that the derived words are evidence for PG origin. The variation between *móklos* and *mókhlos* can be explained by an influence of Anakreon's dialect as well. As he lived in Asia Minor and wrote in Ionic (which was a psilotic dialect), the form without an aspirate could reflect his everyday speech. The aspirate in *mókhlos* has been explained above. For *mókhthos*, one can assume that the suffix *-thos* was added to the stem *mog*.[50] This suffix can be used in words referring to difficult situations or negative feelings such as *ákhthos* 'burden, burden of pain' besides *ákhos* 'pain'. Chantraine (1933: 366–367) considered this to be an inherited expressive suffix, visible in Sanskrit *tha*.[51] Another explanation is

[49] De Vaan stated that *mōlēs* did not have a good etymology.

[50] Rather than a suffix *sdho* (Schulze 1887: 270) or *stho* (Prellwitz 1905: 301).

[51] Frisk (1970: 261–262) agreed with the Greek expressive suffix but did not mention the Sanskrit suffix.

a link with **$d^{h}eh_{1}$* 'put' and then the suffix *thos* would be an original **$d^{h}h_{1}os$* meaning 'bringing, carrying X'. It is therefore not correct to use the variation *g/khth* as proof for Pre-Greekness. Even if *mogéo:* were of non IE origin, the noun *mókhthos* could represent a regular Greek compound.

25. *mudáo:* 'to be humid' (Beekes 2010: 974).
26. *múdros* 'metal roasted in fire, glowing stones' (Beekes 2010: 975).

As *múdros* represents the molten iron, a link with the verb *mudáo:* is very likely (Debrunner 1908: 5, 9). Beekes considered the noun to be PG because of the anlaut *sm* which is also attested and because of the technical meaning. We believe that this is not necessary. Many words of technical meaning are of non-IE origin, but that does not mean that they are all of such origin; second, the anlaut *sm* could be onomatopoeic or could have been influenced by other words with an anlaut *C/sC*. The verb is also attested with a long *u*, which is explained as metrical lengthening (Curtius 1873: 336; Frisk 1970: 263; Beekes 2010: 974). Beekes argued that the verb was PG as well because of the link with the adjective *músos* (variation *d/s* in one word would point to PG) and because the verb was attested with a long and short vowel. We believe that the arguments are not convincing. First, the distinction in vowel length might be ascribed to metrical lengthening (as Beekes admitted himself) and as such, it has no probative value. Second, there is the connection with the Dutch word *mot(regen)* 'light rain' and the Sanskrit word *mudirá* 'cloud'[52] (but this is attested only in the Classical Sanskrit period). As the Indic word also means 'lover' according to the lexica, it is often linked to the noun *mud-* 'joy' (Curtius 1873: 336; Frisk 1970: 263), but this connection is rather doubtful and maybe there was an Indic root *MOD* 'wet' besides *MOD* 'rejoice' (Mayrhofer 1996: 383).[53] As such, we believe that the word is of Indo-European origin (as Germanic, Greek and Indic did not have shared innovations).[54]

27. *múdos* 'voiceless, numb' (according to Hesykhios) (Beekes 2010: 975).
28. *mukós* 'speechless'.

These two words have been explained by Hesykhios as *ápho:nos* 'speechless'. As *múdos* is also attested as *múndos*, Beekes argued that the words were PG because of the so-called "pre-nasalization". This is not certain, however. There is the word *munǰ* in Armenian, which could continue **mund-yo-*. If this were the case, it could be an Helleno-Armenian isogloss and consequently, PG would be excluded. Clackson (1994)

---

[52] Boisacq (1938: 648, without mentioning the Dutch word), Hofmann (1950: 206), Chantraine (1968: 718). Frisk (1970: 263) mentioned the connection, but was doubtful about an etymological link.

[53] He mentioned the Greek words but not the Dutch one.

[54] The recent Dutch etymological dictionary by Van Veen, Van der Sijs (1997: 569, 579) linked the Dutch word with Dutch *modder* 'mud', English *mud* (this connection had already been made by Prellwitz 1905: 301), Greek *mudrós* and Avestan *muthra* 'faeces' and Indic *mutra* 'urine', but this is unlikely as an Indo-Iranian *t*, a Greek *d* and a Germanic *d* cannot be linked with each other.

considered the reconstruction **mundy* or **mūndy* for the Armenian form *somewhat dubious*,[55] but he did not discuss the Greek word. Personally, we do not see why a link between the Greek and Armenian word would be excluded. It also seems that the word *mukós* cannot be separated either. In other languages, there are words with an initial *mu*/*mū* that refer to the absence of speaking such as Latin *mūtus* 'dumb' and Sanskrit *mūka* 'dumb' with different extensions (Ernout, Meillet 1967: 427; Chantraine 1968: 720; De Vaan 2008: 398). A connection with sound imitating *mū* is possible (Frisk 1970: 268), but then the connection with *mûthos* (cf. infra) is difficult, because that word means 'word' and not 'dumb, deaf, soundless'. One single etymology that connects all different words is not available, but it seems that all words are derivations with different suffixes from the onomatopoeic root *mu*/*ū*. As such, there is no need to posit a PG origin for these words.

29. *mûthos* 'word' (Beekes 2010: 976). This word is in formation very similar to the words mentioned above and can -at least formally- be explained by the root **mū* and the suffix **thos* (be it from **$t^h o$-* or **$d^h h_1 o$-*).[56] There are two problems with this explanation: first, there is no indication in the texts that *mûthos* was ever an onomatopoeic word (Chantraine 1968: 719: "mais le sens du mot, dès les plus anciens textes, n'est pas en faveur de cette hypothèse") and second, the other words derived from this root mean 'dumb, not speaking', which is exactly the opposite. The second observation is maybe less problematic, if one assumes an initial meaning 'producing the *mû* sound' which would have evolved into 'producing a sound' and eventually into 'what is produced by the mouth, (namely) word'. Beekes suggested that the word was PG because there were no comparanda, but not every Greek word without an Indo-European etymology is of substrate origin.

30. *mukhós* 'hiding place, innermost place, storage room' (Beekes 2010: 987). According to Wace (1951: 209–210), the word was also used to refer to the private rooms of the master of the house. Fick (1909: 149) linked *múskhon* 'female and male genitalia' with *mukhós* as well and suggested an etymology **mukh-sko*, in which the first velar was lost but the aspiration was transferred to the last velar. Semantically, there is no problem with this explanation, as the genitalia are those parts that remain hidden. Cognates of this word in other Indo-European languages are OCS *smykati se* 'crawl', Lithuanian *smùkti* 'glide', ON *smjúga* 'to slip in' (Frisk 1970: 279; Beekes 2010: 987), with the first two forms derived from **smuk* and the last one from **$smug^h$*. Frisk connected *mokhós* with the Armenian verb *mxem* 'immerse', assumed a basis meaning 'stuff away, hide' from which the Greek word received its meaning 'hiding place' and reconstructed **muk^h*.[57] Clackson (1994: 182) considered the etymology doubtful and Beekes (2010: 987) rejected this reconstruction because

[55] Clackson (1994: 45), talking about the origin of the Armenian *u* (which is often written ‹ow› as well). Similar doubts were already voiced by Frisk (1970: 269).

[56] As was already done by Curtius (1873: 336, without discussing the suffix). See also Frisk (1970: 264–265).

[57] Already Solta (1960: 160) had linked these words. See Clackson (1994: 182).

voiceless aspirates were no longer accepted. As we stated above, voiceless aspirates are indeed much rarer that the other plosives, but there are a few words where another explanation is not possible (the same applies to PIE **b* and **a*) and "rare" does not mean "non-existent". As the root **muk^h* with its voiceless aspirate is only attested in Greek and Armenian, it is not certain that it can be reconstructed for PIE. Other Indo-European languages display roots of a similar form but with different velars (**muk* and **mug^h*). Chantraine (1968: 728) explained this by the expressive nature of the word, and Frisk (1970) suggested that the different roots could be unified into one root with several allophonic variants due to assimilation with contiguous consonants. Maybe the Helleno-Armenian innovation was that from a root **muk* and **mug^h* a root **muk^h* was extracted? Beekes mentioned that Furnée considered this word PG, but that the arguments were lacking (Beekes 2010: 987–988, referring to Furnée 1972: 364). Given the fact that this word is attested in several languages, PG origin is in our opinion excluded.

31. *múo:ps* 'gadfly, goad, spur' (Beekes 2010: 989). Prellwitz (1905: 192) started from the assumption that cattle feared this animal and that it recognized the insect by its tone. He therefore suggested a compound of the *mu* 'mumming sound' and *o:ps*. Boisacq (1938: 65) explained this word as a compound of *muîa* 'fly' and *o:ps* 'seeing'. The meaning would then be 'what looks like a fly'. This etymology was accepted by Hofmann (1950: 209) and Frisk (1970: 281), but rejected by Chantraine (1968: 729). Beekes considered it unlikely and rather suggested PG origin because of the suffix *o:ps* that could be found in other insect names such as *kó:no:ps* 'gnat'. As there is a word *múo:ps* 'short sighted' which is a compound from *múo:* 'I close' and *o:ps* 'seeing' and which literally means 'with closed vision, (hence) 'short-sighted',[58] there is nothing that rules out that and in this case we would have another *o:ps* compound 'with fly-looks'.

3. Conclusion

While it was not our goal to rewrite the dictionary, we hope to have shown that many of the words catalogued as <PG> or <PG?> allowed for other explanations as well (this is the reason why we often decided to discuss earlier etymologies as well). In several instances, there was no agreement on an etymology or there was no established etymology altogether, but in many instances, an Indo-European etymology was available. We never argued (nor will we ever argue) that each and every word in Greek has to have an Indo-European etymology nor that there are no borrowings in the Greek lexicon. The main intention of this article was to show that in establishing etymologies one should look at the evidence and not be searching for borrowings when they are not there, and that strict and falsifiable rules should be used.

[58] This analysis was accepted by Beekes (2010: 989) as well.

References

Bally C., Gautier L. 1922. *Recueil des publications scientifiques de Ferdinand De Saussure*. Genève.

Beekes R. 1969. *The development of the Proto Indo European laryngeals in Greek*. The Hague.

Beekes R. (rev.). 1982. Peters 1980. – *Kratylos* 26: 106–115.

Beekes R. 1988. Laryngeal developments: A survey. – Bammesberger (ed.). 1988: 59–105.

Beekes R. 2010. *Etymological dictionary of Greek*. Leiden.

Benveniste E. 1935. *Origines de la formation des noms en indo-européen*. Paris.

Boisacq E. 1938. *Dictionnaire étymologique de la langue grecque*. Heidelberg.

Brugmann K. 1876. Zur Geschichte der stammabstufenden Declinationen, Erste Abhandlung: Die Nomina auf *ar* und *tar*. – *CS* 9: 361–406.

Brugmann K. 1894. *Moûsa; tríaina, thrínaks, Thrinakíe; eneika*. – *IF* 3: 253–264.

Brugmann K. 1905. Alte Wortdeutungen in neuer Beleuchtung. – *IF* 18: 423–439.

Byrd A. 2015. *The Indo-European syllable*. Leiden.

Chantraine P. 1933. *La formation des noms en grec*. Paris.

Chantraine P. 1948. *Grammaire homérique. Tome I: morphologie*. Paris.

Chantraine P. 1968. *Dictionnaire étymologique de la langue grecque*. Paris.

Chantraine P. 1972. Le témoignage du mycénien pour l'étymologie grecque: *daí, Kopreús, Kukleús, molobrós, mólubdos*. – *Minos* 12: 197–206.

Clackson J. 1994. *The linguistic relationship between Armenian and Greek*. Oxford, Cambridge (MA).

Clackson J. 2007. *Indo-European linguistics*. Cambridge.

Cuny A. 1912. Notes de phonétique historique: indo-européen et sémitique. – *RPhon* 2: 101–132.

Curtius G. 1873. *Grundzüge der griechischen Etymologie*. Leipzig.

Curtius G. 1879. *Grundzüge der griechischen Etymologie*. Leipzig.

Debrunner A. 1908. Die Adjektiva auf *aleos*. – *IF* 23: 1–43.

De Decker F. 2011. Evidence for laryngeal aspiration in Greek. Part 1: The "recent" evidence. – *IF* 116: 87–109.

De Decker F. (rev.). 2015. Beekes 2014. – *IJDL* 13: 1–23.

De Decker F. [forthcoming a]. Evidence for laryngeal Aspiration in Greek? Part 2: The older evidence. – *Orbis*.

De Decker F. [forthcoming b]. What is the Greek counterpart of (Proto-)Indo-Iranian (*)th? – *IJDL*.

De Leeuw J. 1993. *molobrós*. – *LfgrE* 15: 252.

De Vaan M. 2008. *Etymological dictionary of Latin and the other Italic languages*. Leiden.

Dunkel G. 1979. Fighting words: Alcman *Partheneion* 63 *mákhontai*. – *JIES* 7: 249–272.

DWDS = Digitales Wörterbuch der Deutschen Sprache. [Accessible online: www.dwds.de].

Ebeling H. 1885. *Lexicon Homericum* I. Leipzig.

Ehrlich H. 1907. Zur Mythologie II. *Moûsa*. – *KZ* 41: 287–289.

Elbourne P. 1998. Proto-Indo-European voiceless aspirates. – *KZ* 111: 1–30.

Elbourne P. 2000. Plain voiceless stop plus laryngeal in Proto-Indo-European. – *KZ* 113: 2–30.

Elbourne P. 2001. Aspiration by *s* and devoicing of Mediae Aspiratae. – *KZ* 114: 197–219.

Elbourne P. 2011. *epírrhothos*. – *Glotta* 87: 37–57.

Elbourne P. 2012. A rule of deaspiration in Greek. – Probert P., Willi A. (eds.). *Laws and rules in Indo-European*. Oxford: 125–133.

Ernout A., Meillet A. 1967. *Dictionnaire étymologique de la langue latine*. Paris.

Fick A. 1872. Etymologische Beiträge. – *KZ* 20: 161–181.

Fick A. 1890. *Wörterbuch der indogermanischen Sprachen*. [vol. 1; 4th edition]. Göttingen.

Fick A. 1901a. Die griechischen verbandnamen (ethnika). – *BB* 26: 233–265.
Fick A. 1901b. Asklepios und die heilschlange. – *BB* 26: 313–323.
Fick A. 1904. Hesychglossen. – *BB* 28: 84–111.
Fick A. 1909. Hesychglossen VI. – *KZ* 43: 130-153.
Fortson B. 2004. *Indo-European language and culture. An introduction*. Oxford.
Frisk H. 1936. Suffixales *th* im Indogermanischen. – *GHÅ* 42: 3–46.
Frisk H. 1960. *Griechisches Etymologisches Wörterbuch*. [vol. 1: A–Ko]. Heidelberg.
Frisk H. 1970. *Griechisches Etymologisches Wörterbuch*. [vol. 2: Kr–O]. Heidelberg.
Furnée E. 1972. *Die wichtigsten konsonantischen Erscheinungen des Vorgriechischen*. Amsterdam.
Grammont M. 1948. *Phonétique du grec ancien*. Paris.
Grassmann H. 1873. *Wörterbuch zum Rig Veda*. Leipzig.
Greppin J. 1981. Gr. *mallós* "fleece, lock of wool". – *Glotta* 51: 70–75.
Hackstein O. 1995. *Untersuchungen zu den sigmatischen Präsensstammbildungen des Tocharischen*. Göttingen.
Hackstein O. 2002. Uridg. *CH.CC > *C.CC. – *KZ* 115: 1–22.
Hamp E. 1982. *mallós* a clarification. – *Glotta* 60: 61–62.
Harðarsson H. 1998. Mit dem Suffix *-*eh*$_1$- bzw. *-*(e)h*$_1$*-ye/o-* gebildete Verbalstamme im Indogermanischen. – Meid W. (ed.). *Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen Gesellschaft*. Innsbruck: 323–339. [quoted in Piwowarczyk forthcoming and Verhasselt forthcoming].
Hofmann J. 1935. *macerō*. – *Thesaurus Linguae Latinae* 8: 8–10.
Hofmann J. 1950. *Etymologisches Wörterbuch des Griechischen*. München.
Kluge F. et al. 1957. *Etymologisches Wörterbuch der deutschen Sprache*. [Many later reprints]. Berlin.
Krahe H. 1928. *Kháones, Khônes*. – *Glotta* 17: 158–160.
Kuhn A. 1855. Über das alte S und einige damit verbundene lautentwicklungen. Sechster artikel. Der anfall des s vor mutis. – *KZ* 4: 1–46.
Kuiper F. 1934. *Die indogermanischen Nasalpräsentia*. Amsterdam.
Kümmel M. 2001. *$^*u̯erh_1$-*. – *LIV* 2: 689–690.
Kümmel M. [forthcoming]. The lengthened grade in the nominative singular. [Accessible online: https://www.academia.edu/10167321/The_lengthened_grade_in_the_nominative_singular (Accessed August 3rd, 2015)].
Kuryłowicz J. 1927. Les effects du ə en indo-iranien. – *Prace Filologiczne* 11: 201–243.
Kuryłowicz J. 1928. Le type védique gṛbhāyati. – Meillet A. (ed.). *Étrennes de linguistique offertes par quelques amis à Émile Benveniste*. Paris: 51–62.
Kuryłowicz J. 1935. *Études indo-européennes*. Wrocław.
Lewy H. 1895. *Die semitischen Lehnwörter im Griechischen*. Berlin.
LfgrE 15 = Meier-Brügge M. (ed.). 1993. *Lexikon des frühgriechischen Epos*. μά – νεῆνις. Göttingen.
Lindeman F. 2004. On the vocalization of 'laryngeals' in Indo-European. – *KZ* 117: 118–133.
LIV^2 = Rix H. et al. 2001. *Lexikon der indogermanischen Verben*. Wiesbaden.
Macdonell A. 1893. *A Sanskrit – English dictionary*. London.
Malzahn M., Peters M. 2008. How (not) to compare Tocharian and Ancient Greek verbal stems. – Kim R. et. al. (eds.). *Ex Anatolia Lux. Anatolian and Indo-European studies in honor of Craig Melchert*. Ann Arbor: 265–268.
Martirosyan H. 2008. *Studies in Armenian etymology, with special emphasis on dialects and culture*. [PhD thesis]. Leiden.
Mayrhofer M. 1981a. Ferdinand de Saussure. Nach hundert Jahren. – *Sitzungsberichte der Heidelberger Akademie der Wissenschaften:* 7–38.

Mayrhofer M. 1981b. Laryngalreflexe im Indo-Iranischen. – *ZPSK* 34: 427–438.
Mayrhofer M. 1986. *Indogermanische Grammatik*. Heidelberg.
Mayrhofer M. 1996. *Etymologisches Wörterbuch des Altindoarischen*. [Band 2]. Heidelberg.
Mayrhofer M. 2005. *Die Fortsetzung der indogermanischen Laryngale im Indo-Iranischen*. Wien.
Meier-Brügger M. 1993. Zu griechisch *mogéo:* und Familie. – *Glotta* 70: 134–135.
Meier-Brügger M. 2010. *Indogermanische Sprachwissenschaft*. [Zweite erweiterte Auflage]. Berlin.
Meiser G. 2005. Lat. mando, mandi 'kaue'. – Marchesini S., Poccetti P. (eds.). *Linguistica è storia. Scritti in Onore di Carlo di Simone*. Pisa: 121–125.
Meissner T. (rev.). 2014. Beekes 2010. – *Kratylos* 58: 1–31.
Miller G. 1977a. Was Grassmann's Law reordered in Greek? – *KZ* 91: 131–158.
Miller G. 1977b. Some theoretical and typological implications of an IE root structure constraint. – *JIES* 5: 31–40.
Monier Williams M. 1899. *A Sanskrit – English dictionary*. Oxford.
Narten J. 1960. Das vedische Verbum math-. – *IIJ* 4: 121–135.
Neumann G. 1992. Griechisch *molobrós*. – *KZ* 105: 75–80.
Nikolaev A. 2010. Indo-European **demh₂* „to build" and its derivatives. – *KZ* 123: 56–96.
Normier R. 1980. Tocharisch *ñakt/ñakte* 'Gott'. – *KZ* 94: 251–281.
Osthoff H. 1891. Das Praeteritopraesens *mag*. – *PBB* 15: 211–218.
O'Sullivan J. 1993a. Mákhaira. – *LfgrE* 15: 44.
O'Sullivan J. 1993b. Mákhomai. – *LfgrE* 15: 45–49.
Pedersen H. 1893. r/n stämme. Studien über den stammwechsel in der declination der idg. Nomina. – *KZ* 32: 240–273.
Pedersen H. 1926. *La cinquième déclinaison latine*. Copenhagen.
Peters M. 1980. *Untersuchungen zur Vertretung der indogermanischen Laryngale im Griechischen*. Wien.
Peters M. 1993a. Oresq- neben orestā. – *Die Sprache* 35: 135–139.
Peters M. 1993b. Beiträge zur griechischen Etymologie. – Isebaert L. (ed.). *Miscellanea Linguistica Graeco-Latina*. Namur: 85–113.
Pinault G. 1982. A neglected phonetic law: The reduction of the Indo-European laryngeals in internal syllables before yod. – Ahlqvist A. (ed.). *Papers from the 5th International Conference on Historical Linguistics*. Amsterdam: 265–272.
Piwowarczyk D. 2008. *The treatment of laryngeals before yod in Indo-European verbs*. [MA thesis]. Leiden.
Piwowarczyk D. [forthcoming]. Some remarks on Pinault's Rule in Greek.
Pokorny J. 1959. *Indogermanisches Etymologisches Wörterbuch*. Bern.
Prellwitz W. 1899. Lat. *flagitium*, lit. *Blōgas*. – *BB* 25: 280–288.
Prellwitz W. 1905. *Etymologisches Wörterbuch der griechischen Sprache*. Göttingen.
Rasmussen J. 1987. On the status of the aspirated tenues and the Indo-European phonation series. – *ALH* 20: 81–109.
Rasmussen J. 1989. Die Tenues Aspiratae, Dreiteilung oder Vierteilung des indogermanischen Plosivsystems und die Konsequenzen dieser Frage für die Chronologie einer Glottalreihe. – Vennemann T. (ed.). *The new sound of Indo-European: Essays in phonological reconstruction*. Berlin: 153–176.
Renou L. 1966. *Études védiques et paninéennes*. [15]. Paris.
Ringe D. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford.
Rosół R. 2012. *Frühe semitische Lehnwörter im Griechischen*. Frankfurt.
Ruijgh C. (rev.). 1978. Rix 1976. – *Mnemosyne* 4.31: 298–307.

