To *b or not to *b
Proto-Indo-European *b in a phylogenetic perspective
Olander, Thomas

Published in:
Historische Sprachforschung

Publication date:
2020

Document version
Peer reviewed version

Citation for published version (APA):
Olander, T. (Accepted/In press). To *b or not to *b: Proto-Indo-European *b in a phylogenetic perspective. Historische Sprachforschung.
To *b or not to *b

Proto-Indo-European *b in a phylogenetic perspective

Thomas Olander

Everyone knows about the rarity of “IE *b”.
Hamp 1954: 40

It is widely recognised that *b is a rare phoneme in the Proto-Indo-European stop system, but there is no agreement as to the reality of the few alleged examples of Proto-Indo-European *b. This paper approaches the problem from a phylogenetic perspective, assuming that Anatolian and Tocharian were the first two subgroups to separate from the rest of the Indo-European language family. It is concluded that there is no evidence at all for the existence of Proto-Indo-European *b in initial position. In medial position there are a few possible examples, although none of them can be regarded as certain.

1 Introduction

The apparent rarity of *b is a classical problem of comparative Indo-European linguistics.1 Any respectable treatment of the Proto-Indo-European phonological system mentions that the voiced unaspirated bilabial stop *b was found in very few morphemes. Clackson (2007: 42), for instance, provides an overview of the frequency of the various consonant phonemes in the first edition of the Lexikon der indogermanischen Verben (LIV¹), listing *b, with 14 occurrences, on a 25th place out of 25 consonant phonemes, preceded by *gʷʰ, with 17 occurrences. Similar countings of occurrences of *b have been made for Indogermanisches etymologisches Wörterbuch (Pokorny 1959/1994) by Jucquois (1966), Jahowkyan (1982: 60) and Matasović (1994: 141 with n. 11, referring to an unpublished study from 1991).

The rarity of PIE *b has been used as an argument for a radical reinterpretation of the phonological system of Proto-Indo-European. Some authors argue that the phonemes traditionally considered to be voiced unaspirated stops, *b d ţ g gʷ, were actually glottalic, so that in fact it is

1 This study is based on a paper presented at the 38th East Coast Indo-European Conference, University of Pennsylvania, 21–23 June 2019. I am very grateful to the organiser of the conference, Don Ringe, to the participants in the discussion, in particular Jay Jasano and Sara Kimball, who directed my attention to additional material, and to Craig Melchert, who generously sent me his thoughts on the etymological connections between various roots possibly pointing to PIE *b. While preparing the written version I have received kind help from Birgit Anette Olsen and Don Ringe. I am also indebted to Martin Joachim Kümmel for his helpful observations and corrections to the manuscript, including numerous additional references, during the review process. Frederik Kortlandt has kindly sent me a number of useful comments, corrections and references. Finally, I am grateful to Henning Andersen for commenting the paper and providing the pun that ended up as the title of the study.
not *[b] that is rare, but rather *[p’], a situation which is assumed to be rather common cross-linguistically. Thus the question of the existence of PIE *b is of significance for the reconstruction of the Proto-Indo-European phonological system as such.

Other scholars explain the rarity of *b by assuming that the pre-Proto-Indo-European segment that would have yielded PIE *b merged with another segment such as *m (Schindler 1972: 3, retracted apud Mayrhofer 1986: 103; Kümmel 2012: 304), *wu (Sihler 1995: 146–147; Barrack 2005: 238; Weiss 2009b; Kümmel 2012: 304), *bh (Barrack 2005: 238–239) or *p (only in initial position: Thurneysen 1908: 153, followed by Kortlandt 1985/2010a: 198; but cf. Pedersen 1951: “Ich wüsste aber kein anderes Beispiel dafür anzuführen, dass b-sich in irgend einer Sprache nicht parallel mit d- und g- entwickelt hätte”; similarly Kümmel 2007: 332 n. 260; as Frederik Kortlandt reminds me, Pedersen’s objection would not apply if *b was a glottalised [p’]). As far as I am aware, with very few exceptions the problem of Proto-Indo-European *b has not been analysed on the background of a binary-branching family tree, where the branches do not separate all at once but leave the stem one after another.

In this study I do not give a definite answer to the existence of Proto-Indo-European *b, but I will try to reframe the problem by analysing the material under the hypothesis of a binary-branching tree – specifically, a tree where Anatolian and Tocharian are the first two subgroups to separate, a commonly held opinion among specialists today. This, I hope, will not only lead to a better understanding of the particular problem, but also more generally highlight the benefits and even necessity of a phylogenetic approach to problems of Indo-European reconstruction.

Before proceeding to the material and the analysis, we will have a look at some of the more significant takes on the problem from the last two centuries.

2 Brief history of the problem

The field of comparative Indo-European linguistics was only beginning to establish itself as a discipline when Jacob Grimm, in the first volume of the second edition of his Deutsche Grammatik, discussed Old High German f and concluded that “im anlaut kommt es, gleich dem goth. p, selten vor, meistens in fremden wörtern” (1822: 131; cf. Weiss 2009a/2011: 34 n. 18).

The first reference to the actual Proto-Indo-European situation that I have been able to find is in the second volume of Pott’s Etymologische Forschungen from 1836. Pott stated that "Unter den Labialen erscheint b selten zu Anfange oder am Schlusse mit Wurzeln, und wechselt noch dazu oft mit w" (1836: 454; cf. Wackernagel 1896: 182).

In the first volume of the Compendium der vergleichenden grammatic der indogermanischen sprachen Schleicher briefly stated that ”ein sicheres beispiel für disen laut kenne ich nicht” (1861: 139).

