Papers in Historical Phonology
http://journals.ed.ac.uk/pihph
ISSN 2399-6714
Volume 7, 126–143 (2022)
DOI: 10.2218/pihph.7.2022.8959

On the PIE root-structure constraint prohibiting repeated consonants

PAUL S. COHEN
Unaffiliated

BENJI WALD
Unaffiliated

Abstract

This paper confronts and resolves the problem of apparent exceptions to the constraint prohibiting the co-occurrence of identical consonants in both syllable margins of the PIE root: schematically, †… C_i … E … C_i …, where † indicates the prohibition of the root structure following it, C_i = the identical consonant, E = the ablauting vowel, and … = optional additional consonants in the syllable margins. In advancement of previous work addressing this problem — most recently exemplified in Cooper (2009), Corbeau (2013) and Weiss (2020) — it eliminates several potential exceptions to the constraint and proposes that, once a cross-linguistic absence-of-contrast principle is taken into account which determines the relation of laryngeal features (glottalization, aspiration, and voicing) to the syllable margins that contain them, no clear-cut exceptions remain.

1 Introduction

Current scholarship leaves unresolved the extent to which the PIE root-structure constraint prohibiting repeated consonants is exceptionless or merely gradient. In the present paper, we argue that the adducible evidence points strongly to exceptionlessness. The paper has the following general structure: Section 2 gives a brief summary of the history in PIE[1] linguistics of the positing of root-structure constraints and a detailed look at previous work — primarily that of Cooper (2009), Weiss (2020), and, especially,

[1] We use the term *PIE* herein to refer to what some would call *Proto-Indo-Anatolian* (*PIA*) — i.e., the parent-language of Anatolian and what is sometimes termed *Common Indo-European* (including Tocharian). This latter grouping is what is called *North IE* by Ringe (2006, 5). He calls the next oldest grouping — i.e., what remains after Tocharian splits off — *West IE*.

Corbeau (2013) — concerning putative PIE roots containing repeated consonants. Section 3 sets out some methodological principles and then sections 4 to 9 give (i) a reexamination of items with repeated consonants that Corbeau could not exclude (along with an examination of an item he omitted) that we believe we can explicate and eliminate via standard methods, and (ii) an in-depth analysis of the four remaining roots (two verbal, one adjectival/ verbal/nominal, one nominal), leading us to propose an explanation for their being licit (i.e., not actually exceptions) that follows directly from the findings of Kehrein & Golston (2004). Section 10 concludes, with some speculative remarks.

2 The positing of PIE Root-Structure Constraints

The idea of specific PIE root-structure constraints is usually attributed to Benveniste (1935), who propounded that the basic PIE root structure was, in modern notation, $C_i e C_j$.[2] In more recent times, the validity of roots with complex onsets and codas has been fully accepted,[3] and the primary focus has been on the absence (or, at least, very low frequency) of various combinations of consonants, whether identical or merely having certain phonetic features.

Thus, for example, Ringe (2006, 7 with fn. 3) states 'a root could not contain oral stops at the same place of articulation both in its onset and in its coda.'[4] This is footnoted as follows:

> Apparently this constraint classed *m with the bilabial oral stops; that is, there were no roots like '*pem-' and '*mebh-', including both a bilabial oral stop and *m. However, *n was not classed with the coronal oral stops, since we must reconstruct *nadh- 'to tie', *newd- 'to push', *ten- 'to stretch', etc. Three clear exceptions to the constraint, *tewd- 'to beat', *tend- 'to cut', and *mems- 'meat',[5] are securely reconstructable; it is of course not surprising that they involve coronal stops and *m. Other apparent exceptions, such as *bhrem- 'to make a noise', either appear to be onomatopoeic or are not securely reconstructable for PIE proper, so far as I am aware.

---

[2] He writes (1935, 170): 'ı^0 La racine indo-européenne est monosyllabique, trilitère, composée de la voyelle fondamentale *ĕ* entre deux consonnes différentes. 2^0 Dans ce schème constant : consonne + *e* + consonne, les consonnes peuvent être de n'importe quel ordre pourvu qu'elles soient différentes ; seule est exclue la coexistence d'une sourde et d'une sonore aspirée.'

[3] Thus, e.g., Weiss (2020, 49) states, 'The root of a nominal or verbal form has the minimal shape CVC.' Nevertheless, he begins the footnote to this statement 'But the single most common root shape is CVCC.'