Sayce H. 1928. Cybele and Gallos in the Hittite texts. – *CR* 42: 161–163.
Schmidt G. 1973. Die iranischen Wörter für 'Vater' und 'Tochter' und die Reflexe des interkonsonantischen H (ə) in den indogermanischne Sprachen. – *KZ* 87: 36–83.
Schulze W. 1887. Das lateinische v perfectum. – *KZ* 28: 299–274.
Schwyzer E. 1939. *Griechische Grammatik*. München.
Seiler E., Capelle C. 1889. *Vollständiges Wörterbuch zu den Gedichten des Homeros und der Homeriden*. Leipzig.
Solmsen F. 1888. Sigma in verbindung mit nasalen und liquiden im griechischen. – *KZ* 29: 59–124.
Solmsen F. 1909. Odysseus und Penelope. – *KZ* 42: 207–233.
Solta G. 1960. *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*. Wien.
Stella L. 1967. Una recensione di John Chadwick. – *SMEA* 4: 121–124.
Szemerényi O. 1996. *Introduction to Indo-European linguistics*. Oxford.
Trümpy H. 1950. *Kriegerische Fachausdrücke im griechischen Epos: Untersuchungen zum Wortschatze Homers*. Basel.
Van Veen P., Van der Sijs N. 1997. *Groot etymologisch woordenboek*. Utrecht.
Verhasselt G. 2009a. *Het Pre-Griekse substraat. Recente tendenzen binnen het taalkundig onderzoek*. [MA thesis]. Leuven.
Verhasselt G. 2009b. The Pre-Greek linguistic substratum. An overview of current research. – *LEC* 77: 211–239.
Verhasselt G. 2011. The Pre-Greek linguistic substratum. A critical assessment of recent theories. – *LEC* 79: 257–283.
Verhasselt G. [forthcoming]. Die Lex Pinault: Ein idg. Lautregel? – *KZ* 126.
Volkart M. 1994. *Zu Brugmanns Gesetz im Altindischen*. Bern.
Von Miklosich F. 1886. *Etymologisches Wörterbuch der slavischen Sprachen*. Wien.
Wace A. 1951. Notes on the Homeric house. – *JHS* 71: 203–211.
Wackernagel J. 1895. *Moûsa*. – *KZ* 33: 571–575.
Wackernagel J. 1896. *Altindische Grammatik. I. Lautlehre*. Göttingen.
Walde A., Hofmann J. 1954. *Lateinisches etymologisches Wörterbuch*. [M–Z]. Heidelberg.
Weiss M. 2009. *Outline of the historical and comparative grammar of Latin*. Ann Arbor.
Whitney D. 1885. *The roots, verb-forms, and primary derivatives of the Sanskrit language: A supplement to his SanskritgGrammar*. Leipzig.
Wiedemann O. 1904. Etymologien. – *BB* 28: 1–83.
Zehnder T. 2001a. **mag*^{h}. – *LIV*^{2}: 422.
Zehnder T. 2001b. **meld*^{h}. – *LIV* ^{2}: 431.
Zehnder T. 2001c. **menth*$_{2}$. – *LIV*^{2}: 438–439.
Zehnder T. 2001d. **meth*$_{2}$. – *LIV*^{2}: 442–443.
Zubaty J. 1892. Die ursprachliche tenuis aspirata im arischen, griechischen und latein. – *KZ* 31: 1–9.

Studia Linguistica Universitatis Iagellonicae Cracoviensis 133 (2016): 97–#
doi:10.4467/20834624SL.16.007.5153

MAGNÚS SNÆDAL
University of Iceland, Reykjavík
hreinn@hi.is

GOTHIC *BANJA**, *WINJA* AND *SUNJA*[1]

Keywords: etymology, word formation, Gothic, Old Norse, West-Germanic, Germanic

Abstract

The present paper discusses the etymology of three Gothic nouns: *banja** 'sore', *winja* 'pasture', and *sunja* 'truth'. Each of them has a cognate in Old Norse: *ben* 'fatal wound', *vin* 'oasis' and *syn* 'refusal'. None of the West-Germanic languages preserves all three nouns. All are short, feminine *jō*-stems with an *-n-* in front of the stem suffix. The main issue discussed here is the etymology and formation of these nouns.

Introduction

The present paper deals with the etymology of three Gothic nouns: *banja** 'sore', *winja* 'pasture', and *sunja* 'truth'. Each of these nouns has a cognate in Old Norse: *ben* 'fatal wound', *vin* 'oasis' and *syn* 'refusal'. None of the West-Germanic languages preserves all three nouns as will be evident later on. All three are short, feminine *jō*-stems (although ON *ben* can also be neuter and have the meaning 'small wound' outside legal vocabulary). The main issue discussed here is the etymology and derivation of these nouns. It is undisputed that Go. *sunja* is derived from the zero grade of the root of the verb 'to be', but a disputable, unique sound change is generally assumed to have been at work in this noun (see section 3).

Meid (Krahe, Meid 1967: 119–122) mentions some feminine nouns containing the Proto-Germanic suffix **-njō* with a "connecting vowel" ("Bindevokal"), *u* or *i*, e.g. Go. *Saurini* (gen. **-jos*) 'Syrian woman; Σύρα' from *Saur* 'Syrian; Σύρος', and

[1] An earlier Icelandic version of this paper was read at the University of Iceland in Reykjavík on 23 April 2012. I thank my colleague, Jón Axel Harðarson, for his extensive comments on the Icelandic version.

OIc. *Fjörgyn* (gen. *-jar*) 'Mother-Earth' (otherwise the weak form of the suffix is used, *-ynja*, e.g. *ljónynja* 'lioness'), possibly from *Fjörgynn* m. 'name of a heathen god'. According to Chantraine (1979: 107), relics of this formation are also found in Gr. *πότνια* 'goddess' < PIE **pótnih₂* from PIE **póti-* 'lord, husband', Gr. *πόσις* (Beekes 2010: 1226–1227), Skt. *indrāṇī́* 'Indra's wife' from *indraḥ*, and Slav. *bogynji* 'goddess', from *bogŭ* 'god' (e.g. Russ. *богúня*). Meid also mentions that it is uncertain whether the suffix was found in Proto-Germanic without a connecting vowel (cf. also Casaretto 2004: 330–332).

Obviously, in the nouns under discussion the suffix is not used to form a feminine noun from a thematic masculine noun (expressing semantic relations of the type 'goddess' ← 'god'). Here, whether the Gothic nouns *banja**, *winja*, and *sunja* are all derived from the zero grade of the respective roots with the Proto-Germanic suffix **-njō* will be examined. This termination was almost certainly two suffixes originally, PIE **-(e)n-i̯eh₂*. Therefore, the nouns discussed possibly derived from a case form of an old *n*-stem having the zero grade of both root and suffix by adding to it the PIE suffix **-i̯eh₂* > **-i̯ā* > PGmc. **-jō* to a stem form *CC-n-*. Then, the connecting vowel *u* in ON *syn* – and similarly in *Fjörgyn* – has its origin in a syllabic nasal, PGmc. **sunjō* < PIE **h₁sn̥i̯eh₂*. On the other hand, the remaining two nouns, *banja** and *winja*, were derived without a connecting vowel, as their roots ended in a vowel (after losing a laryngeal), so the nasal did not become syllabic. Lastly, Go. *Saurini* must have been derived from *Saur* through a (hypothetical) weak masculine noun *Saura*, i.e.: *Saur* → **Saura* m. 'the Syrian one', dat. **Saurin*, gen. **Saurins* → *Saurini*. The connecting vowel *i* is a part of the stem suffix *-in-*, a descendant of PIE **-en-*.

The discussion in the present paper focuses on the nouns in the following table:

PGmc.	***banjō***	***winjō***	***sunjō***
Go.	*banja**	*winja*	*sunja*
ON	*ben*	*vin*	*syn*
OE	*benn*		*synn*
OS	[*beni-*]		*sunnia*
OHG		[*winne*]	*sunna*

1. Go. banja* 'sore, wound; πληγή, ἕλκος'

The Gothic noun *banja** has its cognates in ON *ben* and OE *ben(n)* '(fatal) wound'; the latter is confined to poetic language. The same stem is found in OS *beniwunda** 'wound (to the bone)' (Heliand 4879; the wound of Malkus Jh 18:10). PGmc. **banjō* did not survive in Old High German. The ON noun *bani* m. 'death, killer' and the verb *bana* 'to kill, sly', along with their West-Germanic cognates, will be discussed below. Two more related nouns should also be mentioned, OS *banethi*

'death, murder' (Heliand 5484; *binithion* dat. pl. 4865) and ON *bend* 'wound' < PGmc. **baniþō*. Most likely these nouns derived from a PGmc. weak verb, **ban-jōn-* or **ban-jan-* 'to wound', cf. ON *benja* and OE *bennian*.

Gothic *banja** with its cognates could derive from PGmc. **banjō* but the PIE form is more difficult to determine. Actually, it has not been ruled out that this noun is from the specifically Germanic vocabulary, although that observation is of little help. Some scholars tried to connect Go. *banja** to the PIE root **bʰen-* with the meaning 'sicken', cf. Lehmann (1986: 61 B23). According to Pokorny (1959: 126), the meaning of this root was 'to wound', but Magnússon (1989: 40 [*bani*]) lists the meaning as 'to hit', cf. Barr (1971: 41 [*bano*]). It has been questioned whether this root existed in Proto-Indo-European, as it is found only in Avestan. Interestingly, Kroonen (2013: xxviii, 51) has restored this etymology, PGmc. **banjō* < PIE **bʰonih₂*, rejecting Seebold's hypothesis, outlined in the next paragraph.

Seebold (1967: 113–114, 115; 1980: 477–478) tried to connect Go. *banja** with the PIE root **gᵘʰen-* 'to beat, to kill'. Casaretto (2004: 152) agrees with that connection. She refers to *LIV*: 75, where the connection of Go. *banja** with PIE? **bʰen* 'sicken' is rejected but reference is made to the PIE root **gᵘʰen-*. In that location it is pointed out that Lloyd, Springer (1988: 460–462) reject Seebold's (1980: 448) hypothesis that PIE *gᵘʰ* and *gʰᵘ* became *b* in Proto-Germanic unless *u* followed; then the outcome was undoubtedly *g*, e.g. OIc. *gunnr* 'battle'. In the referred location, *LIV*: 218–219 Go. *banja** is not mentioned. Actually, Seebold's hypothesis has older roots, as it could be maintained that Grimm (1852: 82) was the first to propose this idea, even if he was not aware of it himself. He equated Go. *banja** with Gr. *φόνος* and *φονή* 'murder'. Later, Schade (1882: 39 [*bano*]) took up this idea. Indeed, Seebold (1980: 450–451, 465, 476) also gives an overview of the discussions so far of the fate of the PIE voiced, aspirated labiovelars in Germanic. The conclusion that they changed to labials is disputed. Normier (Schmidt, Strunk 1989: 273) and Hartmann (2013) are in favour of a revised version of Seebold's theory. On the other hand, Witczak (2012: 88) concludes that PIE *gᵘʰ* became PGmc. *g* when the accent followed; elsewhere it became PGmc. *w*. Here I will stick to the last hypothesis, as it is more attractive. Thus, as a possible etymology will be suggested for Go. *banja**, ON *ben* etc., the initial *b-* in these and related words is proclaimed to have its origin in PIE *bʰ-*. However, a reliable cognate outside Germanic remains to be found.

The basis of Go. *banja** could be a PIE root of the type **bʰeH-*, zero grade **bʰə-*. In *LIV*: 68–69 there is the PIE root **bʰéh₂-*, zero grade **bʰh₂-*, with the meaning 'shine, light, glow', cf. Pokorny (1959: 104). Let us assume that from this root a noun was formed with the compound suffix **-(e)n-i̯eh₂-*. PGmc. **banjō* could have emerged from PIE **bʰh₂-én-i̯eh₂*. This implies that the noun in question possibly derived from a case form of an old (and lost) *n*-stem that had the zero grade of the root but the full grade of the suffix (cf. Kroonen 2011: 33–34). The laryngeal, having coloured the *e*, was lost, which gives us **banī̯ā* > PGmc. **banjō*. Actually, the form **bʰən-* could have developed from one such case form, i.e. the zero grade of the root and suffix with the accent on the second suffix, **bʰh₂n-i̯éh₂-*. There are more possibilities here with the *e-* or *o*-grade of the root or suffix. The combination *VHR* ultimately gave

a short syllable, so it is possible to reconstruct PIE *$*b^heh_2níh_2$*. In the combination *VH*, the vowel is shortened (or not lengthened) ahead of *___RV́*, e.g. *$*suH-nú-$ > *$*sŭ-nu-$ > Go. *sunus* 'son'; *$*u̯iH-ró-$ > $u̯ĭ-ro-$ > Go. *wair* 'man' (Dybo's Law; cf. Schrijver 1991: 351–356). Thus, when the laryngeals disappeared from PIE *$*b^heh_2níh_2$*, the resultant form entered the class of short *jō*-stems.

According to *LIV*: 69 a homophonous root existed with the meaning 'to speak, to tell'. That meaning is held to have grown out of the former, as both roots are morphologically identical. The meaning development is thought to have been 'shine' → *'make bright' → *'make clear' → 'speak' (German: 'leuchten' → *'hell machen' → *'klar machen' → 'sprechen'). Beekes (2010: 1551–1552 [*φάος*]) also mentions this possibility, as does Mayrhofer (1993: 259–260), under the root *bhā* 'shining, beaming', which is possibly connected with other roots in Sanskrit, *bhan* 'speak' and *bhāṣ* 'tell' (possibly connected with ON *banna* 'forbid' and *bón* 'prayer').

The meaning 'shine, light, glow' appears not to fit very well to 'sore' or 'wound'. But maybe PGmc. **banjō* originally had the meaning 'a clean wound'. In Old Icelandic we have the phrase *að fægja sár* 'to polish a wound'. That which has been polished shines – and 'a clean wound shines', a doctor confirmed to me.[2] This could be connected with the reconstructed meaning variants *'make bright' and *'make clear', which could possibly have been used to describe the action 'to polish a wound'. It is also possible that PGmc. **banjō* was used as a euphemism or taboo for a fatal wound, but ultimately 'sore, wound' became its meaning. OE *benn*, in poetic language, means 'fatal wound', as does *ben* has in Old Norse legal vocabulary.

Now, some more nouns, derived from the same root as **banjō*, should be discussed shortly: ON *bani* 'death, killer', OE *bana* 'killer, murderer, death', OS/OHG *bano* 'killer, murderer, executioner' and OHG f. *bana* (2×) 'death, execution'. Most likely these words did not derive directly from the root *$*b^heh_2$*-. Rather, an *n*-stem was derived from its zero grade with the zero grade of the suffix, i.e. *$*b^hh_2$*-*n*-, although the *e*- or *o*-grade cannot be excluded, i.e. *$*b^hh_2$*-*én*-, *$*b^hh_2$*-*ón*-. Supposedly, all these forms, if they entered Proto-Germanic, should have changed to **ban*-, a form that cannot be divided into root and suffix, and is confined to the Germanic languages. It is possible that **ban-jō* was a Proto-Germanic derivation of this new root, as it has no Proto-Indo-European cognate. The same applies to the weak masculine noun **ban-an*-, which is found in ON *bani* etc. PGmc. **banjō* 'wound' has – at least in Old Norse and Old English – developed the meaning 'fatal wound'. This meaning, one could maintain, appears to be the premise for the derivation of ON *bani*, etc., because the original meaning of the root would give **banan*- the meaning 'the shining one, the clear one'. But if we are ready to accept this as a reference to the deathly pale colour of a corpse, 'the cadaverous one' (i.e. death), no such premise is needed. A further derivate from PGmc. **ban*- is the verb **ban-ōn*- 'to kill, to sly', ON *bana*. The meaning 'death' in ON *bani* involves a personification, so it could rather easily be used as an agent noun with the meaning 'killer'.

[2] I thank my friend, the physician Örn Bjarnason, for useful discussion about wounds and other medical issues that turned up here.

To close this discussion, it should be mentioned that Orel (2003: 35 [**banjō*]) derives **banjō* from **banan-*, but he does not make clear whether he assumes the PIE root to be **gᵘʰen-* 'to beat, to kill' (e.g. Seebold 1980: 451) or not. As was made clear above, the relations with this root are disputed, and if the PIE root **bʰen-* existed, it did not have this meaning, as we saw in the beginning.[3]

Finally, a few words concerning the meaning of Go. *banja**. It occurs three times in the Gothic corpus. In two instances (Luke 16:20, 21, "The Rich Man and Lazarus") clearly open ulcers and boils are spoken of, even with exudation, Gr. *ἕλκος*:

(1) Lk 16:20: *banjo fulls* 'full of ulcers' *ἡλκωμένος*
(2) Lk 16:21: *hundos … bilaigodedun banjos is* 'dogs … were licking his ulcers' *κύνες … ἐπέλειχον τὰ ἕλκη αὐτοῦ*
(3) Lk 10:30 *banjos analag[jandans]* 'inflicting blows' *πληγὰς ἐπιθέντες.*

In the third instance (Luke 10:30, "The Good Samaritan") the Gothic phrase *banjos analag[jandans]* is a translation of the Greek phrase *πληγὰς ἐπιθέντες* ('inflicting blows'), so, at a glance, these are not necessarily (open) wounds but could be bruises. In verse 34, which is lost from the Gothic version, we read *κατέδησεν τὰ τραύματα αυτοῦ* 'bound down his wounds'. So, as only open wounds need to be bound down, one can infer that open wounds are meant. It is possible that Wulfila chose *banja** to stress that understanding. In any case, he did not choose Go. *slahs** 'blow', a noun that is used twice as a rendering of the Gr. *πληγή* (2Cor 6:5, 11:23) but that also has the meaning 'plague; μάστιξ'. It is important and should be stressed that Go. *banja** did **not** have the meaning 'fatal wound'.

2. Go. *winja* 'pasture; νομή'

The Gothic noun *winja* is ἅπαξ λεγόμενον. It is attested solely in John 10:9:

(4) *winja bigitiþ* 'finds pasture' *νομὴν εὑρήσει.*

Winja has a cognate in the Modern Icelandic feminine noun *vin* 'oasis'. In Old High German dictionaries we find *winne** (1×) 'pasture', but it is rare and only found in place names, according to Köbler (1993: 1274 [*winne**]) and Trier (1963: 110). This noun is not found in the other WestGermanic languages. As a matter of fact, in Old Norse, *vin* is only used in place names, e.g. *Björgvin* 'Bergen' etc. (see Jansson 1951), ignoring two obscure compounds occurring in Old Norwegian legal vocabulary, *vin(j)artoddi* (*Ólafs saga hins helga*: 73) and *vinjarspann* (Frostaþingslög, *Norges Gamle Love intil 1387* I: 257).[4] Hence, it is not known how old the meaning 'oasis'

[3] Obviously, if Seebold's hypothesis is accepted, the possibility arises that ON *ben* and *bani* are not derived from the same root: Possibly, then, *ben* etc. is derived from **bʰeh₂-*, but *bani* etc. from **gᵘʰen-*.

[4] Trier (1963: 109) gives *vinjartoddi* the meaning, 'Abgabe für die Nutzung eines Grasplatzes', and compares it with MHG *wunnemiete*. The context is indecisive but this is not a bad guess; 'pasture charge, pasture duty'.

is in Icel. *vin*, but the first occurrences in print are from the first half of the 20th century (Written Language Database at the Árni Magnússon Institute for Icelandic studies). In these cases, the noun is used without any explanation, so, apparently, it was known to the general public. Here PGmc. **winjō* can be reconstructed but further etymology is uncertain.

Pokorny (1959: 1146) connects PGmc. **winjō* with the PIE root **u̯en-* (1), **u̯enə-* 'desire, wish, love, get, win', cf. Magnússon (1989: 1138 [*1 vin*]) and Lehmann (1986: 404 W69). The following words are considered to have derived from the same root as Icel. *vin*: ON *vinr* 'friend', *von* 'hope', and *una* 'be happy'; OHG *wunna* 'joy, desire' and OE *wynn* and OS *wunnia* 'delight, pleasure, desire'. Casaretto (2004: 155–156), on the other hand, expresses some doubts about these relations. Below a new – and better – etymology will be suggested.

The meaning 'grassy spot (in a desert)' appears to be confined to Icel. *vin* but the corresponding nouns in Gothic (and Old High German) simply mean 'pasture'. When it is assumed (e.g. Magnússon 1989: 1138 [*1 vin*]) that this noun originally had the meaning 'likeable place', it is an attempt to unite these meanings. Actually, Icel. *vin* does not have the meaning 'likeable place' in general. The place has to be delimitated and in contrast to its (barren) surroundings, i.e. 'oasis' in a direct or transferred sense.