Grassmann, in the very article in which he published his famous law, also discussed PIE *b:

Es ist aus diesem eigenthümlichen verhältnisse [i.e. the lack of correspondences between words beginning with Skr. b, Gk. b, La. b and Go. p – TO] schon an sich wahrscheinlich, dass der anlaut b, ausser in schallnachahmenden wörtern, vor der sprachtrennung noch gar nicht bestanden habe, und daher die in jenen sprachen mit b (germ. p) anlautenden wörter ursprünglich einen andern anlaut gehabt haben werden. (Grassmann 1863: 122)
So in Grassmann’s opinion there was no PIE *b, at least not word-initially.

At this point the problem of the rarity of PIE *b was common knowledge among linguists. It is therefore no surprise that Bickell in 1865 devoted an entire article to it (in my counting, the first one to focus specifically on this problem). He concluded that “Die angeführten beispiele werden gewiss zu dem beweis genügen, dass die labiale media, so selten sie auch erscheint, doch dem indogermanischen sprachstamm keineswegs gefehlt haben kann.”

And thus already 150 years ago there was a vigorous discussion of the problem, with some scholars leaning towards fewer and others leaning towards more PIE *b’s. It is perhaps of interest to note that in the first edition of volume one of his Grundriss, Brugmann stated that “Dieser Laut war von den Verschlusslauten der idg. Urzeit wol der seltenste” (1886: 266), but in the second edition he does not mention the rarity of *b (cf. 1886/1967: 507; see also Brugmann 1904: 150).

Uhlenbeck (1894) treated the problem from a Germanic perspective, and a few years later Johansson (1900, with further references) provided a long list of words containing, in his opinion, word-initial PIE *b (but cf. Mayrhofer 1983: 146 n. 98: “doch ist diese drastisch einzuschränken”). Meillet (1903: 57) was more sceptical:

Le *b est relativement rare; il ne figure dans aucun suffixe important ni dans aucune désinence; il semble secondaire dans une partie des mots où on le rencontre […]; d’autres mots sont initiatifs […]; d’autres sont limités à peu de langues et ont l’air d’emprunt récents.


More than a century after the initial formulation of the problem, Pedersen (1951), at the age of 84, revitalised the interest. Departing from the observation that “*b’s ist fast unmöglich, ein gemeinindoeuropäisches Wort mit anlautendem *b aufzutreiben” (p. 10), he proceeded to a discussion of “das Fehlen des *b im gemeinindoeuropäischen Klusilsystem” (p. 14) and “die Seltenheit des gemeinindoeuropäischen *b” (p. 16). Pedersen concluded that the Proto-Indo-European voiced unaspirated series reflects an earlier unvoiced series, where the absence of *p (which would then have yielded PIE *b) would be typologically common.

Pedersen’s analysis formed the basis of Gamkrelidze and Ivanov’s (1973) reinterpretation of the Proto-Indo-European stop system. In their view the absence of *b is easier to understand if the traditional voiced unaspirated stops were glottalised. Initially stating that “die stimmhafte Serie [gilt] als defektiv, da in ihr das labiale stimmhafte /b/ fehlt bzw. höchst selten vorkommt” (1973: 151), on the next page they discussed the gap in the voiced series “wegen des Fehlens des labialen Verschlusses /b/” (cf. also Gamkrelidze 1981: 604: “we can thus speak of the total absence, of a gap, in the place of /b/ in the voiced series”).

Also in 1973, Hopper stated that “[T]he system [...] displays a surprising gap in the labial series, the plain voiced stop */b/ being absent.” Similar formulations are given by Bomhard (“The three stop system is complete in the dental, palatal, velar, and labiovelar series, but the labial series is defective, lacking a plain voiced member”, 1977: 56) and by Bybee (“[T]he traditional reconstruction of the PIE obstruent system has a gap in it, in that it seems to lack a *b, although *p and *bh are present”, Hooper 1983: 23–24). Thus for many of those arguing for a glottalic series
in Proto-Indo-European, the absence of *b quickly became an established fact (for a critical overview see Szemerényi 1985: 11–12; but cf. Kortlandt 1985/2013a: 198, with an overview of glottalic theories pp. 183–185; see also Kortlandt 2018a; forthc.).

Motivated by the glottalicists’ attempts to erase *b from the Proto-Indo-European phonological system, Jəhowkyan (1982; 1986) tried to rehabilitate the consonant. Similarly, Meid (1989: 6) was of the opinion that “das grundsprachliche Vorhandensein von /b/ [steht] außer Frage”, although the sound mostly occurred in words belonging to a lower register: “Im Vergleich zu dem kräftig gespannten *p und dem beschwingten, federnden *bʰ war *b ein weicher, aber stumpfer Laut” (p. 8). Meid concluded that “[e]ine ‘Leerstelle’ im System hat es nicht gegeben, eine derartige Ansicht ist ein Mythos, zu dessen Überwindung diese Ausführungen beitragen wollen” (p. 13). Approximately at the same time, Westcott (1988) put forward a similar view: many of the roots containing PIE *b are “expressive in character and […] disparaging in tone” (p. 365). Barrack (2005) provided typological arguments against the rarity of *b as evidence for a glottalic system.

Note that for those who depart from the traditional stop system it is possible to argue for the theoretical existence of *b on the basis of typological considerations: since the phonological system of Proto-Indo-European contained *d̑g g* gʷ as well as *p t k kʷ and *bʰ dʰ gʰ gʷ, it must also have contained *b. I have not seen anyone argue explicitly for this stance – in a way the opposite reasoning of that of some glottalicists – but it may in fact be underlying some of the anti-glottalicists’ point of view when they, somewhat uncritically at times, adduce a large amount of material allegedly pointing to PIE *b (e.g. Meid 1989).