[4] Referring directly to this statement, Cooper (2009, 56 fn. 5) writes: 'To be explicit, perhaps a better characterization for Ringe's observation might be "a root could not contain [oral] stops at the same place of articulation *in any prevocalic or postvocalic position.*"'

[5] We account for this item — which, in agreement with, *i.a.*, Vine (1998, 92ff.), Watkins (2000, 54), and Cohen (2015), we reconstruct as **mēms-* — below.

He goes on to write (2006, 8):

> Most surprising of all was a series of constraints on the shapes of root-syllables. A root could not contain two voiced stops,[6] nor could it contain both a voiceless stop and a voiced aspirate unless the former occurred in a root-initial cluster with *s.

Cooper (2009, 56) gives the following summary, gleaned from the relevant literature, of 'PIE Root Consonantal Co-Occurrence Restrictions':

a.	* *C_iVC_i	in which the consonants are identical
b.	* *mVP/PVm	in which *m* is the labial nasal and *P* is any labial oral stop
c.	* *DVD	in which *D* represents a voiced unaspirated stop
d.	* *TVDh/DhVT	in which *T* represents a voiceless stop and *Dh* represents a voiced aspirated stop
e.	* *CVRR	in which *R* represents any sonorant consonant

(where, as he explains in fn. 3, '…one asterisk marks ungrammaticality, the other a reconstructed form'). Then, comparing expected vs. actual frequency counts, he concentrates on the infrequency or near absence[7] of roots

> … not only with homorganic consonantal co-occurrence, but with co-occurrence of consonants sharing manner features as well, as this broader scope is better aligned with the nature of the root structure constraints which have been posited for PIE. (2009, 56)

He is careful to list apparent exceptions to the generalizations he has formulated.

Corbeau (2013) has sections on root-structure constraints against the repetition of a consonant, co-occurrence of two voiced stops, and co-occurrence of a voiceless stop and a voiced aspirated stop (in addition to a section on whether PIE roots could have an initial *r*). Unlike Cooper, he does not use frequency counts; rather, he performs in-depth examination of potential counterexamples and finds that there are a few of these that he cannot eliminate.

Weiss (2020, 49f.) lists constraints against, *i.a.*, roots that begin and end with a plain voiced stop, or begin and end with a liquid, or contain a voiceless stop and a voiced aspirate and no resonant; and writes (2020, 49), that with the exception of **ses-* 'sleep',[8] 'A root could not begin and end with the same consonant: §C_i(R)E(R)C_i' (where § indicates a

---

[6] Cooper (2009, 56 fn. 6) reminds us '… that it in fact could, of course, as long as at least one stop was an aspirate.'

[7] Thus allowing for a position that these constraints might be merely gradient.

[8] We will discuss this item in the next section.

prohibited/nonexistent root; C = consonant; R = resonant; E = ablauting vowel). In a series of email interchanges with the present authors, however, Weiss suggests that this constraint may in fact have been gradient (as exemplified in the approach of Cooper [2009]) rather than absolute, since there are at least a few other apparent exceptions. As stated above, the present paper argues that the constraint is exceptionless. In addition, like Corbeau (2013), we will advocate that the constraint should be framed more generally, such that it eliminates all other types of repetition of a consonant as well. In its most general formulation, it can be stated as †… C_i … E … C_i … (where † indicates the prohibition of the root structure following it, C_i = the identical consonant, E = the ablauting vowel, and … = optional additional consonants in the syllable margins); we will refer to this formulation as the Repeated-Consonant Constraint, henceforth abbreviated as RCC. Establishing our position requires that we eliminate the apparent exceptions that Corbeau either was unable to handle or omitted. This is, in fact, the first major focus of our paper.

3 The corpus

Cooper (2009) and Corbeau (2013) use the verbal roots in *LIV* as their entire corpus, and we too use *LIV* as our primary source. In addition, however, we know of a relevant root that shows up in adjectival, nominal, and verbal stems; and there are also two relevant nominal items. Moreover, one of the key *LIV* entries (Item 5) turns out to have developed from a fundamentally adjectival root. We will handle all of these as well.