When OHG *wunna* is given the meaning 'pasture', it is based solely on the fact that in Middle High German there existed the phrase *wunne und weide*, where *wunne* has replaced *winne*, according to Pokorny (1959: 1147). Kluge (1995: 897 [*Wonne*]) has the following remark: "Das deutsche Wort bedeutet auch eine Art Weide, nach Trier die 'Laubweide' (das frische Baumgrün, das von den Tieren besonders gern gefressen wird)." Trier (1963: 79–82) explains that *wunne* was 'pasture on leaf', but *weide* was 'pasture on grass', and he (1963: 95) stresses that *wunne* was not 'Weide'. Trier's (1963: 110) conclusion is that the meaning of the two nouns, *winne* and *wunne*, had become so similar that it was impossible to keep them apart, but the *wunne* form survived. This means that in Old High German, *winne* also had the meaning 'pasture on leaf', and Trier (1963: 117–118) tries to connect this noun with Skt. *vanam* 'tree', an idea found already with Lidén (1903/4: 27), cf. Lehmann (1986: 400 W52). However, the meaning 'pasture on leaf' is found neither in Gothic nor in Old Norse.

Even though HG *Wonne* 'happiness, pleasure' can also be used for what is called *Laubweide* 'pasture on leaf', it has not, thereby, become equal to Icel. *vin*. Rather, the pasture on leaf is called *wunna* in OHG because domestic animals find happiness, pleasure, etc., in it. Nothing said so far involves OHG *wunna* is necessarily being derived from the same root as Icel. *vin*. The meanings Köbler (1993: 1294 [*wunna**]) lists with *wunna* are 'Wonne, Freude, Lust, Glück, Seligkeit, Wollust, Nutzung (?), Ergötzung, Genuss, Vergnügen'. As can be seen, 'Weide' is not amongst them. On the other hand, meaning of delight is quite clear. To stress that, OHG *wunnigarto* 'park, paradise' can be pointed out.

It is a bit romantic to suppose that Go. *winja* – Icel. *win* had the basic meaning 'likeable place' and that they derived from the root **u̯en-*, as did ON *von* 'hope', *vinr* 'friend', *yndi* 'joy' and others. This idea can be found as early as in the middle of

the 19th century. Munch (1849: X–XI) mentions that *vin* is quite common in place names, continuing:

> Det svarer ganske til det gotiske *vinja*, som hos Ulfilas bruges i Betydningen „Græsgang", og til det angelsaxiske *vynn*, det oldtydske *wunna*, hvilke dog kun anvendes i Betydningen „Velbehag", det nyere tydske „Wonne". Ordet har altsaa aabenbart Hensyn til en Tid, da Fædriften udgjørde vore Forfædres fornemste Levevej, og da den ypperste Græsgang tillige var det behageligste Sted. Hvilken af begge Betydninger er den oprindelige, lader sig vanskeligt afgjøre; men tager man Hensyn til Ligheden, og vel derfor ogsaa Slægtskabet med Ordet *vinr ɔ*: „Ven", skulde man dog nok helde til den Mening, at Ordet egentlig har betegnet „Velbehag", „behageligt Sted", og at betydningen „Græsgang" er afledet.[5]

When looking for a new etymology for Go. *winja* and Icel. *vin*, one can stop with the PIE root **u̯ei̯-/*u̯i-*, meaning 'twist, bend, turn', according to Pokorny (1959: 1120–1121), Magnússon (1989: 1114 [*1 veggur*]), and Orel (2003: 441 [**wajjuz*]); **u̯ei̯H-/*u̯iH-*, according to Jasanoff (1978: 84); **u̯i̯eh₁-/*u̯ih₁-* 'to wrap, encase', according to *LIV*: 695; and **u̯h₁i̯-*, according to Kroonen (2013: 568 [**wajju-*]), cf. Casaretto (2004: 199) and Lehmann (1986: 386 W1). So, then, Go. *winja*, Icel. *vin* < PGmc. **winjō* derived from a case form of an old *n*-stem that had a zero grade of the root and the suffix, PIE **u̯ih₁-n-i̯éh₂* – actually the same root as in Go. *-waddjus*, ON *veggr* 'wall' < PGmc. **wajjuz* < PIE **u̯oi̯h₁-us*. Kroonen (2013: 568) assumes PIE **u̯h₁i̯-u-* with a syllabic laryngeal for PGmc. **wajju-*. It is possible that PIE **u̯h₁i̯-n-* > **uh₁i-n-*; if the most vowel-like sounds became syllabic but the less vowel-like remained non-syllabic, then the laryngeal was lost without traces, yielding **u̯i-n-* etc.

Interestingly, Lehmann (1986: 400 W52; cf. Feist 1939: 559–560 [*weipan*]) also chooses the root **u̯ei̯-* with the 'root enlargement' *-b-*, **u̯ei̯-b-*, when he explains the Gothic verb *weipan** 'to crown; στεφανοῦν'. The Gothic nouns *wipja* and *waips**, both meaning 'crown; στέφανος', derived from *weipan**. They are used for the thorny crown (e.g. in John 19:5). The first shows a zero grade but the second shows an *o*-grade. In *LIV*: 671, this is found under the root **u̯ei̯p-* 'get into swinging/shivering movement'. Another variant of the root enlargement, *-p-*, is thought to be present in Go. *biwaibjan** 'surround, clothe' and ON *veifa* 'be in swinging movement'. In *LIV* the words just mentioned are not assumed to be connected with the same root as Go. *-waddjus* and ON *veggr*. Nevertheless, this assumption is tempting, as the meaning resemblance is great. We can assume that the root had a laryngeal, **u̯oi̯h₁-*, to account for the 'Verschärfung' in Go. *-waddjus* and ON *veggr*,

[5] Translation [MSn]: It corresponds exactly to Gothic *winja*, which by Ulfilas is used in the meaning 'pasture', and to the OE *wynn*, the OHG *wunna*, although these are only used in the meaning 'joy', the Modern German 'Wonne'. So, the word has evidently hindsight to the time when the sheep-breeding was our forefather's most preferable livelihood, and when the best pasture also was the most likable place. Which of the two meanings is the original one is difficult to decide; but if one takes notice of the similarity and, therefore, also the relationship with the noun *vinr ɔ*: 'friend', one should indeed cling to the opinion, that the word has literally signified 'joy', 'likable place', and that the meaning 'pasture' is derived.

and with the root enlargement *-p-/-b-*, e.g. in *weipan**, and an *n*-stem with the zero grade in the root and suffix, resulting in PGmc. **wi-n-jō* in Go. *vinja* and Icel. *vin*. It should also be mentioned that PIE **u̯ih₁níh₂* should give PGmc. **winjǭ-*, as the root vowel is shortened or not lengthened in this position (Dybo's Law; cf. Schrijver 1991: 351–356). Thus, the conclusion would be the same whether the root ended in a laryngeal or not.

The PGmc. **wajjuz* 'wall' was supposedly a twisted wall originally. Then, by extension, it came to denote a wall of stone as well (Go. *baurgswaddjus* 'townwall'). The original meaning of PGmc. **winjō* could have been 'delimited area', or perhaps 'enclosure'. It is also possible that in the beginning it denoted a fence, but then its meaning was extended to include the fenced in area. Probably, it also denoted an enclosure in nature, made by rock, etc. Still, it is impossible to say how Icel. *vin* got its modern meaning 'oasis'.

Actually, though, a similar explanation appears for Vries (1961: 664 [*vin*]), whereby he finds it possible that PIE **u̯en-* had the meaning 'fence', and with a dental enlargement it created **u̯endh-*, which is found in ON *vinda* 'to wring', cf. Trier (1963: 113–114). The problem is that the meaning 'fence, wall' is not easily assigned to the root **u̯en-*.

3. Go. *sunja* 'truth; ἀληθεῖα'

There is no disagreement about the origin in this case. Here we have a derivation from the zero grade of the PIE root **h₁es* 'to be'. It appears to be straightforward that PGmc. **sunjō* < PIE **h₁sn̥i̯éh₂* gave us Go. *sunja* 'truth', ON *syn* 'denial, refutation' (also the goddess name *Syn*; in Modern Icelandic, *syn* is only found in the compound *nauðsyn* 'necessity'), OE *synn* 'offence, sin, hostility', OS *sunnia* 'truth, distress, illness' and OHG *sunna* (2×) 'truth, excuse, justification, legitimate hindrance'. However, it is usually maintained that Go. *sunja* etc. derived from the same protoform as OHG *sunta* and OS *sundia* 'sin', i.e. PGmc. **sunðijō-* < PIE **h₁sn̥tíh₂*.

To connect OE *synn* and OHG *sunta*, Kluge (1881: 106) first floated the hypothesis that OE *synn* 'sin' was created from PWGmc. **sundjō*, which had the cluster *-ndj-* in all inflectional forms; thus, *-ndj-* became *-nn-*. On the other hand, the nom. sg. of OE *bend* 'band, fetters' was originally **bandī*, gen. sg. **bandjōs*. In this paradigm, then, *nd+V* and *nd+j* alternated, and ultimately, the alternation created two variants, OE *bend* and *benn*. The latter form could have easily merged with the homophonous OE *benn* 'soar' (< **banjō*; see section 1). The form *benn-* occurs a few times where *bend-* would be expected (see below).[6] A few years later, Kluge (1885: 444, cf. 1901: 379) altered the hypothesis so that he also reconstructed the nom. sg. as PGmc. **sundī*. From the oblique cases in that paradigm we got Go. *sunja* and OE *synn* because of syncope or assimilation of the dental between *n* and *j*. OHG *sunta* (and OS *sundia*), on the other hand, generalised the nominative form of the stem,

[6] Kluge (1881: 106) also claims that inflectional form of OE *benn* 'wound' with *nd* appear to occur. He does not give references as to where.

so the paradigm ultimately ended up as two separate words: OHG *sunta* and *sunna*, OS *sundia* and *sunnia*.[7] Helten (1905) points out that the premise for a change such as the syncopation of *d* between *n* and *j* is that Sievers' Law had ceased to work, i.e. the Proto-Germanic difference between the *jō*-stems and the *ijō*-stems had disappeared: PGmc. **bandī* then had the gen. sg. form **bandjōs* but not **bandijōs*. This must have happened before the difference between the nom. sg. of short and long *jō*-stems disappeared in West Germanic. Otherwise, the stem form of the nom. sg. of the long stems could not be different from the stem form of the oblique cases.

Accepting this involves the conclusion that words showing the cluster *-ndj-* were either created after the change had occurred or owes the cluster to analogy. As an example, Go. *bandi* 'fetters; δεσμός' should have the form **banjos* in the gen. sg. The actual form, *bandjos*, has been influenced by the nominative. This explanation has been long-lived. Thus, Seebold (1969: 20; cf. also 1980: 452) accepts it, but he admits that it is "nicht streng lautgesetzlich", and actually, the word in question, Go. *sunja*, is "das einzig völlig sichere Beispiel für den Schwund eines Dentals zwischen *n* und *j*" (Seebold 1980: 452), cf. Lehmann (1986: 329 [S163]). Casaretto (2004: 440) maintains that Go. *sunja* shows, in the oblique cases, "Dentalschwund vor **-j-*" almost as though this is a regular sound change. She also agrees with the idea that here one paradigm became two. But the meaning 'sin' or the like is absolutely not the meaning of Go. *sunja* or ON *syn*.

Kieckers (1928: 80–81) appears to be the only to express doubt about the explanation just drafted. He mentions the possibility that Go. *sunja* was created by the loss of a dental between *n* and *j* but considers this uncertain. He points to Go. *sandjan* 'to send' as an exception that has to be younger than the loss of the dental, but such exceptions would be numerous. For example, *nehʋundja* 'fellowman' and *þusundi* 'thousand' should have lost the *d* in the oblique cases, and actually, it should have been lost from the whole paradigm of the former word. Kieckers's point is important. The traditional explanation of Go. *sunja* has the effect that forms such as *sandjan* 'send' and *bandjos* need an explanation, which they do need not if the theory on the *-ndj-* cluster is abandoned, i.e. it is necessary to explain why there are still word forms with the cluster *-ndj-*; to assume a loss from PGmc. **sunjō* only would be hard to support. Here the desire to connect Go. *sunja* with Skt. *satya-* 'true' and OHG *sunta* has been decisive.

High German appears to have lost the noun *sunna*, perhaps because it overshadowed 'the sun', OHG *sunna*. Similarly, the process *syn* → *synja* 'to deny' → *synjun* 'denial' in Icelandic has rendered *syn* superfluous; *synjun* has taken its place. OE *synn* developed the meaning 'sin' (and lives forth in English *sin*). Thereby, the Old English cognate of OS *sundia* and OHG *sunta* became superfluous and was subsequently lost.

The possibility should not be excluded that Proto-Germanic contained both forms, **sunjō* < PIE **h_1s-n̥-i̯éh$_2$* and **sunðī* (obl. **sunðijō-*) < PIE **h_1s-n̥t-íh$_2$*. The former

[7] Brugmann (1897: 707–708) mentions some examples that are meant to support this but they are hard to accept.

could have derived from an old *n*-stem with a zero grade of the root and suffix, just like **banjō* and **winjō*. The latter is an old present participle that did not enter into Gothic or North-Germanic (as ON *synd* 'sin' is generally considered a loan word). The West-Germanic forms with *-nn-* have the geminate as a result of the gemination before *-j-* (Brunner 1965: 186–187; Braune, Eggers 1975: 94–98; Gallée 1910: 157, 182). Therefore, there is no reason to assume a loss of a dental or its assimilation to *n* to account for them. Thus, Campbell (1959: 237) says that OE *synn* comes from **sun̥jō*. Brunner (1965: 159) mentions only OE *synne* 'sin' as an example when he discusses the loss of *d*. Sometimes *d* also drops from the ending of the present participle (Brunner 1965: 279). Brunner refers, on this connection, to Malone (1929), who only discusses the sporadic loss of *d* finally and in a syllable with secondary stress. On the other hand, Campbell (1959: 196) says that the change *-nd-* > *-nn-* occured in Old English only four times in the forms of the noun *bend* 'fetters', i.e. nom. pl. *benne* (1×) and dat. pl. *bennum* (3×). It should be mentioned that in Bessinger (1978: 111) the first occurrence (Daniel 434) has been corrected to *bende*, but the rest has benn kept without conjecture (Genesis 1972; Andreas 962, 1038; Juliana 519). This implies that the meaning 'wound' is not considered impossible in these cases, although the meaning 'fetters' would be much more likely. See also Wülker (1888: 49, note on Andreas 1038.)

The main thing is that there is no need for the traditional explanation. In the Old High German form *sunna*, the *-nn-* could very well be a regular offspring of PGmc. *-nj-* in **sunjō*. According to Köbler (1993: 1050 [*sunna* (1)]), OHG *sunna* had the meanings 'Wahrheit, Entschuldigung, Rechtfertigung, gesetzlich anerkannter Hinderungsgrund'. This is in good agreement with Go. *sunja* 'truth' and ON *syn*. Also according to Köbler (1993: 1050 [*sunta* (1)]), OHG *sunta* had the meanings 'Sünde, Vergehen, Laster, Schuld, Schandtat'. Clearly, the meanings of the two nouns do not overlap, as would be expected if they had grown out of the same paradigm. Summing up, Go. *sunja*, ON *syn*, OE *synn*, OS *sunnia* and OHG *sunna* most likely derived from PGmc. **sunjō*. On the other hand, OS *sundia* and OHG *sunta* (older *suntea*) are descendants of WGmc. **sunðijō-*, which is absent in other Germanic languages. However, it was, most likely, borrowed into Old Norse.

4. Conclusion

The conclusion is that all three nouns – Go. *banja**, *winja,* and *sunja* – could have derived from a case form of an old *n*-stem with a zero grade of both the root and the suffix. To this the additional suffix **-jō* was added. The basis of **banja* is probably the same, but an *e-* or *o*-grade would give the same result: PIE **$b^{h}h_{2}$-n-* or **$b^{h}h_{2}$-en-* or **$b^{h}h_{2}$-on-* > PGmc. **ban-*. Also, **ban-* became the root of ON *bani* and its West-Germanic cognates (and the ON verb *bana*). As mentioned in the beginning, the suffix **-n̥ieh$_2$* was originally a combination of two forms. Perhaps they never melted into a whole. The roots of *banja**, *winja* and *sunja* are found outside Germanic, but probably not these specific derivations.

Go. *winja* and Icel. *vin* are derived from the same root as PGmc. **wajjuz* ‘wall’.

Then, OHG *sunta* and OS *sundia* descended from PGmc. **sunðijō-*. In Old High German this noun has also taken the nominative ending of the *ō*-stems; the regular form would be *sunte*. The derivation of Go. *sunja*, ON *syn*, OE *synn(e)*, OS *sunnia* and OHG *sunne* from PGmc. **sunðijō-* lacks evidence. The geminate *-nn-* in the West-Germanic forms is the result of a well-established West-Germanic rule. Also, the Gothic and Old Norse forms are best understood as descendants of PGmc. **sunjō*.

The derivation and development of the nouns discussed here can be said to be regular; the outcome is in accordance with well-established rules and principles connecting Proto-Indo-European and Proto-Germanic.

References

Barr S. 1971. *An etymological glossary to the Old Saxon Heliand*. Berne.

Beekes R.S.P. 2010. *Etymological dictionary of Greek*. [vol. 1–2]. Leiden.

Bessinger J.B. (ed.). 1978. *A concordance to the Anglo-Saxon poetic rercords*. Ithaca, London.

Braune W., Eggers H. 1975. *“Althochdeutsche Grammatik” von W. Braune*. [13. Auflage bearbeitet von H. Eggers]. Tübingen.

Brugmann K. 1897. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen* […]. [1: Einleitung und Lautlehre. Zweite Bearbeitung]. Strassburg.

Brunner K. 1965. *Altenglische Grammatik*. [Dritte, neubearbeitete Auflage]. Tübingen.

Campbell A. 1959/2003. *Old English Grammar*. Oxford.

Casaretto A. 2004. *Nominale Wortbildung der gotischen Sprache. Die Derivation der Substantive*. Heidelberg.

Chantraine P. 1979. *La formation des noms en Grec ancien*. Paris.

Feist 1939 = Lehmann 1986.

Gallée J.H. 1910. *Altsächsische Grammatik*. [Zweite völlig umgearbeitete Auflage]. Halle, Leiden.

Grimm J. 1852. SCADO. – *Zeitschrift für vergleichende Sprachforschung* 1: 79–83.

Hartmann M. 2013. Anlautendes idg. **g^wh^*- und **(ĝ)ĝ^h^u̯*- im Germanischen. – *International Journal of Diachronic Linguistics and Lingusitic Reconstruction* 10: 1–25.

van Helten W. 1905. Grammatisches. LXVIII. Zu germ. *-ni̯-* (woraus *-nni̯-*) aus *-n* + *t* (oder daraus entstandenen dental) + *i̯*-. – *Beiträge zur Geschichte der deutschen Sprache und Literatur* 30: 248–251.

Jansson V. 1951. *Nordiska* vin-*namn. En ortnamnstyp og dess historia. A.-B.* Uppsala, København.

Jasanoff J. 1978. Observations on the Germanic Verschärfung. – *Münchener Studien zur Sprachwissenschaft* 37: 77–90.

Kieckers E. 1928. *Handbuch der vergleichenden gotischen Grammatik*. München.

Kluge F. 1881. Anglosaxonica. – *Anglia* 4: 105–106.

Kluge F. 1885. Sprachhistorische Miscellen. 12. Etymologien. – *Beiträge zur Gesichte der deutschen Sprache und Literatur* 10: 443–445.

Kluge F. 1901. 2. Vorgeschichte der altgermanischen Dialekte. – Paul H. (ed.). *Grundriss der germanischen Philologie*. [vol. 1; Zweite verbesserte und vermehrte Auflage]. Strassburg: 320–496.