3 The Indo-European family tree and the phylogenetic approach

On the basis of a partly different selection of material from that presented in §§ 4 and 5 in this study, Matasović (1994: 142) remarks that

most of the examples are attested in the so-called ‘Old European’ languages (Italic, Germanic, Celtic, and Balto-Slavic). Indeed IE *b seems to be particularly [sic] abundant in Germanic and Balto-Slavic, two branches of the IE family that are genetically closely related.

It is not clear to me what the occurrence of most examples of *b in “Old European” languages implies. Nor is a close genealogical relationship between Germanic and Balto-Slavic an established fact: the similarities between the two subgroups are likely to be the result of later convergence (cf. Kortlandt 2018b: 224–225). Even more debatable is Matasović’s (1994: 144 n. 18) “southern” group of dialects comprising Italic, Greek, Indo-Iranian and perhaps Anatolian, which is distinct from “another dialect group (comprising all the other IE languages).” But methodologically I believe the way of thinking is a step in the right direction.

The line of reasoning I will utilise in this study is more directly employed by Ringe in his book From Proto-Indo-European to Proto-Germanic, where he states that “[t]he voiced bilabial stop *b was unexpectedly rare […] though at least one example has reflexes in Anatolian and so can be reconstructed even for PIE proper” (Ringe 2006/2017: 9, where “PIE proper” refers to the ancestor of all Indo-European branches, including Anatolian). The example mentioned by
Ringe is "leb- 'to lick' (see § 5.7); in the first edition of the book Ringe also mentioned "h₂ébō 'river' (2006: 8), but in the second edition he reconstructs this word as "h₂ábʰon- (2006/2017: 9 n. 5; see also Kümmel 2007: 332 n. 260). In Olander 2018: 189–190 I briefly mentioned the status of "b as an example showing how the Indo-European family tree may help us towards a better understanding of various problems in reconstruction. Neither Ringe nor I provide a discussion of more than a few words of relevance for the problem of PIE "b. I therefore think it is worthwhile to analyse the problem more thoroughly.

Let us first have a look at the Indo-European family tree. Most people nowadays agree that Anatolian was the first subgroup to separate. There is also some agreement that Tocharian was next. What happened afterwards is less certain, but Italo-Celtic is a likely candidate to be the third group to branch out. In the further ramification it is possible, though by no means certain, that Germanic was the next subgroup to separate, followed by Greek and the satam subgroups: Armenian, Albanian, Indo-Iranian and Balto-Slavic (see fig. 1, based on Ringe, Warnow & Taylor 2002; Nakhleh, Ringe & Warnow 2005; cf. also Kortlandt 2010b; 2016; for the nomenclature of the first-order nodes see Olander 2019).

While most of the internal nodes in this tree are highly debatable, it does seem to be, at the current state of our knowledge, the best binary-branching model we have for the Indo-European language family (for alternative binary-branching trees see e.g. Holm 2008: 635; Bouckaert et al. 2012: 959; Chang et al. 2015: 200).

It is, of course, possible to assume that the first-order branching of Indo-European did not consist exclusively in binary splits – indeed, this is what many treatments of Indo-European linguistics more or less tacitly presuppose (but cf. the instructive considerations in Hale 2007: 238–239). Fortunately, in this study we are mainly concerned with the first two branchings, which happen to be the least controversial: that of Proto-Indo-European into Anatolian and Indo-Tocharian, and that of Indo-Tocharian into Tocharian and Indo-Celtic. Thus for those who do not agree with respect to the internal structure of the Indo-Celtic subtree, but do agree that Anatolian and Tocharian separated first, the premises of this study should be largely acceptable.
And even for those who do not agree on any parts of the tree the methodological aspect should be worth taking into consideration.

If, in a binary-branching tree, we wish to examine the proto-language of the entire tree – in our case, Proto-Indo-European – we must assess the various linguistic features of the individual subgroups and determine whether they are innovations or archaisms. In the case of lexical items, which can easily and very often undetectably be replaced, the only way their existence in the proto-language can be secured is if cognates are found in the outlier of the family – in our tree, Anatolian – and at least one other subgroup. Generally speaking, any lexical item can only be reconstructed as far back as the node where the subgroups in which it has reflexes meet. An item with reflexes in e.g. Tocharian, Armenian and Balto-Slavic can be assumed to have existed in Proto-Indo-Tocharian in the model shown in fig. 1; but we cannot securely reconstruct it back to Proto-Indo-European.

For our problem this reasoning is relevant because the reconstruction of any phoneme in a proto-language depends on the existence of morphemes containing the phoneme. And since, as pointed out already by Meillet (1903: 57), *b cannot be reconstructed for any suffix or ending in Proto-Indo-European, the reconstruction of *b depends entirely on lexical items. Accordingly, in order to reconstruct *b in Proto-Indo-European we must have lexical cognates in Anatolian and at least one other subgroup, and in order to reconstruct *b in Proto-Indo-Tocharian, there must be lexical cognates in Tocharian and at least one other subgroup, etc.

This phylogenetic approach to linguistic reconstruction will be applied in the analysis of the material in §§ 4 and 5.

4 *b in initial position

4.1 Introduction

It is sometimes stated that no Proto-Indo-European root begins with *b (cf. e.g. Szemerényi 1985: 11–12: ‘initially b is rare, perhaps not to be acknowledged at all; but internally it is vigorously represented’; 1996/1999: 145). And indeed, no such roots are listed in NIL (which, of course only presents an “Auswahl” (p. v) of nominal roots) or LIV². It is true that most of the roots with an alleged initial *b are not particularly compelling. For instance, the similar-looking words La. baubor ‘to bark’, Gk. βαβύ (hapax) ‘bow-wow’ (of a dog’s barking) and Li. baũti ‘to bellow’ (of cows) are more likely to be independent onomatopoeic formations than actual cognates; yet Meid (1989: 9) maintains that “man mit ererbtem b rechnen kann”.