Kluge F. 1995. *Etymologisches Wörterbuch der deutschen Sprache.* [Bearbeitet von E. Seebold; 23, erweiterte Auflage]. Berlin, New York.
Köbler G. 1993. *Wörterbuch des althochdeutschen Sprachschatzes*. Paderborn, München, Wien, Zürich.
Krahe H., Meid W. 1967. *Germanische Sprachwissenschaft*. [vol. 3]. [Wortbildungslehre von W. Meid]. Berlin.
Kroonen G. 2011. *The Proto-Germanic* n*-stems. A study in diachronic morphophonology*. Amsterdam, New York.
Kroonen Gus. 2013. *Etymological dictionary of Proto-Germanic*. Leiden.
Lehmann W.P. 1986. *A Gothic etymological dictionary*. [Based on the third edition of Vergleichendes Wörterbuch der Gotischen Sprache by Sigmund Feist]. Leiden.
Lidén E. 1903/4. Blandade språkhistoriska Bidrag I. 2 Slav. *lěsŭ* – feng. *læs*. – *Göteborgs Högskolas Årsskrift* 10.1: 25–31.
LIV = Lexikon der indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen. 2001. [Zweite, erweiterte und verbesserte Auflage, bearbeitet von M. Kümmel und H. Rix]. Wiesbaden.
Lloyd A.L., Springer O. 1988. *Etymologisches Wörterbuch des Althochdeutschen I*. Göttingen.
Magnússon Á.B. 1989. *Íslensk orðsifjabók*. Reykjavík.
Malone K. 1929. Notes on *Beowulf*. – *Anglia* 53: 335–336.
Mayrhofer M. 1993. *Etymologisches Wörterbuch des Altindoarischen*. [vol. 2]. Heidelberg.
Munch P.A. 1849. *Historisk-geographisk Beskrivelse over Kongeriget Norge (Noregsveldi) i Middelalderen*. Moss.
Norges Gamle Love intil 1387 […]. 1846. [Utgivne ved R. Keyser og P.A. Munch. Förste Bind]. Christiania.
Ólafs saga hins helga […]. 1922. [Ved O.A. Johnsen]. Christiania.
Orel V. 2003. *A handbook of Germanic etymology*. Leiden, Boston.
Pokorny J. 1959. *Indogermanisches etymologisches Wörterbuch*. [1. Band]. Bern, München.
Schade O. 1882. *Altdeutsches Wörterbuch*. [vol. 1]. [Zweite umgearbeitete und vermehrte Auflage (Neudruck)]. Halle a. d. S.
Schmidt K. T. & Strunk K. 1989. Toch. B *kwĭpe* "Scham; Schande", A *kip* "Scham" und Germ. **wība*- "Weib". K. Heller, O. Panagl und J. Tischler (ed.): *Indogermanica Europaea*. Festschrift für W. Meid zum 60. Geburtstag am 12. 11. 1989, pp. 251–284. Graz.
Schrijver P. 1991. *The reflexes of the Proto-Indo-European laryngeals in Latin*. Amsterdam, Atlanta GA.
Seebold E. 1967. Die Vertretung idg. $g^{u}h$ im Germanischen. – *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 81: 104–133.
Seebold E. 1969. Germanisch **sanþ-/sund-*, seiend, wahr'. – *Die Sprache* 15: 14–45.
Seebold E. 1980. Etymologie und Lautgesetz. Materialsammlung: Die Vertretung von anlautend idg. $g^{u}h$ und $ghu̯$ im Germanischen. – Mayrhofer M., Peters M., Pfeiffer O.E. (eds.). *Lautgeschichte und Etymologie*. [Akten der VI. Fachtagung der Indogermanischen Gesellschaft, Wien, 24.–29. September 1978]. Wiesbaden: 431–484.
Trier. J. 1963. *Venus*: *Etymologien um das Futterlaub*. Köln, Graz.
Vries J. de. 1961. *Altnordisches etymologisches Wörterbuch*. Leiden.
Witczak K.T. 2012. Indo-European *$g^{u}h$ in Germanic. – *Lingua Posnaniensis* 54.2: 83–90.
Wülker R.P. 1888. *Bibliothek der angelsächsischen Poesie*. [Begründet von Ch.W.M. Grein, neubearbeitet […] von R.P. Wülker. II. Band. – 1. Hälfte]. Kassel.

Studia Linguistica Universitatis Iagellonicae Cracoviensis 133 (2016): 109–114
doi:10.4467/20834624SL.16.008.5154

LAURA STURM
Ludwig-Maximilians-Universität München
sturm.laura@campus.lmu.de

EXPRESSIVENESS AND VARIATION: THE ETYMOLOGY OF GERM. *KLADDER* 'DIRT, MUD'[1]

Keywords: Germanic, etymology, expressive germination, *littera*-rule, *kladder*

Abstract

Although the Germanc dialects offer very ancient vocabulary, the have long been neglected from an etymological perspective. A very old word is e.g. Germ. *Kladder* 'dirt, mud'. Because of its onomatopoetic nature this word shows a considerable diversification and expansion in the Germanic languages: *klatt-* and *klāt-* in Low German, Middle German, Upper German next to *kladd-* only in Low German. Those words ultimately go back to a Proto-Germanic substantive **kladdō* f. 'clot, lump, mud, dirt', leading to the well-known PIE root **gleh*$_1$- 'to be greasy, to be dirty'.

(1) Modern German dialectology is mainly focused on the sociolinguistic aspects and language geography of the German dialects. Although these dialects offer very ancient vocabulary, they have long been neglected from an etymological perspective. In this article, I will demonstrate the considerable diversification and expansion of an onomatopoetic dialectal word.

(2) In many Western and Northern Germanic languages words belonging to the semantic field of "mud, dirt" show expressive variations, especially gemination and lenition, and are very often widely attested. In German those variations can most clearly be found in the dialects. One of these expressive words is *Kladder* 'dirt, mud', which appears in almost all German dialects, with most attestations and phonetic variations displayed in the Low German dialects. There we find the noun *Kladder*

[1] I am grateful for the proofreading by Dr. Maria Kozianka and Lukas Kahl. All remaining mistakes are my own fault.

'dirt, mud', its denominal verb *kladdern* 'to do something untidily and messily' and the adjective *kladderig* 'dirty', which is derived from that noun. Besides those inherited formations with the suffix *-er,* the Low German dialects also have words with other suffixes: Holst. *Kladde* f. 'mud', *kladdig* 'muddy', Pruss., Westphal., Low Sax. *Kladde* f. 'paper for exercise, notebook', Westphal., Low Sax. *kladden* 'to work untidily'. All suffixal derivations of the root *kladd-* also have version with the phonetic long vowel root *klāt-* and the devoiced *klatt-*, but with similar semantics. By contrast, the Upper German dialects mainly show the root *klatt-* instead of *klāt-* and *kladd-*: Bavar., Bad., Swab. *Klatter* f. 'dirt, dung, mud', Bad., Swab. *klattrig* 'dirty'. Only Tyrol. *Klāte* f. 'dirt, dung, mud' has a long vowel variant. Although the Middle German dialects also exhibit a huge number of Low German items in the vocabulary, the presence of the words under discussion is very similar to their attestations in the Upper German dialects, except the fact that the long vowel root appears more often: Thur., Upper Sax., Siles. *Klāter* 'dirt, mud', *klāterig* 'dirty' vs. Rhin. *Klatter* 'mud', *Klatterich* 'soft dung', *klatterig* 'dirty'. Only in the dialect of Luxembourg we find the root *Kladd-* with lenition (widespread in the Low German dialects): *Kladder* 'mud at the back of a cow's leg', *kladdereg* 'dirty'. But those Luxembourgian words only seem similar at first sight: the Low German attestations with geminate *-dd-* represent an inherited Westgermanic root, whereas the Luxembourgian words reflect the Middle German lenition which affects the the inherited Old High German Tenues from early Middle High German onwards ("binnendeutsche Konsonantenschwächung", cf. Schirmunski 2010: 392). In some of these dialects, e.g. in Thuringian, the adjectival formations develop a secondary meaning 'desperate, miserable, poor'. Such a semantic change is often found in adjectives meaning 'mud', e.g. Lat. *sordidus* 'dirty, low, miserable'. *Kladd/tt/t/er* is also frequently attested in composition with the adjective Germ. *nass* 'wet' (i.e. *kladdernass*, *klatternat*), forming a determinative compound with a comparative meaning, like hg. *aalglatt* 'sleak as an eel'.

To sum up, there are three root variants in the modern German dialects:

1. *Klatt-*: Low German, Middle German, Upper German.
2. *Klāt-*: Low German, Middle German, Upper German.
3. *Kladd-*: only Low German.

The Low German dialects not only show all three roots but also most attestations of the words.

(3) This situation is also found in the older Low German languages. In Middle Low German only *klatt-* and *kladd-* exist, and while the root *kladd-* is only verbally represented (Mlg. *klad(d)eren* 'to work untidy, to grease', *kladden* 'to besmear', verb-initial compounds *Kladde-hans*, *Kladde-hack* 'mucky pup'), the word *klatt-* appears in nominal as well as verbal guise (*klatte* f. 'smudge', *klatten* 'to stick together, to felt', *klattich* 'stuck together, felted'). Like Middle Low German, Dutch only shows *klatt-* and *kladd-*: Mdu. *cladde* f. 'mud, dirt', *cladden* 'to clean, to expurgate, to brush', *cladder* 'clothbrush', *clatte* f. 'mud, dirt, lump', *clatten* 'to besmear', Ndu. *klad* n., *kladde* f. 'scratchpaper', *klad(de)* f. 'lump, clot, dirt, mud', *klat(te)* f. 'id.', *kla(t)ter* m., f. 'id.',

klat(t)eren 'to besmear'. Other than that, the only West Germanic correspondences are found in Nwestfris. *kladde* m., f. 'scratchpaper' and *kladderje* 'to besmear'. Nengl. *clat* 'lump, dirt, clot' is not inherited, but rather a Dutch loan. Likewise Nnorw., Nswed. *kladd* and Ndan. *kladde* ‚scratchbook; lump' are borrowed from Dutch (cf. Lühr 1988: 280). Beyond the afore mentioned Dutch borrowings, the modern North Germanic languages only a scant few verbal representatives of the root: Nnorw. dial. *kladda* 'to stick together', Nswed. *kladda* 'to besmear, to blot, to scamp'.

(4) At the first sight it is quite striking that we do not have any older High German correspondences except the High German dialectal attestations. It underscores the vital importance of further etymological study of the German dialects – a field which for long has been underrepresented. Secondly, the distribution of the Germanic cognates exhibits the frequently encountered concentration of relevant items in the West Germanic branch. Beside these general considerations the West and North Germanic cognates show that the root contains geminate *dd*. Taking into account the nhg. *t*-words, a Proto-Germanic substantive **kladđō* f. 'clot, lump, mud, dirt' can be reconstructed, which is the derivational basis for the weak verb PG **kladđōn* 'to besmear, to smudge'. The Low German and Dutch words containing *tt* instead of *dd* clearly represent expressive devoicing, leading to **klattō* and **klattōn* (cf. Lühr 1988: 280). Such variations between *tt* and *dd* can also be found elsewhere in the West Germanic languages, especially in onomatopoetic words, e.g. Upper German *pflattern*, *pflättern* and *pfläddern*, all three mean 'to defecate' (cf. Neri, Ziegler 2012: 217). Even outside onomatopoetics such variations can be found. This is the case e.g. with Mlg. *pitte* 'kernel, core, strength' vs. Mlg. *peddik* 'id.' (cf. Lühr 1988: 283).

Much more difficult is the explanation of the Lg., Hg. and Du. root *klāt-*, whose *t* precludes common origin, since the High German dialects must have undergone the High German Consonant Shift. It is more probable that the Low German and Dutch words with that root have a different provenance than the High German cognates with the long root vowel. The Low German and Dutch examples might, for instance, be a result of the Low German and Dutch alternation between a syllable with the structure *VCC* and and one with the syllable *VR̥C*, a pattern known in Latin as *littera*-rule and exemplified by e.g. Lat. *Iūpiter* next to *Iuppiter*, *littera* vs. *līter*a (cf. Meiser 1998/2010: § 57: 5). In Low German and Middle Dutch we find e.g. Mdu. *vergraamen* next to *vergrammen* 'to get angry' and Mlg. *doder*, *duder* next to *dodder* m. 'yolk' (cf. Franck 1910: § 93). Thus Lg., Du. *klāt-* might be merely an allophonic variant of *klatt-*. Since such phonetic alternations are not attested in High German, the only possible explanation of the High German root *klāt-* is the following: it reflects an inherited ablaut variant PG **klēđ-*, which would thus be a part of an ablauting nominal paradigm **klēđ-/klađ-*. The phonetic variation gave rise to two different nouns with **klēđ-* on the one hand and **klađ,-* on the other. Such a paradigmatic split frequently appears among the Proto-Germanic *n*-stems, where the genitive often served as basis for further derivation, e.g. Mhg. *vinc* m. 'spark' next to Mhg. *vunce* m. 'id.' (< **finkō/funkaz*; cf. Kroonen 2011: 58ff., 159f.).

In light of an ablauting paradigm of this sort, the PG geminated noun **kladđđō* f. 'clot, lump, mud, dirt' must be a Proto-Germanic expressive gemination of the weak stem **klađ-*, which without germination is only attested in the verb Mlg. *kladeren* 'to besmear' (besides homonymous *kladderen*).[2]

(5) The diachronic development can be summarized as follows:
1. Proto-Germanic: inherited ablauting paradigm **klēđ-ō/klađ-*
2. Proto-Germanic: paradigm split
   Paradigm 1: **klēđ-ō* and **klēđ-a-* (~ Tyrol. *Klāte*, Thur., Upper Sax., Siles. *Klāter, klāterig*, Lg. *klāt-*)
   Paradigm 2: **klađ-a-* (~ Mlg. *kladeren*) next to expressively geminated **kladđđō* (~ Lg., Du., Swed., Nnorw. *kladd-*).
3. Old Low German: allophonic *klatt-* vs. *kladd.-*

(6) The Proto-Germanic ablauting paradigm ultimately goes back to a Pre-Proto-Germanic proterodynamic stem with suffix *-eh₂*: PPG **gléh₁dʰ-h₂/glə₁dʰ-éh₂*. For this type of ablaut, cf., for example, PIE **gʷén-h₂/gʷn-éh₂* 'feminity', cf. Beekes (1990: 225) and Harðarson (2014: 23). The root **gléh₁dʰ-* shows a dental root enlargement, which commonly creates root doublets [e.g. Ohg. *glizzan* 'shine' vs. Ohg. *glîmo* 'firefly' ~ PG **glit-* vs. **glī-*; cf. Krisch (1990: 117f.)]. The inherited root, without enlargement, PIE **gleh₁-* might be connected with the Slavonic words: Russ. *želvák* m. 'ulcer', Cz. *žluva* 'ulcer (with horses)' and Sloven. *žę̑łva* f. 'fistula'. The Slavonic nouns are all secondary derivatives of a Proto-Slavonic *u*-stem **gelh₁-u-*. They also presuppose schwebeablaut in the root (hence **gelh₁-*). In Latin there might be a cognate of the root in Lat. *galla* f. 'gallnut'. This connection presupposes a laryngeal dissimilation in the *neh₂*-collective **gₔlh₁-neh₂* > **gₔl-neh₂* (a secondary collective to a *no*-adjective) because otherwise we would expect Lat. **galana* and not *galla*. Taken together, the Latin and Slavonic words point to a basic meaning 'clot, agglomeration' rather than 'dirt, mud'. This semantic difference militates against an etymological connection with the Germanic paradigm. Semantically, the correlation is more likely with Gr. *γλάμων* 'bleary-eyed'. The derived noun *γλάμη* 'eyegum' (only attested at Phot. *lex.* 121) appears as a loan also in Lat. *glamae* 'id.', which is also the derivational basis for the postclassical adjective *glarāns* 'bleary-eyed' [corrupted, instead of **glamāns* (only once attested at Garg. Mart. med. 15 p. 148, 10 im 3. Jh. n. Chr. = Plin. Val. 4, 4; W/H 605)]. The Greek noun continues a PIE amphidynamic *men*-stem **gléh₁-mon-/glə₁-mn-ˊ* 'greasy, lubricious' with generalization of the zero grade root and the *o*-grade suffix (Pre-Proto-Greek **glə₁-mon-*). From *men*-stem a secondary Pie abstract noun **gléh₁-mn-eh₂* 'sliminess' is derived similarly to Lith. *glẽmės* pl. 'slim', Latv. *glẽma* 'id.' [with simplification of the cluster *-mn-* to *-m-*; cf. Matasović (2004: 126)]. The coexistence of the Latvian *ā*-stem and the Lithuanian *ē*-stem proves the Lithuanian stem as secondary, since there is the tendency in Lithuanian to systematically replace with productive *ē*-stems the inherited *ā*-formations, e.g. Lith.

[2] Mhg. *bekletzen* 'to besmear' does not belong to the PG Paradigm, cf. Lühr (1988: 280).

garbẽ next to *garbà* 'honor', cf. Lühr (1999: 304). The short vowel of the Baltic root results from laryngeal deletion before a consonant cluster according to Schmidt's Law, cf. Hackstein (2002: 1ff.) and Neri (2011: 292f.). Finally, this *men*-stem is also attested in the Oengl. noun *clām* 'mud, dirt', which goes back to a secondary thematic PPG **gléh$_1$-mn-o-*. In conclusion, the Germanic and the other Indo-European words are ultimately continuants of the PIE root **gleh$_1$-* 'to be greasy, to be dirty', only attested in nominal form.

Abbreviations

Bad.	=	Badian	Nengl.	=	New English
Bavar.	=	Bavarian	Nhg.	=	New High German
Cz.	=	Czech	Nnorw.	=	New Norwegian
dial.	=	dialectal	Nswed.	=	New Swedish
Du.	=	Dutch	Ohg.	=	Old High German
Germ.	=	German	PG	=	Proto-Germanic
Gr.	=	Ancient Greek	PIE	=	Proto-Indo-European
Hg.	=	High German	PPG	=	Prae-Proto-Germanic
Holst.	=	Northern Low Saxon	Pruss.	=	Prussian
Lat.	=	Latin	Rhin.	=	Rhinelandic
Latv.	=	Latvian	Russ.	=	Russian
Lg.	=	Low German	Siles.	=	Silesian
Lith.	=	Lithuanian	Sloven.	=	Slovenian
Low Sax.	=	Low Saxon	Swab.	=	Swabian
Mdu.	=	Middle Dutch	Swed.	=	Swedish
Mhg.	=	Middle High German	Thur.	=	Thuringian
Mlg.	=	Middle Low German	Tyrol.	=	Tyrolian
Ndan.	=	New Danish	Upper Sax.	=	Upper Saxon
Ndu.	=	New Dutch	Westphal.	=	Westphalian

References

Beekes R. 1990. *Vergelijkende taalwetenschap. Tussen Sanskrit en Nederlands.* Utrecht.

Franck J. 1910. *Mittelniederländische Grammatik mit Lesestücken und Glossar.* [2nd edition]. Leipzig.

Hackstein O. 2002. Uridg. *CH.CC > *C.CC. – *HS* 115: 1–22.

Harðarson J.A. 2014. Das andere Wort für ‚Frau' im Urindogermanischen. – Neri S., Schuhmann R. (eds.). *Studies on the collective and feminine in Indo-European from a diachronic and typological Perspective*: 23–56.

Krisch T. 1990. Zur Etymologie von nhd. *gleiten* (mit Bemerkungen zu weiteren Etymologien). – *HS* 103: 116–131.

Kroonen G. 2011. *The Proto-Germanic n-stems in diachronic morphophonology.* Amsterdam.

Lühr R. 1988. *Expressivität und Lautgesetz im Germanischen.* Heidelberg.

Lühr R. 1999. Zum „Sekundärwortschatz" des Baltischen: Alte und neue feminine Abstraktbildungen. – Eggers E. et al. (eds.). *Florilegium Linguisticum*. [Festschrift für Wolfgang P. Schmid zum 70. Geburtstag]. Frankfurt a. M.: 299–312.

Matasović R. 2004. *Gender in Indo-European*. Heidelberg.
Meiser G. 1998/2010. *Historische Laut- und Formenlehre der lateinischen Sprache*. [Unchanged reprint of the 2nd edition 2006]. Darmstadt.
Neri S. 2011. *Wetter. Etymologie und Lautgesetz*. [Unpublished dissertation].
Neri S., Ziegler S. 2012. *„Horde Nöss". Etymologische Studien zu den Thüringer Dialekten*. Bremen.
Schirmunski V.M. 2010. *Deutsche Mundartkunde. Vergleichende Laut- und Formenlehre der deutschen Mundarten*. [Ed. by Larissa Naiditsch and Peter Wiesinger. Transl. by Wolfgang Fleischer]. Frankfurt a. M.

Studia Linguistica Universitatis Iagellonicae Cracoviensis 133 (2016): 115–123
doi:10.4467/20834624SL.16.009.5155

RANKO MATASOVIĆ
University of Zagreb
rmatasov@ffzg.hr

THE ACCENTUATION OF BALTO-SLAVIC *VṛDDHI* FORMATIONS AND THE ORIGIN OF THE ACUTE[1]

Keywords: Balto-Slavic *Vṛddhi*, Balto-Slavic accentuation, laryngeals, lengthened grade, Indo-European

Abstract

There is still no scholarly consensus about the origin of the Balto-Slavic intonations. The traditional view is that all long vowels and diphthongs receive the acute in Balto-Slavic, while short vowels and diphthongs are circumflexed. On the other hand, according to the Leiden school, the only source of the Balto-Slavic acute is the glottal stop, which is either a reflex of the PIE laryngeals, or of the following glottalized stops (traditional voiced stops) in syllables that underwent Winter's law. We believe that the traditional view that PIE lengthened grade vowels receive the acute in Balto-Slavic can no longer be defended. It is contradicted by such examples as PIE **dʰugh₂tēr* 'daughter' > Lith. *duktẽ*, PIE **(H)rēk-s-o-m* 'I said' > Croat. *rijêh*, PIE **h₂ōwyom* 'egg' > Croat. *jâje*. It should also be taken as proved that syllables closed by laryngeals and voiced stops (or glottalics, by Winter's law) received the acute intonation in Balto-Slavic. However, the fact that the PIE lengthened grade long vowels are circumflex in Balto-Slavic does not prove that all lenghtened grade long vowels in Balto-Slavic are circumflex. In the present paper we attempt to show that a number of *Vṛddhi* formations, that were not inherited from PIE, received the acute in Balto-Slavic. These are the words with reflexes in both Baltic and Slavic languages, derived from PIE roots by means of *Vṛddhi*, which remained a productive pattern of derivation during the period of Balto-Slavic unity, and probably later. Such words have the lengthened grade only in Balto-Slavic, but not in other IE languages, which shows that their *Vṛddhi* is not inherited from PIE. This paper systematically analyzes such material in order to show that the Balto-Slavic *Vṛddhi* formations, in contradistinction to the inherited PIE long vowels, received the acute intonation.

[1] This paper was read at the conference of the Indogermanische Gesellschaft in Leiden, in July 2013. Thanks go to Tijmen Pronk, Frederik Kortlandt and Sasha Lubotsky for their comments.

1. Theories of the origin of Balto-Slavic acute

There are three current theories about the origin of the Balto-Slavic acute intonation:

1. The traditional theory: acute developed on all long vowels, whether apophonic or secondarily lenghtened after the loss of laryngeals (e.g. Carlton 1991).
2. The Leiden school approach (Derksen 1996; Kortlandt 2011; Pronk 2012): the acute developed from glottalization, which in turn occurred on vowels preceding laryngeals and voiced (< glottalized) stops (by Winter's law). All apophonic long vowels are circumflexed.
3. Villanueva-Svensson's theory: apophonic long vowels are circumflexed in non-initial syllables and monosyllables, but acuted in initial syllables. Vowels lengthened by the loss of laryngeals are acuted (Villanueva-Svensson 2011).
4. The present writer's opinion (Matasović 2008): the Leiden school is correct with respect to PIE lengthened grade vowels, which are circumflexed. However, new, morphologically derived lengthened grades in Balto-Slavic receive the acute. This is the 'Balto-Slavic *Vr̥ddhi*', to which this paper is dedicated.

2. *Vr̥ddhi* in PIE?

Vr̥ddhi is a formation of denominal adjectives by lengthening the root syllable. The process is best attested in Indo-Iranian, especially in Sanskrit, cf. Skt. *sákhi-* 'follower, friend' vs. *sākhyá-* 'society', *víś-* 'village, settlement' vs. *váiśya-* 'member of the vaiśya caste', *mā́nuṣa-* 'connected with men' vs. *mánuṣa-* 'man'.

It is unclear whether Indo-Iranian *Vr̥ddhi* is a process inherited from Proto-Indo-European, or it developed in the already differentiated Indo-Iranian proto-language (or a group of Indo-European dialects, to which Indo-Iranian belonged). The existence of *Vr̥ddhi* in PIE is rather disputed. If it existed as a derivational process, it was certainly rare. A possible PIE instance of *Vr̥ddhi* is the word for 'egg':

PIE *$h_2$ewi-* 'bird' > Lat. *avis,* Skt. *ví-*: PIE *h_2ōwyom* ('that of the bird, bird's' > 'egg') > Lat. *ōvum,* OHG *ei,* ON *egg* (< PGerm. **ajjaz-* n), Croat. *jâje* (Novi), Pol. obs. *jajo, jaje*, ULus. *jejo* < PSl. **âje*, (AP c), Derk. 27, ESSJa I: 61–2). Alternative reconstruction of the word for 'egg' may be PIE *h_2oh_2w(y)o-*, with the first reduplicated syllable, as in PIE *h_1oh_1k'u-* 'quick' vs. **h1ek'wo-* 'horse', *k^wek^{w}lo-* 'wheel' (< *'turning') from *k^wel(H)-* 'turn'. If this is accepted, there is no need to posit a *Vr̥ddhi* formation.[2]

3. Balto-Slavic *Vr̥ddhi*?

Just as it is uncertain whether *Vr̥ddhi* existed in PIE, it is at present unclear whether it should be posited in Balto-Slavic. Since Balto-Slavic shares a number of isoglosses with Indo-Iranian (e.g. the operation of RUKI-rule and the satemization of

[2] For possible other instances of *Vr̥ddhi* in PIE and Germanic see Darms 1978.

palatalized velars), this question cannot be answered without a careful examination of the available evidence. Collection of (mostly implausible) examples of the derivational lengthened grade in Slavic can be found in Gołąb (1967). More probable cases have been collected below, and they share three defining features: firstly, they contain either an acuted long vowel, or an acuted diphthong, and are derived from roots that do not end in laryngeal. Secondly, they usually have a derived meaning with respect to the meaning of the base noun. Their meaning is usually possessive, relational, or collective: 'belonging to X', 'pertaining to X', 'descending from X', or 'a collection of X'es', where X is the base noun. Finally, formations with "Balto-Slavic *Vr̥ddhi*" usually do not show evidence for lengthened grade except in Balto-Slavic. Here is the relevant material:

1. PSl. **bérmę* (a) 'load, burden' (OCS *brěmę*, Russ. dial. *berémja,* Pol. *brzemię*, Croat. *brȅme*) < PIE **b^{h}er-* 'carry, bear' (Lat. *fero*, Gr. *phérō*, Skt. *bhárati*, etc.), cf. Skr. (L sg.) *bhā́rman* 'by bringing', RV 8,2,8 < **b^{h}ērmen-* (NIIL 16). Alternatively: PSl. **bérmę*, Skr. *bhárīman-* 'burden, maintenance' < **b^{h}er-H-men-* (but the suffix *-īman-* in Skr. cannot be used as evidence for the laryngeal, cf. *dhárman-* besides *dhárīman-* 'support' from **d^{h}er-* 'support, fix', IEW: 252–3, Wackernagel-Debrunner 1954: 756).

2. Lith. *káimas, káima* 'village' vs. *kiẽmas* 'farmstead', from the root of Goth. *haims* 'village'. Ultimately, they may be from the same root as number 9 below. Derksen (1996) explains the acute of *káima(s)* by metatony, pointing out that the reflex *-ai-* for original **-oy-* means that the root was originally unstressed, so that the acute may be attributed to the retraction of the stress from the last syllable. If this is correct, the circumflex of *kiẽmas* would be original, and Lith. *kiẽmas* and *káima(s)* would represent PIE **kóymo-* and **koymó-* respectively.

3. PSl. **kórsta* (a) 'crust' (Russ. *korósta*, Pol. *krosta* 'pustule', Croat. *krȁsta*) may be a *Vr̥ddhi* derivative from the root of Lith. *karšti, karšiu* 'card, comb (wool)', Latv. *kārst* 'id.' < PIE **sker-* 'cut, scratch' (OHG *skerran* 'scratch', Lat. *carro* 'card (wool)', DV: 95).

4. PSl. **lápa* (a) 'paw' (Russ. *lápa*, Pol. *łapa*, Croat. dial. *làpa*, Slov. *lápa* 'snout', ESSJa XVI: 26–28), Lith. dial. *lópa* (1) 'paw', Latv. *lāpa* 'paw'. Like Goth. *lofa* 'flat of the hand', this appears to go back to PIE **leh$_2$p-*, or **lōp-*, which may be a *Vr̥ddhi* formation to **lop-* seen in OCS *lopata* 'shovel', Russ. *lopáta*, Croat. *lòpata* (ESSJa XVI: 39–43), Lith. *lāpas* 'leaf'. However, Lith. *lópeta* 'shovel' and Latv. *lâpsta* show the word for 'shovel' with the long vowel and the acute (Smocz.: 363). Derk. (268–269) proposes that there were two different roots, **leh$_2$p-* (PSl. **lápa*, Lith. dial. *lópa*) and **lop-* (> OCS *lopata*, Lith. *lāpas*), but this seems like an ad hoc solution. Latv. *lę̃pa* 'paw' proves that we are indeed dealing with the lengthened grade (**ē*) rather than a root in laryngeal.

5. PSl. **lípa* (a) 'lime-tree' (Russ. *lípa*, Cz. *lípa*, Pol. *lipa*, Croat. *lȉpa*, Bulg. *lipá*, Slov. *lípa*, Derk.: 279, ESSJa XV: 114–116), identical to Lith. *líepa*, Latv. *liẽpa*. These words can be derived from the PIE root **leyp-* 'smear, glue' (Skt. *limpáti*, Lith. *lìpti, limpù*).

The semantic connection is in the sticky juice of the lime-tree. We may want to posit the original thematic noun **loypos* 'glue' (PSl. **lěpъ* (c), cf. CSl. *lěpъ*, Cz. *lep*, Croat. *lijêp*, Derk.: 273), and a BSl. *Vr̥ddhi* derivative **lēypos* 'sticky', substantivized as **lēypā* 'sticky one' > 'sticky tree'.

6. Lith. *lúobas* 'bark', Latv. *luobas* 'id.'; these may represent *Vr̥ddhi* derivatives of the root **lewb^h-* / **lub^h-* > Goth. *laufs* 'leaf', Lith. *lubà* 'plank', Latv. *luba* 'linden bark', Lat. *liber* 'bark', perhaps also in ORuss. *lъbъ* 'front of the head, skull'. A Slavic parallel **lúbъ* (with the acute) is found in Croat. dial. *lȕb* 'bark' (Vodice, Istra), *lȕba* 'lump' (Istra, Rijeka), Russ. *lúb* (G sg. *lúba*). Standard Croat. *lûb* (G sg. *lûba*) points to the circumflex. If we start from PIE **lowb^{h}o-* 'bark', the BSl. *Vr̥ddhi* formations may represent **lāuba-* '(made of) bark' > Lith. *lúobas* and PSl. **lúbъ*.

7. PSl. **pálica* 'stick, staff'' (OCS *palica*, Russ. *pálica* 'club', Cz. *palice* 'baton', Croat. *pȁlica*, Derk. 390) vs. PSl. **políca* 'shelf' (CSl. *polica*, Russ. *políca*,Cz. *police*, Croat. *pòlica*, Derk. 410), cf. also Russ. *pol* 'floor' (< **'plank') < PIE **(s)pol-* 'plank, staff' (OIc. *fjǫl* 'plank', Latv. *spals* 'handle', perhaps also Skt. *phálakam* 'plank'). Another *Vr̥ddhi*-derivative could be PSl. **pálьcь* 'finger' (CSl. *palьcь*, Russ. *pálec*, Cz. *palec* 'thumb', Croat. *pȁlac* 'thumb', Derk.: 390), which has been related to Lat. *pollex* 'thumb'; the derivational relationship might exist between **polo-* (> PSl. **polъ*) 'staff, plank' and **pōlo-* (> **palica*, perhaps **palьcь*).

8. PSl. **pítja* (a) 'nourishment, food' (OCS *pišta*, Russ. *píšča*, Croat. dial. *pȉća*, Cz. *píce* 'fodder', Derk.: 401) < PIE root **peyt-* (Lith. *piētūs* (N pl.) 'dinner', OIr. *ithid* 'eats', Skr. *pitú-* 'nourishment'). Derksen's (1996) assumption that the acute is due to the contamination with the root **peyH-* (Skt. *pīvan-* 'fat') is ad hoc.

9. Latv. *siẽva* 'wife' vs. OHG *hīwo* 'husband', Lat. *cīvis* 'citizen' and Skt. *śivá-* 'dear'; further connections to the root **k'ey-* 'lie', or the deictic particle **k'i-* 'this, here' are possible, but rather speculative. We might posit a derivational relationship between **k'ey-wo-* 'local, member of the local community' (DV: 116) and **k'ēywo-* 'belonging to the local community, own' > '(own) wife'.

10. PSl. **sláva* (a) 'glory' (OCS *slava*, Russ. *sláva*, Croat. *slȁva*, Cz. *sláva*, Pol. *sława*, Derk.: 453), Lith. dial. (Žemaitian) *šlóvė*. As Pronk (2012: 18–19) points out, many (but not all) derivatives from the root **k'lew-* in Balto-Slavic are acuted, so one must count with the possibility that a laryngeal was added to that root as a dialectal innovation. In Lith. *kláusti* 'ask' the laryngeal may be a part of the desiderative suffix **-Hs-*, and this may also be the source of the acute in PSl. **slúšati* 'listen' (OCS *slušati*, Russ. *slúšat'*, Croat. *slȕšati*, Pol. *słyszeć*, Derk.: 455), **slýšati* 'listen'. Note, however, Latv. *klàust* 'ask' without the acute, pointing to the conclusion that the acute in Lith. *kláusti* is secondary, as assumed by LIV. It may have been introduced to avoid homophony with *klausýti* 'listen', where the root is not acuted, cf. 3 sg. pres. *klaũsė* 'he listened' vs. *kláusė* 'he asked'. BSl. **k'lowo-* 'fame' (East Lith. *šlãvė*, *šlavẽ,* Latv.

slava, slave 'fame', unless these were influenced by *-a-* in Slavic cognates) vs. **k'lōwo-* 'having fame' ≫ **k'lōwā* > PSl. **sláva* 'famous deeds, glory'.

11. Lith. *sróvė* (1) 'stream' (besides *srovẽ* (4), both forms in LKŽ): this might be a *Vr̥ddhi* formation derived from the same root as Lith. *sraūja* 'stream' (OCS *struja*, Croat. *strúja*, Russ. *strujá*, etc., Vasm. III: 32–33), from PIE **srew-* 'flow' (Skr. *srávati,* Gr. *rhéō,* etc.). We might posit a derivative **srōwo-* 'flowing (water)' opposed to **srowo-* 'flow' (Gr. *rhóos*, Skt. *srāva-*, OCS *-strovъ* in *ostrovъ* 'island'), but the problem is that the evidence for the acute intonation in Lithuanian is scant. Derksen (1996: 59) considers this to be an instance of *métatonie douce* and derives *srovẽ* (the only form he cites) from the root of *srū́ti* 'flow'. However, there is no evidence for a laryngeal in the root **srew-*.

12. PSl. **tą̑ča* (a) '(snow-)storm' (Russ. *túča* 'dark cloud', OCS *tą̑ča* 'snow-storm', Croat. *tùča* 'hail', Vasm. III: 158–159), derived from the same root as Lith. *tánkus* 'thick', Skr. *tañc-* 'be solid', MHG *dīhte* 'thick'. The acute in Balto-Slavic seems to point to *Vr̥ddhi* (**tānk-jā* > PSl. **tą̑ča*). LIV reconstructs the root as **temk-* because of Hitt. *tamekzi* 'fixes'.

13. PSl. **ú(s)tro* (a) 'morning' (OCS *utro*, Russ. *útro*, Pol. *jutro*, Croat. *jȕtro*) may be a *Vr̥ddhi* formation build on the same root as Lith. *aušrà* 'dawn' (PIE **$h_2$ewsōs* > Gr. Hom. *ēṓs*, Lat. *aurōra*, etc.).

14. Lith. *vìlkė* 'she-wolf'. Pronk (2012) justly points out that this cannot be an old formation because of its fixed initial acute in light of Skt. *vr̥kī́*. Precisely: if this is an instance of *Vr̥ddhi* with respect to Lith. *viĨkas* 'wolf', it has to belong to a younger stratum, cf. also Lith. *žùikė* 'she-hare' vs. *žuĩkis* 'hare'. It probably belongs to the same Balto-Slavic stratum as the following item.

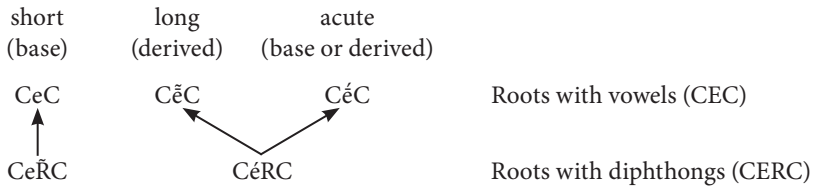
15. PSl. **vórna* 'crow, *corvus corone*' (Russ. *voróna*, Bulg. *vrána*, Croat. *vrȁna*, Cz. *vrána*, Vasm. I: 229) and Lith. *várna* appear to be a *Vr̥ddhi* formation with respect to PSl. **vornъ* 'raven, corvus corax' (Russ. *vóron,* OCS *vranъ,* Croat. *vrân*, Cz. *vran*, Vasm. I: 228) and Lith. *vařnas.* The similarity with Gr. *kóraks* 'raven' and *korṓnē* 'crow' is probably accidental, and does not testify to the difference of suffixes (masc. *-no-* vs. fem. **-Hno-*). In Baltic (though not in Slavic) this pattern of opposing masculines to feminines derived from the same root must have been productive, cf. Lith. *šeřnas* 'wild boar' vs. *šérnė* 'wild sow', *ántis* 'duck' vs. *añtinas* 'drake', cf. Petit (2004: 174–176).

16. PSl. **žę́tva* 'harvest' (OCS *žętva,* Russ. *žátva*, Cz. *žatva*, Croat. *žȅtva,* Cz. *žatva*, Vasm. I: 411) vs. **žęti* 'reap, mow' (OCS *žęti,* Croat. *žȅti*), Lith. *genė́ti* 'prune, hem', Derk.: 561. The PIE root is **g^{wh}en-* 'strike' (Hitt. *kuenzi,* Skt. *hánti,* Gr. *theínō*). The unexpected acute of **žę́tva* may be the result of the BSl. derivation process: **gentwo-* 'striking, mowing' ≫ **gēntwo-* '(time) of the mowing' > 'harvest'.

Any objective discussion of this material would have to admit 1) that the instances of acuted lengths in possible *Vr̥ddhi* derivatives are not numerous, 2) there are very few exact lexical cognates in Baltic and Slavic (**vórna* and *várna*, **sláva* and *šlóvė*, **lápa* and **lópė*, **lípa* and *líepa*), and 3) as noted by Petit (2004: 179ff.), most of the examples of alleged BSl. *Vr̥ddhi* do not involve long vowels, but rather diphthongs opposing acute intonation to the circumflex. The explanation of this opposition offered by Petit (2004: 180–181) for Baltic might work for Balto-Slavic as well:

> En d'autres termes, un degré long morphologique [ē] peut avoir en baltique, dans les voyelles, une existence distincte à la fois du degré bref [e] et du degré long d'origine glottale [ḗ], tandis que, dans les diphtongues, un degré long morphologique ne peut se distinguer du degré bref, s'il aboutit à une intonation douce [eȓ], ou du degré long d'origine glottale, s'il aboutit à une intonation rude [ér]. Il me semble que, dans ces conditions, l'économie du système favorise l'assimilation du degré long morpholoqigue plutôt au degré long d'origine glottale (d'où [ér] dans les deux cas) qu'au degré bref (d'où [eȓ] dansl es deux cas) : cette dernière hypothèse empêcherait toute possibilité d'un degré long morphologique distinct du degré bref, dans les radicaux à diphtongue du baltique.

Here is how we can represent this development:



If this is correct, then the acute, e.g. in Lith. *líepa* and PSl. **lípa* does not imply a proto-form **lḗypā*; rather, the derived form **léypa-* 'sticky' was opposed to the base form of the root **leȳp-* / **laȳp-* 'glue'. After this pattern, the association of the acute with the derived morpheme was transferred to proper vowels, so that the pattern **a* (base) : **á́* (derived) was established, e.g. in PSl. **lopa* vs. **lápa*, or **polica* vs. **pálica*.

5. Root nouns with lengthened grade in Balto-Slavic?

Most root-nouns in BSl. with cognates in other branches of IE are *i*-stems, cf. OCS *myšь* 'mouse' vs. Lat. *mūs*, *mūris*, OCS *noštь* 'night', Lith. *naktìs* vs. Lat. *nox, noctis*, etc. (Larsson 2001).

In some original *i*-stems we have the lengthened grade in BSl., and the root vowel is regularly non-acuted; however, in words belonging to AP (c) the acute may have been eliminated by Meillet's law. Here is a tentative list of Slavic *i*-stems that should be derived from earlier root-nouns.

1. PSl. **dalъ* 'distance' (Russ. *dal'*, Pol. *dal*, Croat. *dâlj*, ESSJa IV: 186–7), from the root of **dъliti* 'last' (Russ. *dlít'sja,* Cz. *dlíti*, Derk.: 133). The connection with PIE **dlh$_1$g^{h}o-* 'long' (Skr. *dīrghá-*, Gr. *dolikhós*) is possible, but uncertain.

2. PSl. **granъ* 'edge, boundary' (Russ. *gran'*, Pol. *grań*, Vasm. I: 304) and **grana* 'branch (Croat. *grána* 'branch', ULus. *hrana* 'edge', ESSJa VII: 106–107); the lengthened grade points to the vocalism of the Nom. sg. (PIE **grōn-*); the *o*-grade is preserved in OHG *grana* 'beard', and the *e*-grade in OIr. *grend* 'beard'.

3. PSl. **rěčъ* (c) 'speech' (OCS *rěčъ*, Russ. *reč'*, Pol. *rzecz* 'thing' Croat. *rijêč* 'word', Derk.: 434). From the root of **rekti* 'say' (OCS *rešti*, Croat. *rèći*, etc.).

4. PSl. **mělъ* (beside **mělъ*) 'sand bank' (Russ. *mél',* SerbCSl. *mělъ* 'chalk', Pol. *miel* 'shallow water', Vasm. II: 115), Lith. *smẽlis, smėlỹs* 'sand', cf. also Germ. dial. *māle* 'der feine Staub der Landstrassen' < **mēl-* and ON *melr* 'sand bank' < **mel-*.

5. PSl. **tvarъ* (c) 'creation, creature' (OCS *tvarъ*, Russ. *tvar'*, Pol. *twarz* 'face', Croat. *tvâr*), parallel to Lith. *tvorà* 'fence'. The same root is attested in **tvorъ* 'creation' with the full grade.

6. PSl. **žalъ* 'grief, pity' (OCS *žalъ* 'tomb', Russ. *žal'*, Pol. *żal*, Croat. *žȁo*, Derk.: 554), Lith. *gėlà* 'acute pain', from the root of OHG *quāla* 'violent death' (< **g^{w}ēlH-*) and OIr. *at-baill* 'dies' < **g^{w}elH-* (IEW: 471).