More felicitously Kimball has suggested a relationship between Hi. palša-comm. ‘way; time’, OIr. belach neut. ‘gap; road, path’ and Arm. pelem ‘to hollow out’ (1999: 265, 450; I am grateful to Sara Kimball for pointing this possible connection out to me), pointing to a root PIE *bel- ‘split, cut out, excavate’. However, the distribution of the root across Indo-European subgroups is too sporadic to exclude chance similarity, the semantics are not compelling, and the reconstruction of PIE *b depends exclusively on the Armenian form. I therefore do not include these forms here (cf. also Kloekhorst 2008: 622).

As pointed out to me by Martin Joachim Kümmel, another potential root with word-initial *b is *beu̯-, if Mr. biàs ‘bag; stomach’, OHG pfoso, Gk. βούζα ‘stuff’ (perhaps from *bus-né-) and Alb. m-bush ‘fill’ derive from the same root (Schumacher & Matzinger 2013: 982; Orel 1998: ...
Proto-Indo-European *b, however, as the material collected in Pokorny 1959: 98–102 shows, chances are we are dealing with a subset of a group of formations imitating the act of blowing that are not necessarily genealogically related (Frisk 1960: 277; Beekes’ argument that “this connection fails on the necessity to assume a root with PIE *b-” (2010: 248) would of course be circular in the framework of this study). I exclude this putative root from the survey.

As we have already seen, however, the absence of *b in initial position is by no means an established fact. A few roots that often recur in the discussion seem to rest on safer ground. They are considered in the following subsections.

4.2 *bel- ‘strong’: Proto-Indo-Celtic

- Italic: La. débilis adj. ‘weak’
- Greek: βελτίστος adj. ‘better’, βελτίων adj. ‘better’, βελτίστος adj. ‘best’
- Indic: Ved. bála- neut. ‘power, strength’, bályas- (Atharvaveda+) adj. ‘stronger’, báliṣṭha- (Brāhmaṇa+) adj. ‘strongest’
- Slavic: OCS bol’ii adj. ‘bigger, better’

At first glance the evidence points unequivocally to a (nominal) root *bel-, and this reconstruction is often regarded as Proto-Indo-European (e.g. Meid 1989: 6; cf. also Mayrhofer 1976: 416–418 with references: “Die Zusammengehörigkeit der angeführten Formen kann nicht bezweifelt werden”; 1983: 146 n. 98 with references). Had it not been for the *b, the Proto-Indo-European status of the root would perhaps never have been questioned. But the rarity of PIE *b has led to the suggestion that La. débilis may be derived from débeò ‘I owe; I ought’, like habilis ‘easy to handle’ from habeò (Pedersen 1951: 11–12; Dieu 2008: 260; Eichner 2013: 26–27 with references). Greek βελτίων etc. may contain PIE *g and be related to βελτίσταν ‘want’ (Pedersen 1951: 12; Dieu 2008: 258–260; cf. also Kölligan 2013: 151–152). Vedic bála- has been derived from a Proto-Indo-European root *uval- presumably also seen in La. ualeò ‘am strong’ (Burrow 1987: 69, citing Grassmann), and from a Dravidian source (Burrow 1946: 19; see also Emenau 1954: 27; cautiously Witzel 1999: 17; Kümmel 2007: 332; but cf. Burrow 1987: 63; Mayrhofer 1986: 99: “ganz ungläubhafte Herleitung”). OCS bol’ii may be related to Gk. ἐξελλω ‘increase’ from the root *h₂bel- ‘increase’ (Dieu 2008). Kortlandt (pers. comm., April 2021) suggests a reconstruction PIE *dbel- ‘[bel], which he derives from the zero grade of the root *dveb- (see the end of § 5.1).

It has been suggested that the root is also attested in Iranian as Ossetic bal ‘group’ (for references see Mayrhofer 1996: 215); cf. also the meaning ‘large crowd of individuals; army’ in Kalasha baw, from Sanskrit bāla- (Trail & Cooper 1999: 34). However, as pointed out to me by Martin Joachim Kümmel, Abaev (1958: 232–233) prefers to connect the Ossetic word with *bār- ‘to ride’ (cf. Avestan bar-, which includes the meaning ‘ride’), while Cheung (2002: 170) derives it from a noun *bāra- ‘(also in Modern Persian bār ‘burden; load’ and Ved. bhāra- masc. ‘burden; load’), both ultimately from PIE *bʰer- ‘carry’. Since the etymology of the Ossetic word is uncertain I have not included Iranian in the list.

While each of the individual alternative explanations may be possible, it is in my opinion more economical to assume that the words listed above are cognate. But how far back can we trace the root *bel- in the light of the binary-branching model in fig. 1? Since there are no cognates in Anatolian or Tocharian, even the most ardent supporters of an inherited root (as long as they accept Anatolian and Tocharian as the first two subgroups to separate) can only assign Proto-Indo-Celtic age to it. It cannot be shown that the root existed before that stage, i.e. in
Proto-Indo-Tocharian or Proto-Indo-European itself. If La. *dēbilis* were shown to be unrelated, the root would only be reconstructible back to Proto-Indo-Greek (note that Stifter 2019a: 184 with n. 20 also suggests an Old Irish cognate, remarking, however: “Denkbar ist auch, dass [OIr.] *dibell* überhaupt aus lat. *debilis* [for *dēbilis*] entlehnt ist”).