However, we also find some *ā*-stems that are good candidates for root-nouns,[3] cf. OCS *gora* 'mountain' vs. Lith. *girià* (2) 'wood' < PIE **g^{w}orH-* (Skt. *girí-* 'mountain'). These do not appear to have direct cognates in Baltic. Some such *ā*-stems show the long vowel in the root syllable. Here is a tentative list:

1. PSl. **travà* (b) 'grass' (OCS *trava*, Russ. *travá*, Croat. *tráva*, Cz. *tráva*, Derk.: 496) < PIE **trewH-* 'rub, spend' (Gr. *tryṓ* 'wear down, exhaust', OCS *tryti* 'rub', Lith. *trūnė́ti* 'spoil, putrefy, decay', LIV **trewH-*). Perhaps **trava* is a deverbal formation based on **travìti* (b) 'digest, feed with grass' (Russ. *travít'* 'exterminate by poisoning', Pol. *trawić* 'digest', Croat. *tráviti* 'feed with grass'), rather than vice versa as assumed by Derk.: 496. In its turn, **traviti* is an intensive/iterative of **truti* 'feed+ (OCS *natruti*, ORuss. *truti* 'consume', Pol. *truć* 'poison'), and Slavic intensives/iteratives regularly have the circumflex root (cf. PSl. **davìti* 'suffocate', Russ. *davít'*, Croat. *dáviti* < PIE **d^{h}ōh$_2$u-*, ESSJa IV: 198–199, Derk.: 97), PSl. **dirati* 'touch' (Croat. *dírati,* originally an intensive formation from the same root a s **derą*, **dьrati*).

2. PSl. **děra* 'crack, hole' (ORuss. *děra* 'opening', Cz. *díra* 'hole', ESSJa V: 12), from the root **der-* 'flay' (OCS *dьrati,* Lith. *dìrti,* Gr. *dérō*).

[3] See Matasović (2014: 21–24).

3. PSl. **dira* 'crack, hole' (OCS *dira*, Russ. *dirá*, Croat. dial. *dìra,* Derk.: 107, ESSJa V: 30–31); from the root **der-* 'flay'. The AP cannot be determined.

4. PSl. **kara* 'punishment' (Russ. *kára*, Pol. *kara*, Croat. *kára,* ESSJa IX: 151); derived from the root of **karati* 'punish' (Russ. *karát'*, Pl. *karać*, Croat. *kárati*), which is from **koriti* (Croat. *kòriti* 'reproach').

In Baltic, we find several *ā*-stems with long non-acuted vowel, e.g. Lith. *lomà* (2/4) 'hollow', *bylà* 'case, speech', *gėlà* 'pain', etc. Pronk (2012: 9) thinks these are best derived from old collectives. However, this type is very rare in other IE languages. Pronk cites only Lat. *cella* (derived by 'littera-rule' from *cēla, from PIE **k'el-* 'hide', cf. OIr. *celid*), which he thinks is the regular development of the collective **k'el-h_2* > **k'ēl-h_2*, with a "regular" lengthening in monosyllables. I find this too speculative, not only because there are too few lengthened grade *ā*-stems in other IE languages (even *cella* could represent **kelsā* or **kelnā*, among other things), but also because the lengthening in monosyllables is not a sound law established beyond doubt.

Larsson (2004ab) points out that the long circumflex vowel in Lithuanian *ē*-stems is often the result of a retraction from the following syllable (*VC-*íyā* > V̄:C-*iyā*). This pattern was extended analogically to many *ā*-stems, which are often parallel formations to *ē*-stems, with little difference in meaning, cf. Latv. *tvāre* 'fence' vs. Lith. *tvorà* 'id.' (from *tvérti* 'close'), Lith. *bėgẽ* 'run' vs. Lith. *begà* 'id.' (from *bė́gti* 'run'), Lith. *piovẽ* 'cutting' vs. *piovà* 'id.' (from *piáuti* 'cut'), Lith. *kõvė* 'fight' vs. *kovà* 'id.' (from *káuti* 'strike'), cf. Larsson 2004b: 166.

However, Larsson's lengthening (and the analogical spread to *ā*-stems) will allow us to explain the long vowel in Lith. *lomà*, *tvorà*, etc., but not in Slavic **lamъ* 'hollow, bend' (Russ. dial. *lam* 'wasteland', Pol. *łam* 'quarry, bend', Croat. dial. *lâm* 'knee-joint, underground passage', Slov. dial. *lam* 'quarry', Derk.: 268), **tvarъ* 'creature', etc., since there was no parallel retraction of the ictus in Slavic that would trigger the analogy. The Slavic nouns thus probably represent old root-nouns.

It is also possible that Balto-Slavic preserved a number of root-nouns with lengthened grade in the Nom. sg. and that some such nouns became *ā*-stems either in Balto-Slavic, or separately in Baltic and Slavic. In Slavic, these nouns mostly belong to AP c), so it is impossible to establish whether they were originally acuted or circumflexed. In Baltic, long vowels in the root of circumflexed *ā*-stems can always be the result of Larsson's lengthening. Therefore, it is impossible to establish the original accentuation of root nouns in BSl.

6. Conclusion

In our opinion, then, Balto-Slavic *Vr̥ddhi* is not inherited from PIE; rather, it is a parallel innovation in word-formation, similar to, but independent of Indo-Aryan and, possibly, PIE *Vr̥ddhi*. Nouns that can be characterized as showing Balto-Slavic

Vr̥ddhi regularly have the acute intonation, in contradistinction to nouns that have long vowels inherited from PIE, which are circumflexed. Moreover, the acute, rather than the vowel length, is the primary marker of the Balto-Slavic *Vr̥ddhi* as a process of nominal derivation.

References

Derk. = Derksen R. 2008. *Etymological dictionary of the Slavic inherited lexicon*. Leiden.
DV = De Vaan M. 2008. *Etymological dictionary of Latin and the other Italic languages*. Leiden.
ESSJa = Trubačev O.N. 1974(-2011). *Etimologičeskij slovar´ slavjanskix jazykov*. [vol. 1(-37)]. Moskva.
IEW = Pokorny J. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern.
LIV = Rix H. et al. 1998/2001. *Lexikon der indogermanischen Verben*. Wiesbaden.
NIIL = Wodtko D.S. et al. 2008. *Nomina im indogermanischen Lexikon*. Heidelberg.
Smocz. = Smoczyński W. 2007. *Słownik etymologiczny języka litewskiego*. Wilno.
Vasm. = Vasmer M. 1953–1958. *Russisches etymologisches Wörterbuch*. [vol. 1–3]. Heidelberg.

Carlton T. 1991. *Introduction to the phonological history of the Slavic languages*. Columbus.
Darms G. 1978. *Schwäher und Schwager, Hahn und Huhn. Die Vr̥ddhi-Ableitungen im Germanischen*. München.
Derksen R. 1996. *Metatony in Baltic*. Amsterdam.
Gołąb Z. 1967. The traces of vr̥ddhi in Slavic. – *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday, 11 october 1966*. The Hague, Paris: 770–784.
Kortlandt F. 2011. *Selected writings on Slavic and general linguistics*. Amsterdam, Atlanta.
Larsson J.H. 2001. Proto-Indo-European root nouns in the Baltic languages. – Huld M.E. et al. (eds.). *Proceedings of the 12th annual UCLA Indo-European conference, Los Angeles May 26–28, 2000*. Washington: 50–64.
Larsson J.H. 2004a. Metatony and length in Baltic. – Hyllested A. et al. (eds.). *Per aspera ad asteriscos. Studia Indogermanica in honorem Jens E. Rasmussen*. Innsbruck: 305–322.
Larsson J.H. 2004b. Length and *Métatonie douce* in Baltic deverbative nouns. – Clackson J., Olsen B.A. (eds.). *Indo-European word formation*. Copenhagen: 159–170.
Matasović R. 2008. *Poredbenopovijesna gramatika hrvatskoga jezika*. Zagreb.
Matasović R. 2014. *Slavic nominal word-formation*. Heidelberg.
Petit D. 2004. *Apophonie et catégories grammaticales dans les langues baltiques*. Leuven.
Pronk T. 2012. Indo-European long vowels and Balto-Slavic accentuation. – *Baltistica* 47.2: 205–247.
Villanueva-Svensson M. 2011. Indo-European long vowels in Balto-Slavic. – *Baltistica* 46.1: 5–38.
Wackernagel J., Debrunner A. 1954. *Altindische Grammatik. II. 2. Die Nominalsuffixe*. Göttingen.

Studia Linguistica Universitatis Iagellonicae Cracoviensis 133 (2016): 125–147
doi:10.4467/20834624SL.16.010.5156

IDA RAFFAELLI, DANIELA KATUNAR
University of Zagreb
ida.raffaelli@ffzg.hr, dkatunar@ffzg.hr

A DISCOURSE APPROACH TO CONCEPTUAL METAPHORS: A CORPUS-BASED ANALYSIS OF SPORTS DISCOURSE IN CROATIAN

Keywords: conceptual metaphors, discourse analysis, specialized digitized corpora, Croatian language

Abstract

This paper deals with the analysis of sports discourse in Croatian through the theoretical framework offered by conceptual metaphor theory. Within this framework, certain metaphorical expressions found in sports discourse are analyzed as expressions of two conceptual metaphors: SPORT IS WAR and SPORT IS FORCE. The analysis of these metaphorical expressions combines the methodology of cognitive linguistics with corpus linguistics, resulting in the proposal of a new method for discourse analysis in general. In our research, we introduce the notion of the specialized digitized corpus as a basis for further quantitative and qualitative research. On the basis of the specialized digitized corpus created for the purposes of this research, it is shown how the formation of sports discourse is dependent on three categories of metaphorical expressions relative to the degree of their conventionalization within sports discourse: (a) conventionalized, (b) semi-conventionalized, and (c) innovative metaphorical expressions. Each of these categories is analyzed according to their frequency and various aspects of meaning that it entails. Through the introduction of the semi-conventionalized metaphorical expression category, we aim to examine the gradable line between language creativity and conventionality as it is formed within the discourse of sports.

1. Introduction

Discourse as a complex structure of an utterance is a site where social forms of organization engage with systems of signs in the production of texts, thus reproducing or

changing the sets of meanings and values which make up a culture (Hodge, Cress 1988). From today's perspective discourse analysis is undoubtedly regarded as a method which provides insight into structures of knowledge and the speaker's understanding of the world. On the other hand, discourse not only reflects the speaker's knowledge of the world, but also actively builds the person's identity in his or her cultural environment. Discourses are not about objects; they constitute them (Foucault 1972). It is through discourse that meanings, subjects, and subjectivities are formed (Wright 2000).

Discourse analysis distinguishes between two types of context, as suggested by Malinowski (1923), who separated the *context of situation* from the *context of culture*. The context of situation refers to the structure of the immediate situation in which an utterance takes place and which a speaker must be familiar with in order to understand a particular instance of language properly. Context of culture, on the other hand, represents the broader structure of a culture a speaker must know in order to understand what is said or written (cf. Lee, Poyton 2000).

In view of these two definitions, it is obvious that every study of a discourse must include an analysis of the context of situation as well as an analysis of the context of culture. The former is necessary in determining the immediate content of a message that the hearer can decode from an utterance. The latter is essential because every specific discourse is situated within a cultural context, and thus a speaker must decode not only the immediate content of a message but also the structures of a specific context of culture. It is thus necessary to regard every discourse as a complex structure of knowledge shared by the speakers of a specific culture.

Although the importance of sports can be seen on many levels of everyday life, we are mostly interested in the linguistic aspects of sports, specifically the discourse practices that surround it. Papers dealing with linguistic phenomena in sports discourse view sport mostly as a source domain used to understand other domains of human activities, such as the media and politics (cf. Blain, Boyle, O'Donnell 1993; Bairner 2001; Callies 2011). They also view sport as a domain that actively shapes the cultural identity of a speech community. Callies (2011), for instance, analyzes the degree to which various metaphors used in American popular sports (such as baseball) form various aspects of American culture and identity on the basis of the conceptual metaphor LIFE IS A GAME.

This paper has a somewhat different way of approaching the subject of metaphors related to the sports domain. We will not regard sports as a source domain but conversely as a target domain. Our goal is thus to analyze the basic domains that form our understanding of sports as a part of our culture. If we go back to Malinowski's definition of the context of culture, it becomes clear that an analysis of sports discourse will enable us to extrapolate important elements of the context of culture a certain discourse is related to. Furthermore, since the context of culture is essentially a complex structure of knowledge, an analysis of sports discourse will enable us to describe conventional knowledge about sports that we all share as members of a specific culture.

We also intend to show how these parts of background knowledge participate in the formation of sports discourse. Furthermore, we believe that a careful examination

of metaphorical expressions in sports discourse will contribute to the study of conventionalized and innovative metaphors as gradable categories and their mutual interaction within a discourse. We assume that some of our conclusions might be relevant for the broader Western context of culture being background knowledge for sports discourse shared by different but related speech communities.

The discourse approach to conceptual metaphors we propose is based on building and analyzing a *specialized digitized corpus* (henceforth SDC) of texts exclusively related to the sports domain.

Since the main goal of the proposed analysis is to define how metaphors constitute knowledge about sports in Croatian culture and how they are used to convey this knowledge in sports discourse, the SDC offers many benefits for this kind of research, which will be pointed to in the next sections.

2. Conceptual metaphors, discourse and culture

Since the publication of Lakoff and Johnson's *Metaphors we live by* (1980), conceptual metaphors have been one of the main topics in cognitive linguistics. They have been studied from various perspectives that include philosophical, cross-linguistic, grammatical, corpus, cognitive, and psycholinguistic approaches (e.g. Lakoff, Johnson 1980; Langacker 1987; Lakoff, Turner 1989; Žic Fuchs 1992; Mahon 1999; Steen 1999; Kövecses 2000; Johnson, Lakoff 2002; Kövecses, Palmer, Dirven 2003; Charteris-Black 2004; Deignan 2005; Steen 2007; Stanojević 2013). Many works in cognitive linguistics deal with the way in which conceptual metaphors reflect knowledge and reveal a view of the world as constructed by a specific culture. An analysis of conceptual metaphors contributes to the understanding of the culture itself (e.g. Lakoff 1987; Sweetser 1990; Kövecses 2005).

The initial thesis of this paper is that conceptual metaphors can be viewed as knowledge structures that are integral to a culture. This means that defining metaphors within a discourse presupposes a reconstruction of conventional knowledge characteristic of a certain speech community. Methodologically, this means that an analysis of metaphors in the discourse provides an insight into the metaphorical structures of the context of culture.

Therefore, we assume that the analysis of metaphors in sports discourse enables a reconstruction of metaphorically based knowledge of sports in the Croatian speech community.

Parallel to this assumption is another aspect of metaphorical discourse formation. The thesis that every context of culture is metaphorically structured is similar to Lakoff's thesis that speakers metaphorically understand and structure every situation (Lakoff 1990). According to Lakoff (1990), metaphorical structuring of situation consists of two parts. The first part is made up of a relatively stable set of metaphors that determine our view of a situation, while the second part is actually our ability to apply these metaphors when communicating about that situation. Furthermore, the speaker is able to linguistically form these metaphors in different

ways. It is thus necessary to find and analyze different expressions that function in a discourse as different facets of the same conceptual metaphor. Lakoff's thesis can be incorporated into our analysis of metaphors in sports discourse as well as any other type of discourse. It relates to the fact that discourse reflects the way we view a specific situation. Conversely, the principles of the metaphorical understanding of a situation build and shape the discourse itself.

In our research we will discuss two conceptual metaphors which in different ways shape the sports discourse in Croatian. These are SPORT IS WAR and SPORT IS FORCE.[1] Defining these two conceptual metaphors as very relevant for the sports discourse in Croatian was a result of a careful examination of a specialized digitized corpus of sports texts and the observation of lexical units which systematically draw their meaning from the same source domains. In English, for instance, lexemes such as *offense, defense* and *shooter* also show systematic relations with the notion of conflict (i.e. war), each in their own way (through the notions of "attack", "defense" and "attacker", respectively). Also, in expressions such as

(1) Ljubo srušio Nadala (SDC, www.vecernji.hr)
'Ljubo knocked down Nadal' (tennis)

the notion of "force" is used to denote the victory of one tennis player over the other. Many verbs used in these context, such as *potopiti* 'to sink', *pomesti* 'to sweep up', *otpuhati* 'to blow away' refer to some kind of animate or inanimate agency (the sea, a person handling a broom and the wind, respectively) exerting force over another entity. Put together they point out that metaphorical links are not only confined to individual lexemes, but that their metaphorical use in the sports discourse relies on a wider background conceptual structure that links the two domains and that structure, i.e. a conceptual metaphor. This is the reason we can find many different lexical units related to the same source domain. We must point out that although it is possible to analyze various scenarios related to different sports, our goal was to establish broad conclusions about the structure of the domain of sports in Croatian within the limits of one paper, and for this reason we discuss the sports domain as a whole, and put forth the two conceptual metaphors as an important part of that whole.

It is also important to point out that the analysis of particular lexical units as metaphorical within our corpus was made by cross-referencing existing Croatian dictionaries.[2] For instance, the lexeme *napad* 'attack; offense' has the following senses listed in Anić (1996): (a) an act of aggression with the intent of causing physical or psychological harm; (b) a short physical or psychological disturbance (e.g. *napad kašlja* 'a cough attack') and (c) a sports action with the goal of achieving a score. The lexicographic data therefore points to its metaphorical meaning in the domain of sports.

---

[1] Based on the analysis of metaphorical instantiation in sports discourse, it has become evident that the largest number of examples belong to these two conceptual metaphors. Some methodological details will be given in the sections that follow.

[2] Anić (1996) and Šonje (2000).

Therefore, we intend to (a) analyze expressions that are related to either of the two conceptual metaphors within sports discourse, (b) point out how sports discourse is in large part shaped on a different degree of inclusion and linguistic instantiation of these metaphors, and, consequently, (c) to give a classification of metaphorical expressions based on different qualitative and quantitative criteria.[3]

3. Discourse analysis and corpora

In her book *Metaphor and corpus linguistics* (2005), Deignan points to differences between the cognitive and discourse approaches to metaphor research,[4] stressing the existence of two different discourse approaches to metaphor research. The cognitive approach analyzes speech or writing of a particular text-type, generally with the agenda of showing how metaphors are used to present a particular message or ideology. The discourse approach, represented by a smaller group, looks at how speakers use metaphor to develop shared understanding as a spoken discourse unfolds. What the two approaches share is the close analysis of text as a product.

The discourse approach to metaphor research that we propose is somewhat different from those described by Deignan (2005).[5] In a way, it has some common features with the cognitive approach, since the main intention of the proposed discourse analysis is to define how metaphors are used in line with knowledge about a certain human activity and how they are used to convey this knowledge through discourse communicating this specific activity. However, methodologically it proposes some new procedures for discourse analysis and for metaphor research as related to discourse analysis.

---

[3] One could analyze both the source and the target domain in detail but we decided to delimit our research in two ways: (a) give an overview of the sports domain in general, not particular sports, e.g. tennis vs football etc. and (b) examine texts that describe the competition aspect of sports, i.e. the game itself, and not for instance the politics of sports (such as player tranfers and the like).

[4] Deignan (2005: 123) points to two basic differences between the discourse and the cognitive approach to metaphor research. First, the discourse approach tends to take Conceptual Metaphor Theory as a starting point and does not attempt to test the theory itself. The main intention of this approach is to see how speakers create meaning using metaphor as a tool. Second, the discourse approach is based on naturally occurring texts and therefore does not ask speakers for any metalinguistic interpretation of their own utterances. Speakers are not supposed to invent sentences not produced in natural linguistic and extralinguistic circumstances.

[5] Cameron and Deignan (2003) and Charteris-Black (2004) conducted their research on metaphors on small corpora that were either hand-sorted corpora or a sample gathered from a larger corpus. Both corpus approaches to metaphors were based on searching data by hand. It has to be pointed out that Cameron and Deignan's small corpus consists of 28,285 words and that it consists of transcribed talk in a primary school in the UK. The large corpus consists of the 9-million-word collection of spoken data from the section of the Bank of English. Cameron and Deignan's corpora were not specialized corpora, whereas Charteris-Black's corpus was topically related to 9/11 and especially to the metaphorical use of the word *crusade*.

3.1. Building a specialized digitalized corpus for discourse analysis purposes

The main benefits of building an SDC for discourse analysis purposes is the possibility for a researcher to do the following:

1. constitute a corpus based on texts that are of special interest to his or her research. Since a specialized corpus is a corpus of domain-related texts, it can give us a coherent overview of the sports discourse as a whole because it serves as a representative sample of the entire discourse;
2. define the extent as well as the structure of the corpus.[6] It provides insight into a considerable amount of most recent sports-related texts that cannot be obtained any other way, except for manually searching articles on the Web;[7]
3. conduct a statistical analysis of a certain linguistic phenomenon within a discourse. An SDC is more easily manageable in terms of the statistical analysis of frequencies, especially with respect to the analysis of metaphorical expressions.

Furthermore, this methodology significantly differs from the methodology proposed by Charteris-Black (2004), who restricts a general corpus for discourse analysis purposes to texts related to a certain domain.[8] A convenient aspect of this methodology is that a general corpus serves as a control corpus. However, his findings on metaphorical expressions are based on a manually conducted collection of metaphors from several small corpora extracted from a general corpus. An SDC is built by a researcher interested in texts related to a certain domain and collected from the Web. However, a main disadvantage of such a corpus would be the necessity that texts are available on the Internet, i.e. that they are digitally processed.

In building our SDC, we used the BootCat (short for Bootstraping Corpora and Terms from the Web) tool engine to gather and sample texts from the Internet. BootCat is a recently published software that collects various texts from the Internet and builds a corpus through a set of instructions provided by the researcher (for a detailed description, see http://bootcat.ssmlt.unibo.it). These instructions can be modified according to the type of corpus a researcher is interested in. In our case, we wanted to get the most recent synchronic overview of sports discourse. Also, we wanted to further restrict our search only to highly relevant sports related texts, so we gathered articles from newspaper portals of *Jutarnji list*, *Večernji list* and *Sportske novosti*.[9] This means that we were interested primarily in texts describing

---

[6] Building specialized corpora leaves it up to researchers to define how large the corpus should be and what kind of texts it should consist of.

[7] The Croatian National Corpus ver. 2.0, a large, 100-million-word general corpus, has a limited overview of sports-related texts, none of which are more recent than 2005.

[8] Methodologically, it is not clear enough whether the extracted texts form separately manageable small corpora. From a corpus linguistic point of view, this should be an important methodological question. Since there are no statistical data (even the size of the small corpora is not mentioned) we assume that these were not separately manageable small corpora.

[9] We wanted to collect texts that are explicitly part of public communication about sports. We therefore decided not to include texts from forums and blogs, though these could be added in further analysis. Judging by the number of readers and their overall popularity, these are the most influential newspapers and magazines in Croatia.

sports competitions, and not peripheral topics such as transfers of football players in order to delimit our research. The corpus contains 1,195,883 tokens. From a corpus linguistics point of view, this is considered a small corpus sufficient and relevant (Sinclair 2001) for this type of research. It can also be expanded by adding new texts if the need arises.

In this respect our methodology is more in line with the research conducted by Skorczynska and Deignan (2006) on metaphors in the economical domain. Skorczynska and Deignan (2006) compiled their texts from two journals dealing with economical topics, one in a scientific and the other a popular way, in order to compare the two periodicals and draw conclusions on their functions in the economical discourse. However, there are a few differences between our approach and the one made by Skorczynska and Deignan. Whereas the comparison made by Skorczynska and Deignan is made between two already specialized corpora our comparison will be based between the SDC and a general corpus of Croatian. Also, our work results in formulating two salient conceptual metaphors in the sports discourse and their comparison in terms of the degree of conventionality, whereas Skorczynska and Deignan focus more on the metaphor vehicles stemming from various domains without positing conceptual metaphors in terms of X IS Y (e.g. SPORT IS WAR). Our specialized digitized corpus is furthermore compiled and processed semi-automatically from a variety of texts gathered from general newsportals, not specialized sport periodicals.

4. Classification of metaphors

Implementation of a corpus-based analysis of sports discourse is especially useful for examining and studying metaphorical linguistic expressions as instantiations of the conceptual metaphors SPORT IS WAR and SPORT IS FORCE. As we have already pointed out, these two conceptual metaphors are highly representative in the SDC via diverse linguistic instantiations, but with different frequencies, which further confirms our initial thesis of their different salience and relevance not only in sports discourse but in construing knowledge about sports as part of the context of culture as well.

This brings us to another important aspect of this study. Since we stressed that our discourse approach to metaphor research is similar to that of scholars who are interested in use of metaphors in communicational purposes as conveying knowledge structure about something in the world, an SDC can provide some interesting statistically supported evidence about the linguistic instantiations of metaphors. An overview of the SDC revealed that conceptual metaphors SPORT IS WAR and SPORT IS FORCE participate on different levels and with differing frequency based on their linguistic instantiations in sports discourse.[10] With respect to their linguistic

[10] According to Kövecses (2000), source domains of WAR and FORCE are separate but closely related conceptual domains. This relation can be illustrated by words such as *demolition* and

instantiation, these two conceptual metaphors are not conventionalized to the same degree. Based on these observations, we propose three different levels of their instantiations which determine the linguistic actualization of any conceptual metaphor, not just of the ones studied in this paper. These are (a) *conventionalized metaphors*, (b) *semi-conventionalized metaphors*, and (c) *innovative metaphors*.

Authors usually talk about the gradability or the continuum between conventionalized and innovative metaphors because the degree of conventionalization can be determined on various levels (e.g. Žic Fuchs 1992; Deignan 2005; Kövecses 2002; Stanojević 2013).[11] The proposed classification of conventionalized metaphors in this paper differs in part from other classifications (cf. Deignan 2005). It is our opinion that metaphor research based on searching an SDC could provide some new evidence in the principles by which metaphors structure (sports) discourse. In the following sections we will examine the criteria of our classification in more detail. We believe that with this kind of classification of metaphors we will be able to (a) demonstrate the various degrees of their inclusion in sports discourse and (b) contribute to a more systematic reconstruction of the context of culture of sports discourse.

Therefore, we state that the intermediate category of semi-conventionalized metaphors is necessary for explaining more systematically the way conceptual metaphors participate as background knowledge in the structuring of (sports) discourse.

4.1. Criteria for the classification of metaphors

Our classification is based on several quantitative as well as qualitative criteria.[12] These are (a) *frequency of use*, (b) *affective markedness*, and (c) *individual use*. Table 1. shows the way these criteria participate in the classification of metaphors in the sports discourse.

a) Frequency

Frequency of use is the most typical quantitative feature of an analyzed expression and the only one of the three criteria that is related to quantitative analysis. The other two criteria are related to the semantic features of the analyzed expression. Since our corpus

destruction, whose meanings are closely tied to both a WAR scenario and a scenario of exerting FORCE. Such is the example of *srušiti* 'destroy, tear down', e.g. *u ratu su srušeni mnogi spomenici* 'many monuments were destroyed in the war'. Thus the lexical (sub)system related to the two source domains of WAR and FORCE indeed serves as a window into the structure of the underlying system of metaphors (see also Kövecses 1986). On the other hand, in our research we also present arguments that support the division of the two conceptual metaphors SPORT IS WAR and SPORT IS FORCE.

[11] It is worth stressing that Žic Fuchs' criteria for the classification of metaphors as conventionalized and innovative are strictly qualitative, whereas Deignan's are a combination of qualitative and quantitative criteria although it seems she considers them as separate entities and not as a group.

[12] Charteris-Black (2004) and Deignan (2005) point out that the corpus approach to metaphor research provides us with a quantitative and qualitative analysis of linguistic data. Quantitative analysis mostly deals with the frequency of an expression, and qualitative analysis deals with its meaning.

	Conventionalized metaphors	Semi-conventionalized metaphors	Innovative metaphors
Frequency	more than 100 tokens	between 100 and 10 tokens	less than 10 tokens
Affective markedness	–	+	+
Individual use	–	–	+

Table 1. The classification of metaphorical expressions according to the presence of several criteria. Frequency criteria is based on a one million token corpus.

is an SDC, the meaning issue of the expression does not concern the relation between literal and metaphorical meaning since the analyzed expressions appearing in this SDC are exclusively metaphorical. For example, in sports discourse, the word *napad* 'offense; lit. attack' cannot appear in its literal meaning, but exclusively metaphorically. This is because the SDC was manually checked after compiling to exclude examples such as, for example, fights between football hooligans or players on the field, leaving only the instances of "attacks" that refer to the game strategy. Therefore, our qualitative criteria are related to the function certain expressions have in construing a discourse, their additional semantic features, and individual and creative usage.

In view of our own classification, the research presented by Deignan (2005) was of special interest to us. Deignan's corpus research was based on the corpus analysis of a general corpus of English (Bank of English). To demonstrate the frequency of metaphorical uses of a certain linguistic expression, she uses the *absolute frequency approach* (see Deignan 2005). This approach enables to determine the frequency of metaphorical uses of a certain expressions with respect to literal uses and to the total number of uses.[13] However, frequency criteria are not equally essential for determining conventionalized metaphors and innovative metaphors. Thus, Deignan (2005: 40) determines innovative metaphors as those linguistic expressions where a particular metaphor is found occurring less than once in every thousand citations of the word form. On the contrary, frequency criteria are not relevant in determining conventional metaphors; i.e. it seems that every expression found more than once in every thousand citations of the word is a conventionalized metaphor. According to Deignan (2005), conventionalized metaphors are dependent on a literal sense and they tend to invoke at some level a literal counterpart.

[13] In her research of the expression *in the running* Deignan (2005: 28–29) states that from a total number of 124 citations, 61 citations are metaphorically motivated. However, the number of metaphorical expressions could be considered larger with respect to the use of this expression in the sports discourse where this expression is often used in a more general metaphorical sense of competition.

It should be stressed that these data tell us nothing about the relevance of metaphorical expressions structuring the discourse.

If we look at Deignan's criteria as applied to the definitions of conventionalized and innovative metaphors, it is quite obvious how the two criteria (quantitative and qualitative) are used independently to define the two categories of metaphors. Consequently, they do not elaborate on the way innovative metaphors interact with more conventionalized ones or the way one can observe and study the process of conventionalization itself.[14] We therefore decided to use mutually interacting criteria in our analysis of metaphorical expressions. In this sense, they have to be taken into consideration as a group, not separately. In this way, the degree of absence or presence of any of these three criteria is also indicative (a) of the values of the other two criteria and (b) of the degree of conventionalization characteristic of a metaphorical expression. As stated by Bybee and Hopper (2001), frequency is one of the factors that conditions functional change of a word or any linguistic element. Although their main focus is on grammaticalization, we think that frequency is inseparable from the changes that could affect a certain word, such as its function in a discourse or the (non)existence of certain semantic features.

For example, high frequency of use is indicative of the low affective markedness of the metaphorical expression simply because a phrase heard quite often in the same context does not have as strong an emotional impact on the hearer as does an expression heard once or only a few times. Thus high frequency is in contrast with the individual use of a word.

Whereas Deignan used absolute frequency as a criterion, we decided to use the *relative frequency approach* instead, with some references to the absolute frequency of the word when needed. Since relative frequency shows the number of times an expression appears in relation to the absolute number of tokens in a corpus, we believe that it enables us to better examine the relation between a metaphorical expression and a discourse in its entirety.[15] High relative frequency of certain metaphorical expressions is indicative of conventionalized metaphors and their importance in constituting the sports discourse itself. Transitional stages with lower frequency are in the same way indicative of semi-conventionalized metaphors which have a lower impact in structuring a discourse but significantly higher than innovative metaphors. It is important to stress that metaphorical expressions with lower frequency are affectively more marked than those with a higher frequency status.

---

[14] Deignan (2005: 40) stresses that all conventional linguistic metaphors must have been innovative at some point in history. Innovative language uses are related to individual creative *ad hoc* uses of a language in a certain communicational situation. From a diachronic semantics point of view, this is an important issue elaborated by cognitive linguists (Nerlich and Clarck 1988 and Györi 2002) as much as by structural linguists (Coseriu 1973) or the prestructural linguist Sterne (1931). For more details see Raffaelli (2009).

[15] It must be noted that such calculations are made much easier by the use of a specialized corpus where one works with a smaller set of tokens and where the expressions are expected to appear mainly in their metaphorical meaning pertaining to a single domain, in our case sport. It is much harder, as Deignan shows, to hand-pick various metaphorical meanings one by one from a general corpus. In other words, in a SDC related to the sports domain, we can hardly expect that a word such as *napad* 'attack; offense' and *obrana* 'defense' would be used in their literal meaning. Therefore, this methodology makes it much easier to define the "real" status of metaphorical expression in construing sports discourse.

Finally, in accordance with Deignan's definition, innovative metaphors appear rarely and thus their relative frequency is negligible, though also indicative of their category status.

Based on this observation and according to our analysis, a provisional frequency limit between conventionalized and semi-conventionalized metaphorical expressions is a frequency of 100 citations in the whole corpus of 1 195 883 tokens.[16] The frequency limits we suggest are in a way different from those proposed by Deignan (2005). First, the criterion to determine innovative metaphors as those expressions which appear in less than one citation in every thousand citations of the word form would categorize all the metaphorical expressions in the SCD as innovative metaphors.[17] This is also a matter of the small size of the SDC as compared to a general corpus. Based on Deignan's frequency limit, even the most frequent metaphorical expressions in the SDC should therefore be classified as peripheral members between innovative and conventionalized metaphorical expressions. Therefore, the frequency limits we propose are adjusted to the absolute frequencies of the analyzed metaphorical expressions in the SDC.

Since the main goal of the presented research is to determine to what degree expressions belonging to different conceptual metaphors participate in construing sports discourse in Croatian we propose the following frequency limits: (a) most of expressions that are categorized as conventionalized metaphors appears in more than 100 citations in the corpus,[18] (b) expressions appearing in citations between 100 and 10 are categorized as semi-conventionalized, and (c) those expressions that appear in less than 10 citations in a 1-million-word corpus are categorized as innovative metaphors.

b) Affective markedness

The second criterion, affective markedness, is what Leech (1974) defines as affective meaning, the type of meaning which includes the speaker's attitudes and feelings towards the thing he or she is talking about. Affective meaning is often explicitly conveyed through the meaning of the words used. In discourse analysis, affective meaning, or affective markedness in our terminology, can be correlated with a larger degree of innovation of certain expressions in the discourse. The less frequent the expression is, the less it is conventionalized and more it is affective. This is the part of the discourse where variability and the speaker's creativity come into play. Relatively frequent expressions such as *razbiti* 'to demolish', *slomiti* 'to break', and *zgaziti* 'to stamp out' (an opponent) in the sense of *to defeat an opponent* all bring additional semantic features to the utterance, intensifying the meaning of *defeat* and adding certain emotional overtones to sports discourse. Affective markedness thus shows us that such a metaphorical expression brings some additional semantic information

[16] For the clarity of presentation, we will henceforth refer to the corpus as a one-million token corpus.

[17] Its absolute frequency in the SDC is 1,439 citations.

[18] The nonmetaphorical word *utakmica* 'match' appears 4,430 times. These are the two most frequent expressions related to the sports domain and according to Deignan's frequency limits, hardly conventionalized.

into the discourse in contrast to those expressions that are categorized as conventionalized. Metaphorical expressions such as *razbiti* 'to demolish', *slomiti* 'to break', and *zgaziti* 'to stamp out', with the meaning of *to defeat an opponent*, are more affective than the expression *pobijediti* 'to win', which lacks any affective features and is highly schematic.[19] Such metaphorical expressions are less frequently used than conventionalized metaphors. Therefore, they will be placed in the category of semi-conventionalized metaphors.

c) Individual use

The third criterion is individual use, which simply refers to the uniqueness of the expression the speaker uses, i.e. the speaker's creativity. It correlates well with the frequency criterion since expressions that could be defined as individual speaker's use have very low frequency (less than 10 citations in the one-million-word SDC), often highly affective and belong to the category of innovative metaphors. As we will demonstrate in the sections to follow, innovative metaphors in sports discourse exhibit some specific semantic and syntactic features that have to be pointed out.

As mentioned previously, the three criteria are interrelated and do not form absolute categories themselves. In other words, all three are also gradable in the sense that, for instance, semi-conventionalized metaphors have a greater degree of affective markedness and individual use than conventionalized metaphors, which have none or close to none. But also, semi-conventionalized metaphors also have a lower degree of affective markedness and individual use than innovative metaphors, which in turn have both as their distinguishing features.

As will be pointed out, the categorization of metaphorical expressions should be based on the interaction of all three criteria at the same time. Frequency as a quantitative criterion is relevant because it correlates significantly with some functional changes of words in a discourse and their semantic features. It should be regarded in interaction with the other two, qualitative criteria.

5. Three types of metaphorical expressions

In this section we turn to look at the types of metaphorical expressions based on their degree of conventionalization in sports discourse. As we have already stated, our classification differs from existing classifications because it introduces a new intermediate category between conventionalized and innovative metaphors, that of semi-conventionalized metaphors.

---

[19] *pobijediti* 'to win' has two senses in Anić (1996): a) to overcome an opponent and b) to contain something or someone. Because it is highly schematic in its meaning we will use it as a reference point for the comparison of actual metaphorical extensions from the domain of WAR, but it will not be itself analyzed as a metaphorical expression.

5.1. Conventionalized metaphors

Conventionalized metaphors form the backbone of sports discourse. They are an essential part of the vocabulary used to talk about sports, and as such have very high citation frequencies (see Table 2).[20] If compared with the frequencies of semi-conventionalized expressions (see Tables 3 and 4, below) it is obvious that frequencies of conventionalized metaphors are significantly higher compared to those of semi-conventionalized metaphors. The two most frequent metaphorical expressions are *osvojiti* 'to seize' and *dvoboj* 'duel'. The expression *osvojiti* appears in 983 citations and *dvoboj* in 859 citations in the SDC.

Metaphorical expressions	No. of tokens	Relative frequency
osvojiti 'to seize; to win'	983	0.0825%
dvoboj 'duel'	859	0.0718%
obrana 'defense'	852	0.0717%
napad 'offense', lit. 'attack'	535	0.047%
napadač 'striker', lit. 'attacker'	525	0.044%
strijelac 'shooter'	506	0.0423%
izboriti lit. 'fight out'	482	0.0403%
savladati / *svladati* 'overcome'	442	0.0369%
izbaciti 'throw out'	125	0.0289%
obraniti 'to defend'	214	0.0178%
zaustaviti 'stop'	179	0.0149%
veteran 'veteran'	129	0.0107%

Table 2. Conventionalized metaphors in SDC

Although a scalar structure of the category of conventionalized metaphor is pertinent with respect to the frequency criterion, almost all of these conventionalized expressions also function as key-words of the sports discourse. What we mean by key-word function of certain metaphorical expressions will be explained in the next section.

5.2. Conventionalized metaphors as key-words of sports discourse

The relevance of conventionalized metaphors for sports discourse is also seen in the way they reveal important elements of the context of the culture of sports discourse.

[20] The only two expressions with significantly lower frequencies are *veteran* 'veteran' and *bunker* 'bunker, a style of defensive play'.

It is through this perspective that they can be viewed as key-words of sports discourse. What is meant here by key-words is as Matoré defines them: "lexicological units expressing a society… a person, a feeling, an idea which are alive insofar as society recognizes in them its ideal," (Matoré 1953: 68). An example of this is the French term *bourgeois*, which according to Matoré (1953: 69) compiles the dominant cultural meanings of its historical period.

Many studies based on Matoré's assumptions deal with diachronic research whose goal is to describe the dominant ideas of a certain historical period, an author's work, or a specialized historical activity (Ullmann 1962: 252–253). From a synchronic perspective, key-words were the subject of research lead by Wierzbicka (1997), who based her comparative cultural analysis on the discovery and analysis of culturally specific lexemes.

What both the diachronic and the synchronic approach have in common is the wide scope of phenomena they include in their investigations. They both focus on key-words that directly represent the society as a whole, extant in various discourse practices and contexts. On the other hand, they exclude key-words that shape a single discourse practice within a society. This way of approaching key-words is highly important because we consider key-words as elements that represent the way speakers view and understand a single activity which is the topic of a specific discourse. It was thus our goal to limit ourselves to those words that can be interpreted as key-words in sports discourse and which thus represent the core structure of the context of culture involved in Croatian speakers' understanding of sports. These key-words are metaphorical instantiations of conceptual metaphors by which sports are mostly conceptualized.[21]

Such is the example of *izboriti finale* lit. 'fight out the final'. Even in cases where this metaphorical expression can be substituted with another metaphorical expression based on another conceptual metaphor such as SPORT IS A JOURNEY, e.g. *izboriti finale* lit. 'fight out the final' vs *ući u finale* 'enter the final', the metaphorical expression which is the instantiation of the conceptual metaphor SPORT IS WAR is seen more frequently, having semantic features (such as endurance and supremacy over the opponent) that the other metaphorical expression, *ući u finale* 'enter the final', does not possess because it lacks the concept of 'competition' coded in the expression *izboriti finale* 'fight out the final'.

Some other examples are expressions such as *osvojiti* 'to seize' (which occurs with objects such as *prvenstvo* 'championship', *zlatnu medalju* 'gold medal', *naslov prvaka* 'title'), *izbaciti* 'to throw out (*iz prvenstva* 'from the championship', *iz finala* 'from the final', *iz natjecanja* 'from the competition')', *obraniti* 'to defend (*naslov* 'title', *gol* 'the goal')' and *strijelac* 'the shooter (*pobjedničkog pogodka* 'of the winning shot')', which also lexicalize some parts of sports competitions that are fundamental when communicating about the sport and cannot be expressed in some other way. It is impossible to refer to these events with any other lexical unit or expression, and as such, they are affectively completely neutral.

[21] Charteris-Black (2004: 37).

Key-words are not marked for affective meaning and are not part of an individual usage. To say that someone is a *football veteran* (*nogometni veteran*), will bear no special unique meaning to the hearer and will be a common way to talk about the persons denoted by the expression. It is clear that both of these metaphorical expressions are taken from the domain of WAR, but through the process of conventionalization they have lost semantic features related to the source domain and formed new ones pertaining to the domain of sports. Thus, when we talk about Davor Šuker as a *football veteran,* we do not think of a possibly troubled, physically and psychologically scarred individual, but of an experienced retired sportsman with notable prior achievements in his profession. These features – experience, retirement, and earned appreciation – motivate the metaphorical meaning of the lexeme *veteran*.

By defining conventionalized metaphors as key-words of sports discourse, we point out the reciprocity of the context of culture and discourse as mediated through the key-words of the discourse. This means that by examining the expressions which are instantiations of the conceptual metaphors SPORT IS WAR and SPORT IS FORCE, we are (a) examining the dominant ways in which sports are understood and talked about in Croatian culture, i.e. (b) getting an insight into the background knowledge that is fundamental in shaping sports discourse in Croatian.

5.3. Semi-conventionalized metaphors

Semi-conventionalized metaphors are the intermediate and also the most unstable category of metaphors in sports discourse. Their frequency is significantly lower (less than 100 citations in the SDC) than that of conventionalized metaphors but still relevant to their presence in sports discourse (see Tables 3 and 4).