This conclusion, based on a phylogenetic approach, differs fundamentally from that of e.g. Mayrhofer (1986: 99–100), who regards *bel- as “[e]in in mehreren archaischen Sprachen fortgesetztes, also gemeinindogermanisches Lexem” (following Schindler 1972: 1–2, Mayrhofer 1986: 87 uses “gemeinindogermanisch” to refer to reconstructions “die aus irgendwelchen Gründen für alle Einzelsprachen vorausgesetzt wird”, i.e. corresponding to “Proto-Indo-European” in our terminology).

4.3  *bak- ‘stick’: Proto-Indo-Celtic

- Italic: La. *baculum* ‘stick’
- Celtic: OIr. *bacc* ‘hook’
- ‘Germanic: ?MEng. *pegge* ‘peg’
- Greek: βακτηρία *staff; stick*, βάκτρον *staff; stick*

Another nominal root that recurs in the discussion of PIE *b* in initial position is *bak- ‘stick’. Note that a reconstruction *bh₂k- would also yield the attested forms, eliminating the need to reconstruct PIE *a* (this possibility does not seem to be acknowledged by Beekes 2010: 194: “Is it a European loanword, given the a-vocalism?”; cf. also Matasović 1994: 142 n. 14; Schrijver 1991: 100 argues that since the root may be of non-Indo-European origin, it probably did not contain a laryngeal).

Due to the *b*, the apparent lack of verbal forms of the root (but cf. Frisk 1960: 212), the technical meaning and the uneven geographical distribution of the apparent reflexes, *bak- is often considered not to be Proto-Indo-European; for instance, Mayrhofer (1986: 100) maintains that *bak- belongs to the “nachindogermanische’ Schicht” (cf. also e.g. de Vaan 2008: 67; Beekes 2010: 194; Kroonen 2013: 395). Other authors are more positive towards the word’s status as inherited: Matasović (1994: 143) enumerates it among “the few really convincing etymologies on our list” (cf. also Walde & Hofmann 1906/1938: 92; Frisk 1960: 211–212; Chantraine 1968–1980/1999: 159).

Klingenschmitt (1982: 241) tentatively suggests a reconstruction *gʷak- (again, *gʷh₂k- would also be possible), which would yield the Greek form directly; La. *baculum* would be a borrowing from Oscan or Umbrian, and the Germanic and Balto-Slavic forms would presumably not belong to this root. On the other hand, Arm. *kal* ‘threshing floor’ could be added to the reflexes (see also Olsen 1999: 35–36; I am grateful to Birgit Anette Olsen for supplying this reference).

Like many others I do not consider it particularly likely that the apparent reflexes of a root *bak- or *bh₂k- are inherited, but if, for the sake of the argument, we assume that it is, it can only be reconstructed back to Proto-Indo-Celtic, not to Proto-Indo-European itself, in the tree in fig. 1.
5 *b in non-initial position

5.1 Initial sifting

The evidence for *b in non-initial position is more promising. As a discussion of all suggested instances of PIE *b would take us too far, I base the following overview primarily on the material in LIV² and NIL, supplemented by other handbooks. It seems reasonable, in the context of the present study, to discuss only roots that are represented in either Anatolian, Tocharian or Italo-Celtic plus at least one other subgroup. Thus I do not discuss the following roots reconstructed in LIV² (the question marks have been retained from LIV²):

- ?*dʰreuub-‘to drip’ (LIV² 156): Germanic only
- ?*g/ɢ⁴ebb-‘seize, take’ (LIV² 195): Italic, Celtic
- ?*k/keuubh₂-‘to lie down’ (LIV² 357–358): Italic, Celtic
- ?*stebb-‘to make stiff’ (LIV² 592): Greek, Armenian, ?Baltic
- ?*dʰreb-‘to punch, to hit’ (LIV² 153): Germanic, Slavic
- *kleuub-‘to stumble’ (LIV² 364): Germanic, Baltic, Slavic
- *(s)leb-‘to become limp’ (thus LIV² addenda; LIV² 565 reconstructs ?*sleh₂b-): Germanic, ?Indic, Baltic, Slavic (for putative Celtic reflexes see below)
- ?*steib-‘to sharpen’ (LIV² 566): Germanic, ?Greek
- *(s)leub-‘to turn’ (LIV² 700): Germanic, Greek

Some of these roots may be reconstructed back to the Indo-Germanic stage in fig. 1, but since none of them are present in Italo-Celtic, Tocharian or Anatolian plus another subgroup, we cannot reconstruct them further back in the tree.

For the root *(s)leb-‘to become limp’ possible Celtic reflexes are sometimes adduced (e.g. in LIV² addenda), including OIr. lóbar ‘weak’, Welsh lhwfr ‘cowardly; mean’ and Middle Breton lofr ‘weak, miserly, leprous’. However, as argued by Stifter (2019b), the Welsh form requires, and the remaining Celtic languages are compatible with, PCelt. *lubro or *luφro-, the vocalism of which is not reconcilable with a root PIE *leb-. Instead Stifter suggests, reasonably in my opinion, that the Celtic forms belong to PIE *leuu-p-‘peel, strip’.

Similarly, in some roots the reconstruction of *b is only one possibility, or the relationship between the suggested cognates is too uncertain to be of significance for the present study. For instance, the correspondence set La. lūbris ‘slippery’, Goth. sliupan ‘to slip’, OE slūpan may point to *(s)leuub(h)b- (Weiss 2009a/2011: 34), but a reconstruction *sleuubh₄ is at least as likely (LIV² 567; Meiser 1998/2006: 87; de Vaan 2008: 349).

A root *(s)krebb- ‘to scrape’ is posited in LIV² 562 based on Italic, Celtic, Germanic, Baltic and Slavic material, but it is noted that the absence of Winter’s Law in Balto-Slavic might suggest an alternative reconstruction *(s)krebb₄, in which case the Germanic *p would be secondary (thus Kroonen 2013: 444, 447, followed by Derksen 2015: 407). The presence of *b in this root is thus highly uncertain.

In LIV² 411 a root *(s)lemH- ‘to hang down’ is reconstructed on the basis of Tocharian, Germanic and Indic material. In LIV² addenda, on the other hand, it is stated that ‘Ansatz der Wurzel wird fragwürdig, wenn toch.B lyama ‘saß’ usw. zu trennen sind (s. *h₃lemH-)’, and so it seems safe to leave this root out of the survey.
The root *seiqb- ‘to drip’ (LIV² 521; Malzahn 2010: 947), with alleged reflexes in Tocharian A, Middle High German and Greek, is based on material that is too uncertain to be of relevance for the present study (Beekes 2010: 379 considers a Pre-Greek origin of the Greek word).

Cop (1956) has suggested that the Hittite verb kap- ‘to plan, to conspire’ might be related to a root found in West Germanic, e.g. Osax. tō-hopā fem. ‘hope’, OEEng. hopian ‘to hope’, pointing to a root PIE *keuwa- ‘to think’. I do not find this idea compelling, primarily because of the very limited distribution of the cognates (cf. also Puhvel 1997: 257; see Kluge & Seebold 1883/1999: 379 for alternative cognates to the Germanic root).

And finally, a Proto-Indo-European root *dʰeb- ‘strong’ has been reconstructed on the basis of evidence from Anatolian (CLuw. tapar- ‘to rule’), Germanic (Oltc. dapr ‘firm’; OHG tapfar adj. ‘firm; heavy’; MHG tapfer ‘massive; firm; brave’) and Balto-Slavic (Lv. dabľš adj. ‘strong’; OCS dobľ’s adj. ‘strong’, debels adj. ‘fat’) (Melchert 2003: 19 with n. 18; I am grateful to Craig Melchert for directing my attention to this possible cognate set; as mentioned above, according to Kortlandt (pers. comm., April 2021) *bel- ‘strong’, from *dʰeb-, is derived from *dʰeb-). However, the Anatolian and non-Anatolian forms may not be related at all (thus e.g. Tischler 1991: 118; Kloekhorst 2008: 831), and the reconstruction of PIE *b is only supported by Germanic, whereas the Balto-Slavic forms, with no trace of Winter’s Law, rather point to *bʰ.

A sizeable number of relevant items remain, however. In the following subsections I provide a more detailed analysis of them.

5.2 *h₂ābel- ‘apple’: Proto-Indo-Celtic

- ?Italic: ?Osc. Abella fem. place name
- Celtic: OIr. ubull neut. ‘apple’
- Germanic: Crimean Go. apel ‘apple’; Oltc. epli neut. ‘apple’; OHG apful masc. ‘apple’; OEEng. appel masc. ‘apple’
- Baltic: Li. obuolys masc. a.p. 3a ‘apple’; Lv. ābuõls masc. ‘apple’
- Slavic: OCS abško neut. ‘apple’; OPo. jabło neut. ‘apple’

One of the most prominent *b items is the word for ‘apple’ (it is the only example of PIE *b in Beekes 1995/2011: 164; for further forms, discussion and references see NIL 262–266; Piwowarczyk 2014; de Vaan 2008: 20; Stifter 2019a, for Celtic; cf. also Kroonen 2016 on other Indo-European words for ‘apple’). Some authors adduce the ablaut pattern of the word as evidence for its Proto-Indo-European status (Stifter 2019a: 207; see also Rasmussen 1989: 232–233; NIL 264 n. 1), but it should be remembered that loanwords are sometimes incorporated into existing archaic-looking patterns (thus also Kortlandt (pers. comm., April 2021), despite the formulation in Kortlandt 1985/2010a: 65; cf. e.g. the Germanic loanwords that become ū-stems in Slavic; see Pronk-Tiethoff 2014). On the other hand, many authors consider ‘apple’ to be non-Indo-European, often with reference to the *b (e.g. Markey 1988; Hamp 1979; cf. also Kroonen 2013: 32).

It is possible, though not certain, that reflexes of this word are preserved in East Iranian. As the Iranian forms seem to reflect *amarnā-, we would have to assume a sporadic change of *b to *m in Iranian (e.g. Blážek 1995: 17, with references; cf. also Stifter 2019a: 206–207 with n. 44; by contrast, a change of *m to *b outside of Iranian is suggested by Cheung & Aydemir 2015: 85–86, 88–89). Since Iranian is one of the more deeply embedded branches in our tree, it is not crucial for our purposes if reflexes of *h₂ābel- are attested in Iranian or not.
Potentially of high significance is the assumption that *h₂ábel- has reflexes in Anatolian. Ivanov (1976: 163–162) has put forward the idea that Hi. šam(a)lu- neut., probably meaning ‘apple’ or ‘apple tree’, is a cognate of the non-Anatolian words for ‘apple’. If correct, this would mean that the root would be of Proto-Indo-European age. It appears, however, that the Hittite word cannot be formally reconciled with the non-Anatolian words (Kloekhorst 2008: 713; Kroonen 2016: 87–88; Stifter 2019a: 197); and even if Ivanov’s hypothesis could be shown to be correct, the *b would have arisen from *m after the dissolution of Proto-Indo-European.

As the non-Anatolian ‘apple’ words may be inherited from a common ancestor and do contain *b, they can be traced back to the Proto-Indo-Celtic stage, being attested in Italo-Celtic as well as in other non-Anatolian and non-Tocharian subgroups.