The relation between the conventionalized and semi-conventionalized metaphors can be viewed as highly indicative of the relation between the conceptual metaphors SPORT IS WAR and SPORT IS FORCE. As both categories are instantiated through relatively stable sets of lexemes, they are the ones which reveal most about the metaphorical structures of the context of culture as background knowledge of the sports discourse.

Almost all of the examples of conventionalized metaphors listed in table 2 are metaphorical expressions based on the conceptual metaphor SPORT IS WAR. This shows us how deeply embedded the conceptual metaphor SPORT IS WAR is in the Croatian sports discourse. If we return to our discussion of conventionalized metaphors as key-words of the sports discourse it can be said that viewing sport as WAR is one of the commonest ways of understanding sports in Croatian culture. On the other hand, it is clear that the conceptual metaphor SPORT IS FORCE is not as conventionalized in our understanding of sports as the conceptual metaphor SPORT IS WAR is. This is seen in the fact that metaphorical expressions based on the conceptual metaphor SPORT IS FORCE mostly fall within the category of semi-conventionalized expressions (see Table 3 below). Verbs such as *razbiti* 'to break; to smash', *slomiti* 'to break', *srušiti* 'to knock over', *potopiti* 'to sink' all refer to

Parasynonyms of 'to win'	No. of tokens	Relative frequency
razbiti 'to break; to smash'	98	0.0081%
pokoriti 'to conquer'	62	0.0051%
potopiti 'to sink'	45	0.0037%
poraziti 'to defeat'	35	0.0029%
demolirati 'to demolish'	32	0.00267%
pomesti 'to swipe out'	29	0.0024%
srušiti 'to knock over'	26	0.00217%
pregaziti 'to run over; to stomp on'	22	0.00183%
slomiti 'to break'	21	0.0018%
nokautirati 'to knock out'	19	0.00158%
razvaliti 'to destroy; to break'	10	0.0006%

Table 3. Verbal semi-conventionalized metaphors. Parasynonyms of *pobijediti* 'to win' and their frequencies

Parasynonyms of 'match'	No. of tokens	Relative frequency
sukob 'conflict'	91	0.0076%
sraz 'collision'	90	0.0075%
bitka 'battle'	81	0.0067%
sudar 'clash; crash'	62	0.0051%
okršaj 'skirmish; clash'	30	0.0025%

Table 4. Nominal semi-conventionalized metaphors. Parasynonyms of *utakmica* 'match'

the kind of force (natural or physical).[22] On the other hand nominal expressions that fall within this category are instantiations of the conceptual metaphor SPORT IS WAR which points to the interesting distribution between nominal and verbal expressions as instantiations of the two conceptual metaphors.

[22] Although Kövecses (2003) distinguishes between natural and physical force, we do not consider such a distinction relevant for our analysis. Second, we consider that this distinction is not clear cut and as such not plausible for our research.

With respect to the analysis of linguistic expressions it has become obvious that the status of the two conceptual metaphors is not equal in structuring the sports discourse and thus do not represent in the same way background knowledge representative for understanding sports in Croatian culture. Therefore, one must regard conceptual metaphor SPORT IS WAR as more entrenched and more conventional in structuring sports discourse in Croatian language and culture. Conversely, it tells us that the conceptual metaphor SPORT IS FORCE has become more and more salient in the way we conceptualize sports, producing novel metaphorical expressions, thus becoming more and more entrenched in the sports discourse. Verbs such as *savladati* 'to overcome' or *izbaciti* 'to throw out' have become conventionalized as to their frequencies in the SDC (see Table 2) and are also losing the feature of affective markedness.

As we have already mentioned, based on our analysis, nominal metaphorical expressions are instantiations of the conceptual metaphor SPORT IS WAR. This could be explained by the fact that the novelty and metaphorical productivity is not exclusively related to the conceptual metaphor SPORT IS FORCE, but to the conceptual metaphor SPORT IS WAR as well. The category of semi-conventionalized metaphors points to the fact that once a certain conceptual metaphor is considered to be a fully conventionalized structure of knowledge (due to expressions that are instantiated in a certain discourse) its productivity and motivating input could be variable. Based on the analysis of Croatian sports discourse, it has become evident that the conceptual metaphor SPORT IS WAR motivates new metaphorical expressions only within the nominal lexical category. Thus semi-conventionalized nominal expressions such as *sukob* 'conflict', *bitka* 'battle', and *okršaj* 'skirmish' are related to the domain of WAR (on the basis of their lexicographic definitions), whereas two metaphorical nominal expressions that are connected to the domain of FORCE are *sudar* 'crash' and *sraz* 'collision'.

The reason why we introduced the novel category of semi-conventionalized metaphorical expressions is that some metaphorical expressions are a constant of sports discourse, appearing regularly, less frequently than conventionalized metaphors and far more frequently than innovative metaphors. Further, they do not function as key-words. They mostly function as parasynonyms of key-words with a high degree of affective markedness. The expression *potopiti (protivnika)* 'to sink (an adversary)' is much more affectively marked than the schematic expression *pobijediti (protivnika)* 'to defeat (an adversary)'. With respect to the individual use criterion, these expressions are more innovative than conventionalized expressions but are becoming more and more regular in sports discourse.

What is important about semi-conventionalized metaphors is that the bulk of the expressions in this category exhibit different patterns of their use in the discourse. These patterns become pertinent through the comparison of the SDC and a general corpus. The Croatian National Corpus (CNC) ver. 2.0, a general corpus consisting of 100 million tokens, is to some degree diachronic[23] in comparison to the SDC, which consists of more recent sports texts. Comparing the two corpora, the difference

[23] The CNC ver. 2.0 consists of texts from 1990 to 2005, which makes our data collected after that period comparable to the CNC in a diachronic way.

between the relative frequencies of certain words in the two corpora has become evident. The relative frequency of the word *okršaj* 'skirmish' in the SDC is 0.0025% and its relative frequency in the CNC is 0.0011%.[24] This statistical data show us that the expression *okršaj* 'skirmish' has undergone a process of conventionalization. To be more precise, the noun *okršaj* 'skirmish' appears in 30 citations in every one million tokens in the SDC, whereas it appears in 11.79 citations in every 1 million tokens in the CNC. This data shows that the word *okršaj* 'skirmish' has become twice as frequent over a period of twenty-five years and thus more conventionalized and entrenched in our background knowledge.

Verbs such as *pomesti* 'to sweep up' and *demolirati* 'to demolish' exhibit similar patterns in the discourse formation. The verb *pomesti* 'to sweep up' in the metaphorical sports sense appears in 50 citations in the CNC. This means that its relative frequency in the CNC is 0.00005%, which classifies this metaphorical expression as being in the category of innovative metaphors because its appearance is totally insignificant. However, its relative frequency is significantly different in the SDC. In the SDC, its relative frequency is 0.0024%, which means that it has become more conventionalized than it was in the CNC. This means that the frequency that the verb *pomesti* 'to sweep up' in the CNC is 0.5 citations which is statistically irrelevant data, contrary to the data obtained by the SDC (where there are 29 citations of the verb *pomesti* 'to sweep up'.

The verb *demolirati* 'to demolish' has undergone an even more significant process of conventionalization. Its relative frequency is 0.0026% in the SDC and 0.000013% in the CNC. This shows how the verb could be categorized as a highly innovative metaphorical expression in the CNC, appearing in a negligible number of citations in every 1 million tokens (0, 13 times), whereas in the SDC it appears in 32 citations. Thus comparison of the general and specialized corpora gives us very precise evidence of how a metaphorical expression became more conventionalized over twenty-five years.

Semi-conventionalized metaphors exhibit some interesting features with respect to qualitative data as well. They function mutually as parasynonyms and in relation to some conventionalized metaphorical expressions in the sports discourse. As seen in the examples in tables 3 and 4, most of the lexemes have very similar meanings with different affective markedness and can be replaced by one another in the same context. In other words, a speaker can choose one or the other as part of his own stylistic motivation. For example, when talking about a match between two teams or players, a speaker can report the outcome of the game with various lexemes, e.g.

(2) Ljubo srušio Nadala (SDC, www.vecernji.hr)
'Ljubo knocked down Nadal' (tennis),

(3) Lyon razbio Bordeaux (SDC, www.vecernji.hr)
'Lyon smashed Bordeaux' (football),

(4) Milan potopio Genovu (SDC, www.jutarnji.hr)
'Milan sank Genoa' (football).

[24] This is the relative frequency of the word *okršaj* used exclusively in sports discourse in the general corpus. The method was to manually separate meanings related to sports from those related to other domains.

These can be freely replaced with one another and still the main meaning of defeat will be properly denoted, e.g.

(5) Milan zaustavio / slomio / potopio / nokautirao Genovu
'Milan stopped / broke / sank / knocked out Genoa'.

The choice seems to be left to the speaker, as he can choose from a set of parasynonymous lexemes with more or less the same denotational meaning. It could be noticed that they are affectively marked when compared to the affectively neutral conventionalized expressions *pobijediti* 'to win' and *poraziti* 'to defeat', e.g.

(6) Milan porazio Genovu
'Milan defeated Genoa'

since these say nothing about the gravity of the defeat itself as seen through the eyes of the speaker. The same is true of the nominal expressions *sukob* 'conflict', *sraz* 'collision', *bitka* 'battle', and *okršaj* 'skirmish', which are parasynonyms to the word *utakmica* 'match'.

5.4. Innovative metaphors

The term *innovative metaphor* refers to a metaphorical expression based on a conventionalized conceptual metaphor, in our case SPORT IS WAR and SPORT IS FORCE. Innovative metaphors have extremely low frequencies, appearing less than 10 times in the SDC, and are examples of the creative, individual use of language by the speaker. Because of this, they have the highest degree of affective markedness.

Within the category of innovative metaphors we find expressions such as *otpuhati* 'to blow away', *torpedirati* 'to torpedo', and the verb *razvaliti* 'to ruin', which appears 10 times in the SCD. Therefore, the latter could be considered as a peripheral member between semi-conventionalized and innovative metaphors. The expressions *otpuhati* 'to blow away' and *razvaliti* 'to ruin' are related to the conceptual metaphor SPORT IS FORCE, whereas the expression *torpedirati* 'to torpedo' is related to the domain of WAR, expressing in its literal sense the way a certain weapon is used to defeat an enemy.

An interesting property of innovative metaphors in sports discourse that should be pointed out is the fact that very often these expressions are complex, and are produced either by combining novel metaphorical expressions from both conceptual metaphors or by introducing new elements into the utterance. The following examples will illustrate the way in which innovative metaphors as complex instantiations are produced in the sports discourse:

(7) Amerikanci su Kanađanima bacili rukavicu uraganskom prvom trećinom protiv Finaca (SDC, www.vecernji.hr)
'The Americans threw down the glove before the Canadians with a hurricane-like first third against the Finns',

(8) Branitelj naslova maršira do finala (SDC, www.jutarnji.hr)
'The defender of the title marches to the final',

(9) Igra Danaca je kao mlin koji melje (SDC, www.sportskenovosti.hr)
'The Danes' game is like a grinding mill'.

Example (7) is a combination of two differently motivated metaphorical expressions. The first part is the existing idiomatic metaphor *baciti (INF) rukavicu rukavicu* 'to throw down the glove' not connected to any specific type of discourse but to the more general everyday use of language, although it must be noted that it also originally relates to the domain of WAR. The second, *uraganska prva trećina* 'hurricane-like first third' is directly based on the conceptual metaphor SPORT IS FORCE, as represented by the force of the hurricane, i.e. natural force. To use hurricanes to specify the type of strength involved in the game is an innovative way of talking about the match in itself. The innovative metaphorical expression is thus a combination of an existing metaphor (*baciti rukavicu* 'to throw down the glove') introduced into sports discourse in a novel way and a metaphorical expression based on a conceptual metaphor already present in the sports discourse but instantiating a novel concept: a hurricane.

Example (8) is a combination of an existing conventionalized metaphor in the sports discourse *branitelj naslova* 'the defender of the title', which is combined in a new way with the expression *maršírati* 'to march'. The literal sense of the verb is related to the domains of MOVEMENT and ARMY because it refers to a special, brusque, energetic way in which an army moves. In sports discourse this verb is a novel metaphorical expression and thus introduces some new meanings and new ways of understanding sports.

Example (9) is a case of a purely individual introduction of elements from general language use which changes their meaning simply because of their placement in a discourse which they are not normally a part of. They are still not randomly selected because their understanding in this context is based on and motivated by the conceptual metaphor SPORT IS FORCE, in this case the strength being the force of grinding.

These examples show us how novel expressions can be drawn into a discourse from: (a) already existing conceptual metaphors or (b) other domains that are peripherally related to existing conceptual metaphors (e.g. the verb *maršírati* 'to march').

6. Some concluding remarks

The main goal of our research was to show how Conceptual Metaphor Theory could be integrated into discourse analysis and what it can reveal about the way speakers structure discourse. We focused on sports discourse for two reasons: (1) sports are highly metaphorically understood and (2) sports are commonly used as a source domain for understanding other concepts; therefore, we wanted to conduct a research to see what domains mostly serve as source domains in understanding sports.

The theoretical framework used for this research was in correlation with Malinowski's (1923) distinction between context of situation and context of culture,

as well as Lakoff's (1992) statement that speakers understand every situation metaphorically and accordingly have an ability to communicate about this situation using different linguistic expressions. Our analysis showed that the two most prominent conceptual metaphors used in understanding sports are SPORT IS WAR and SPORT IS FORCE because most of the metaphorical expressions found in sports discourse are instantiations of these two metaphors. However, corpus based analysis showed that these two metaphors, although prominent, are not entrenched in speakers' background knowledge to the same degree. This statement was enabled by using corpus based analysis.

Implementation of corpus based analysis of sports discourse appeared to be very useful in pointing to some new evidence in the way conceptual metaphors participate in the formation of a discourse as part of background knowledge. Our approach to corpus-based analysis was somewhat different from already-existing research (cf. Cameron and Deignan 2003; Charteris-Black 2004; Deignan 2005), since we used a specialized digitalized corpus consisting of more than 1 million tokens for discourse analysis. Corpus analysis based on the SDC pointed to some interesting evidence on how metaphorical expressions are used when communicating about a certain situation. Unlike Deignan (2005), we used a relative frequency approach to show the number of times a certain metaphorical expression appears in relation to the absolute number of tokens in a corpus. This provides evidence of the representation of metaphorical expressions in (sports) discourse.

This enabled us then to propose a modified classification of metaphors with semi-conventionalized metaphors as an intermediate category. The three criteria used in the classification of metaphorical expressions are characterized as quantitative and qualitative but are regarded as a coherent, inseparable group. Frequency strongly correlates with the other two qualitative criteria. This means that expressions with the highest frequency function as key-words of sports discourse, lacking affective markedness.

The category of semi-conventionalized metaphors was introduced into the continuum from conventionalized to innovative metaphors because, based on our analysis, it has become evident that there is a set of metaphorical expressions that differ from both conventionalized and innovative metaphors.

Furthermore, with respect to two principal conceptual metaphors that structure the background knowledge of sports in Croatian culture, it has become clear that they do not participate in the same way in the structuring of sports discourse. Firstly, metaphorical expressions related to the conceptual metaphor SPORT IS FORCE exhibit lower frequency than expressions related to the conceptual metaphor SPORT IS WAR. Secondly, they function as parasynonyms either of some conventional metaphors (mostly to the verb *pobijediti* 'to win') or of each other.

Innovative metaphors are expressions that are statistically irrelevant in a corpus exhibiting a high degree of the speaker's individual and creative use. In sports discourse they are often complex metaphorical expressions structured from conventionalized metaphors and completely innovative metaphorical expressions sometimes related to less entrenched domains.

The proposed analysis showed some aspects of metaphorical discourse formation that are in high correlation with Lakoff's (1990) assertions. It gives some new empirically (statistically) corroborated evidence on how conceptual metaphors, in our case SPORT IS WAR and SPORT IS FORCE, structure the speaker's background knowledge about a certain situation (sports within the Croatian speech community), what their instantiation are, and how they function in structuring (sports) discourse.

As we mentioned before, we consider this kind of analysis based on the SDC approach suitable for the analysis of any other kind of discourse. This approach thus paves the way for further anthropological, linguistic, and sociological research stemming from discourse analysis.

Bibliography

Anić V. 1996. *Rječnik hrvatskoga jezika*. Zagreb.

Bairner A. 2001. *Sport, nationalism and globalization: European and North American perspectives.* Albany.

Blain N., Boyle R., O'Donnell H. 1993. *Sport and national identity in the European media*. Leicester.

Bybee J., Hopper P. 2001. Introduction to frequency and the emergence of linguistic structure. -Bybee J., Hopper P. (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam: 1–27.

Callies M. 2011. Widening the goalposts of cognitive metaphor research. – Callies M., Keller W.R., Lohöfer A. (eds.). *Bi-directionality in the cognitive sciences: Avenues, challenges and limitations*. Amsterdam: 57–81.

Cameron L., Deignan A. 2003. Combining large and small corpora to investigate tuning devices around metaphor in spoken discourse. – *Metaphor and Symbol* 18.3: 149–160.

Charteris-Black J. 2004. *Corpus approaches to critical metaphor analysis*. Hamsphire.

Coseriu E. 1973. *Sincronía, Diacronía e Historia – El problema del cambio lingüístico*. Madrid.

Deignan A. 2005. *Metaphor and corpus linguistics*. Amsterdam.

Foucault M. 1972. *The archaeology of knowledge*. London.

Györi G. 2002. Semantic change and cognition. – *Cognitive Linguistics* 13.2: 123–166.

Hodge R., Cress G. 1988. *Social semiotics*. New York.

Johnson M., Lakoff G. 2002. Why cognitive linguistics requires embodied realism. – *Cognitive Linguistics* 13.3: 245–263.

Kövecses Z. 1986. *Metaphors of anger, pride and love: A lexical approach to the structure of concepts*. Amsterdam, Philadephia.

Kövecses Z. 2000. *Metaphor and emotion: Language, culture, and body in human feeling. Studies in emotion and social interaction*. Cambridge.

Kövecses Z. 2002. *Metaphor: A practical introduction*. New York.

Kövecses Z. 2003. The scope of metaphor. – Barcelona A. (ed.). *Metaphor and metonymy at the crossroads: A cognitive perspective*. Berlin: 79–92.

Kövecses Z. 2005. *Metaphor in culture: Universality and variation*. Cambridge.

Kövecses Z., Palmer G.B., Dirven R. 2003. Language and emotion: The interplay conceptualisation with physiology and culture. – Dirven R., Pörings R. (eds.). *Metaphor and metonymy in comparison and contrast*. Berlin: 133–160.

Lakoff G. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago.
Lakoff G. 1990. Metaphor and war: The metaphor system used to justify war in the Gulf. [Distributed by electronic mail. Reprinted in Kreisler H. (ed.). (1992). *Confrontation in the Gulf: University of California professors talk about the war*. Berkeley.].
Lakoff G., Johnson M. 1980. *Metaphors we live by*. Chicago.
Lakoff G., Turner M. 1989. *More than cool reason: A field guide to poetic metaphor*. Chicago.
Langacker R. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford.
Leech G. 1974. *Semantics*. London.
Mahon J.E. 1999. Getting your sources right: What Aristotel *didn't say*. – Cameron L., Low G. (eds.). *Researching and applying metaphor*. Cambridge: 69–80.
Malinowski B. 1923. The problem of meaning in primitive languages. Supplement 1. – Ogden C.K., Richards E.A. (eds.). *Meaning of meaning*. London: 451- 510.
Matoré G. 1953. *La méthode en lexicologie. Domaine français*. Paris.
Nerlich B., Clarke D. 1988. A dynamic model of semantic change. – *Journal of Literary Semantics* 17: 73–90.
Lee A., Poyton C. 2000. Culture and text: An introduction. – Lee A., Poyton C. (eds.). *Culture and text discourse and methodology in social research and cultural studies*. Lanham: 1–18.
Raffaelli I. 2009. *Značenje kroz vrijeme. Poglavlja iz dijakronijske semantike*. Zagreb.
Skorczynska H., Deignan A. 2006. Readership and purpose in the choice of economics metaphor. – *Metaphor and Symbol* 21.2: 87–104.
Sinclair J. 2001. Preface. – Ghadessy M., Henry A., Rosberry R.L. (eds.). *Small corpus studies and ELT – theory and practice*. Amsterdam: 57–81.
Stanojević M. 2013. *Konceptualna metafora: Temeljni pojmovi, teorijski pristupi i metode*. Zagreb.
Steen G.J. 1999. Metaphor and discourse: Towards a linguistic checklist for metaphor analysis. – Cameron L., Low G. (eds.). *Researching and applying metaphor*. Cambridge: 81–105.
Steen G.J. 2007. *Finding metaphor in grammar and usage*. Amsterdam.
Sterne G. 1931. *Meaning and change of meaning*. Bloomington.
Sweetser E. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantics*. Cambridge.
Šonje J. 2000. *Rječnik hrvatskoga jezika*. Zagreb.
Ullmann S. 1962. *Semantics: An introduction to the study of meaning*. Oxford.
Wierzbicka A. 1997. *Understanding cultures through their key words: English, Russian, Polish, German, Japanese*. New York.
Wright J. 2000. Disciplining the body: power, knowledge and subjectivity in a physical education lesson. – Lee A., Poyton C. (eds.). *Culture and text discourse and methodology in social research and cultural studies*. Lanham: 152–169.
Žic Fuchs M. 1992. Konvencionalne i pjesničke metafore. – *Filologija* 20.2: 585–593.