5.3 *treb- ‘to build’: Proto-Indo-Celtic

- Italic: La. trabs fem. ‘wooden beam’; Osc. acc. sg. tríbúm ‘house’
- Celtic: OIr. treb fem. ‘house’; OWelsh treb fem. ‘hamlet, village’
- Germanic: Go. þaurp neut. ‘field’; OSc. þorp neut. ‘village’; OHG dorf neut. ‘village’; OEng. þorp masc. ‘village’
- Baltic: Li. trobà fem. a.p. 3 ‘house, cottage’; Lv. trãba fem. ‘hut, hovel’

Another nominal root that is often mentioned as an example of PIE *b is *treb- (e.g. Jahowkyan 1982: 62; Meiser 1998/2006: 99; NIL 705–706). Again, partly because of the *b, but also due to the geographical distribution of the root, some authors consider the word not to be inherited (Huld 1990: 398; see also de Vaan 2008: 626).

Assuming the forms are inherited from a common ancestor, the root may be reconstructed for Proto-Indo-Celtic.

5.4 *píbe/o- ‘drink (prs.)’: Proto-Indo-Celtic

- Italic: La. bibō ‘drink’; Faliscan fut. 1sg. pipafi ‘will drink’
- Celtic: Gaul. ipv. 2pl. ibetis ‘drink (pl.)’; OIr. ibid ‘drinks’; OWelsh ipf. 1pl. iben ‘we drank’
- Armenian: Arm. əmpem ‘drink’
- ?Albanian: ?OAlb. 3sg. pi=p ‘drinks’
- Indic: Ved. pibati ‘drinks’
- Iranian: Wakhī pōv- ‘drink’ (see Mayrhofer 1963: 252)

The present stem *píbe/o- is attested in several subgroups, and the *b it contains may be original, as argued by Kortlandt (1985/2010a: 65 and apud Beekes 1994: 452–453) on the basis of a suggestion by Thurneysen (1908). Other authors derive the present stem from *pi-ph₃- under the assumption of a presumably pre-Proto-Indo-European development of *ph₃ to *b (e.g. Mayrhofer 1986:100, 143–144; LIV² 462–463: “grundsprachlich”; cf. the discussions in Kümmel 2007: 332 with n. 260; 2012: 325: “the possibly most certain case of PIE *b”).

If it is accepted that the *b was phonemic in *pibēti, the form provides evidence for */b/ at the Proto-Indo-Celtic stage.
The root *dʰeuu̯b- ‘deep’ is amply represented across the Indo-European subgroups (see NIL 122–124). The problem with this reconstruction is whether the root actually contained a *b: Germanic *p seems to point to PIE *b, but Kortlandt (1991/2010a: 177) has suggested that it may rather reflect *bʰ followed by *n in accordance with Kluge’s Law (see also Kroonen 2011: 253, 255; Hill in Hock & al. 2015/2019: 270). The lack of lengthening by Winter’s Law in Balto-Slavic forms like Li. dubùs would also speak against *b and in favour of *bʰ (Kortlandt 1991/2010a: 177).

On the other hand, as pointed out to me by Jay Jasanoﬀ, the reconstruction of the root as *dʰeuub- is rendered diﬃcult by the Tocharian form: since it appears that Grassmann’s Law took place in Tocharian (Winter 1962: 24–25; Ringe 1996: 47; Pinault 2008: 423), a form *dubʰ- > *dub would have yielded TochB, TochA **ts, not the attested t-. As the meaning of the Tocharian words is ‘high’, not ‘deep’ as in the reconstructed form, it is possible that they have a diﬀerent etymology, meaning that the root could only be reconstructed back to Proto-Indo-Celtic. However, as ‘deep’ and ‘high’ can often be used with overlapping meanings (e.g. when used for liquids; see Adams 1999/2013: 296), it seems overly cautious to separate the Tocharian forms from this group. A less convincing argument for original *b is constituted by Gk. βυθός masc. ‘depth (of the sea)’, if from *ṯвуβός by metathesis (Pokorny 1959/1994: 267).

All in all, while some uncertainty persists, the reconstruction *dʰeuub- cannot be easily dismissed. Since the root is attested in Tocharian as well as in other subgroups, it can be reconstructed back to Proto-Indo-Tocharian.

The root *gʰrebʰ₁/h₂- ‘to seize’ is possible reflexes in both Anatolian and non-Anatolian subgroups and thus aspires to actual Proto-Indo-European status. The reconstruction is not unproblematic, however. First, the evidence for *b as opposed to *bʰ is restricted to the long vowel of Balto-Slavic, which is most naturally explained by Winter’s Law, requiring a voiced unaspirated stop (Olsen 1984: 141–142, 144 n. 9; Rasmussen 1987/1999: 219). The p of OIr. grápa would point to PIE *b as well, but it is disputable if the word belongs here, and even if it does the Germanic evidence
is ambiguous; compare forms like modern English *grab, pointing to PIE *bh (see e.g. Kluge & Seebold 1883/1999: 335; Kortlandt 1988/2009: 69). Not all long vowels of Balto-Slavic are the result of Winter’s Law, however, and we even find evidence of a short vowel in the apparent Latin cognate grebt ‘to scrape; to seize’, which would point to *bh. This form may, on the other hand, also reflect the root *ghrebh- ‘to dig’ (Olsen 1984: 142; LIV² 201–202; Derksen 2015: 186).

The second, and more serious, problem with this root in our context is the possible Anatolian cognate – or rather cognates, as there are two competing hypothetical reflexes. Since the root has no reflexes in Tocharian or Italo-Celtic, the absence of an Anatolian reflex would mean that the root would be demoted to the Indo-Germanic stage.

In LIV² 201 *ghrebh₁/h₂- is thought to have yielded Hittite kar(ap)p- ‘to lift’. Kloekhorst (2008: 442–444, 452–454), by contrast, objecting to the phonetic developments of this scenario, follows Sturtevant and Hahn (1933/1953: 31) in connecting the root with Hittite karap- ‘to devour’, which is semantically less attractive but fits better formally.

It seems that there is no straightforward solution to this problem, but – although I acknowledge that this is not an entirely satisfying conclusion – we may consider it relatively likely that one of the Hittite candidates is a true reflex PIE *ghrebh₁/h₂-. As it is also slightly more economical to reconstruct *b than *bh in view of the Balto-Slavic (and possibly Germanic) material discussed above, the most straightforward reconstruction, at least within a theory that is not a priori disfavouring PIE *b, seems to be *ghrebh₁/h₂-. With the proviso that the evidence is far from conclusive, we can assume that this root provides an example of PIE *b.

As Birgit Anette Olsen notes (pers. comm., November 2019), it is striking that this root, if reconstructed correctly, has a somewhat peculiar structure, with the *b followed immediately by a laryngeal. For those trying to avoid PIE *b it would be possible to speculate that the root originally ended in *bh, and that the *b required by Balto-Slavic might be the result of a (possibly phonologically regular) reinterpretation of PIE *bh as *bh in pre-Proto-Balto-Slavic, which would have triggered Winter’s Law.

5.7 *leb- ‘to lick; lip’: Proto-Indo-European

- ?Anatolian: ?Hi. lip(p)- ‘to lick’
- Italic: La. labium, labrum neut. ‘lip’
- Germanic: OSw. leipi masc. ‘lip’; ODa. leapa masc. ‘lip’; OHG lefs masc. ‘lip’; OEng. lippa masc. ‘lip’
- ?Greek: ἱπποσ masc. ‘lobe’

Another candidate for PIE *b is provided by the root *leb- ‘to lick; lip’. Although reflexes of the root are only found in a small number of subgroups, the phylogenetic relationship among these subgroups is such that it suggests Proto-Indo-European status of the root. The reconstruction of *b rests only on the Germanic forms and, if related at all, Greek ἱπποσ, but it is not contradicted by the Italic and Anatolian material.

It is questionable if the Hittite verb for ‘to lick’ is an actual cognate of the Italic and Germanic nouns meaning ‘lip’. Kloekhorst (2008: 528) regards the Hittite verb as “onomatopoetic in origin” and does not list the Italic and Germanic forms, thus implying that there is no etymological connection between Hi. lip(p)- and the remaining words. Similarly, no verbal root *leb- is included in LIV². According to Sihler (1995: 146) sets of words of the ‘lip’ type “are probably not in fact PIE, but come from such sources as slang and regional loan-words from non-IE sources”.

To *b or not to *b: Proto-Indo-European *b in a phylogenetic perspective

13
Similarly, de Vaan (2008: 319) thinks "*lab- 'lip' may be a borrowing from an unknown adstrate". The semantic difference, 'to lick' in Anatolian vs. 'lip' in Italic and Germanic, is also worth noting: licking is usually connected with the tongue, not the lips. Interestingly, as Martin Joachim Kümmel points out to me, a common West Iranian form *lapi-(čV-) with the meaning 'lip' may be reconstructed on the basis of Modern Pers. lab and Kurdish lev. Since this form, with its *p, is hardly related to the Italian and Germanic words for 'lip', it might suggest that words for 'lip' containing a lateral and a labial may arise independently.


It is difficult to determine whether the Hittite verb belongs here and thus to which stage the root, with its *b, is reconstructible. If we accept the connection between the Hittite, Italic and Germanic words, the root belongs to Proto-Indo-European; if we exclude the Hittite verb, the root can only be reconstructed back to Proto-Indo-Celtic.

6 Conclusions

Our analysis has shown that there are no candidates for initial *b in Proto-Indo-European. The only roots with initial *b that may claim some antiquity, *bel- 'strong' and the less appealing *bak- 'stick', do not have cognates in Anatolian and Tocharian and can thus only be reconstructed back to Proto-Indo-European. *bak- probably cannot even be traced that far back.

For non-initial *b the evidence is better, as has been noted by several scholars. Two roots, *gʰrebh₁/h₂- 'to seize' and *leb- 'to lick; lip', have possible (but not certain) reflexes in Anatolian and non-Anatolian subgroups and may thus be reconstructed back to Proto-Indo-European. The root *dʰeuub- 'deep' is not found in Anatolian, but since it has cognates in Tocharian and other subgroups it can be relatively safely reconstructed back to Proto-Indo-Tocharian; however, the reconstruction of *b, as opposed to *bʰ, is questionable in this root. The present stem *pɪbe/o- 'drink' as well as two further roots, *h₂ábel- 'apple' and *treb- 'to build', have a distribution that makes it likely that they were present in Proto-Indo-Celtic.

The roots containing *b for which some antiquity may be claimed, including those listed in § 5.1 but not discussed in detail, are shown in fig. 2.
To *b or not to *b: Proto-Indo-European *b in a phylogenetic perspective

All in all, we can be fairly certain that at the Proto-Indo-Celtic stage at least some roots contained *b – in the case of *bel- even in initial position – whereas in earlier stages there is no evidence whatsoever for initial *b and only disputable reconstructions containing non-initial *b.

It is possible to argue, on this basis, that *b was an innovation of Proto-Indo-Celtic after the separation of Anatolian and Tocharian. On the other hand, since it cannot be excluded that there were Proto-Indo-European roots containing *b that have simply been lost or are not attested in these subgroups, this idea cannot be more than speculation.

References


Fig. 2. Phylogenetic tree with roots containing PIE *b.


To *b or not to *b: Proto-Indo-European *b in a phylogenetic perspective

LIV\(^1\) = Rix et al. 1998.
LIV\(^2\) addenda = Kümmel 2015.

NIL = Wodtko, Irslinger & Schneider 2008.


To *b* or not to *b*: Proto-Indo-European *b* in a phylogenetic perspective