As a preamble to our analysis of the apparent exceptions to RCC, we offer an explicit categorization of *LIV* roots we have ruled out and ruled in:

- We have disqualified all items that have only cover-symbols (rather than specific phonemes) as their potentially repeated consonants, since these items cannot *ipso facto* be counterexamples.
- Like Cooper (in his methodology) and Corbeau (in his analysis, generally speaking), we have eliminated all roots annotated with a question mark, which *LIV* (43) defines as indicating that '… der uridg. Status der betreffenden Wurzel ist fraglich.'
- We have followed *LIV* in allowing roots that may be (or almost assuredly are) formed from other roots via the addition of one or more consonants at the end (i.e., 'root enlargements/extensions').[9]
- We have included roots beginning with *s*-mobile that have another *s* in them.

[9] *LIV* (6f.): '… bleibt der Begriff der Wurzelerweiterung rein synchron-deskriptiv und damit ohne Bedeutung für eine historische Analyse. Im vorliegenden Werk sind so die scheinbar oder anscheidend erweiterten Wurzeln als eigene Lemmata registriert; wo es angebracht schien, ist in einer Anmerkung auf die Möglichkeit einer Erweiterung einer einfacheren Wurzel hingewiesen.'

4 The apparent exceptions

An exhaustive search of *LIV* yields the following alphabetic list of apparent exceptions to RCC:

(1) **h₁eh₁s*- 'sitzen' (p. 232)
(2) 1. **h₁reh₁*- 'fragen' (p. 251)
(3) 2. **h₁reh₁*- 'rudern' (pp. 251f.)
(4) **h₂elh₂*- 'ziellos gehen' (p. 264)
(5) **h₂seû̯s*- 'trocken werden' (p. 285)
(6) **prep*- 'in die Augen fallen, erscheinen' (p. 492)
(7) **ses*- 'ruhen, schlafen' (pp. 536f.)
(8) **(s)gᵘ̯esh₂*- 'erlöschen' (pp. 541ff.)
(9) **skek*- 'sich schnell bewegen, springen' (pp. 551f.)
(10) **(s)kers*- 'kratzen, (Wolle) krempeln' (p. 559)
(11) **tetk̂*- 'erzeugen, herstellen' (pp. 638f.)

The adjectival/verbal/nominal root alluded to above is:

(12) **h₁erh₁*- 'quiet, at rest' (*EIEC* [474 *s.v.* QUIET]) (with reflexes in Baltic, Celtic, Germanic, Greek, and Indo-Iranian)

The two primarily nominal forms are:

(13) **h₁î̯eh₁*- 'year; (period of) time; new season'[10] (with reflexes in Anatolian, Germanic, Greek, Slavic, and, possibly, Italic and Tocharian)
(14) **mēms*- 'meat, flesh' (with reflexes in Albanian, Armenian, Balto-Slavic, Celtic, Germanic, Greek, Indo-Iranian, Italic, and Tocharian)

5 Analysis of the apparent exceptions

Item 1 **h₁eh₁s*-, with reflexes in Anatolian, Greek, and Indo-Iranian

LIV footnotes this as follows: 'Ungewöhnliche Wurzelstruktur vielleicht liegt doch ursprunglich **h₁es*- (= 1. **h₁es*- 'sein' ?) vor …, mit Reduplikation dann **h₁e-h₁s*-.'

[10] For etymological details of forms derived from **h₁î̯eh₁*-, see, especially, Pinault (2018, 49), Stifter (2017, 224). This root is itself generally held to be an extended form of **h₁eî̯*- 'to go'; see, e.g., Pinault (2018, 49), Stifter (2017, 224, with reservations), and *EIEC* (654 *s.v.* YEAR), though *EIEC* gives the extended form with optional **h₁* in initial position and **ē*, rather than **eh₁*, medially. As Michael Weiss points out (pers. comm.), however, **ē* is ruled out by phonological developments attested in the Luvian reflexes. Thus Melchert (1994, 241, 245 resp.) writes for Luvian: '/i:/ < */e:/' (with exx.) and '/a:/ < */ae:/ (**eh₁*): … *āra/i-=á+ra/i-* 'time' < PA[nat.] **ǽro*-.'

EIEC (522 *s.v.* SIT), offers another derivational possibility for the relevant PIE form, namely *$h_1ēs$- 'sit', writing:

> Though only modestly attested, its geographical distribution guarantees its PIE antiquity. Originally a lengthened-grade intensive of *h_1es- 'be'. *h_1es- may have originally meant 'sit' and on 'weakening' to 'be' was replaced in its original semantic sphere by the derived intensive *$h_1ēs$-.

Corbeau (2013, 13) finds *LIV*'s explanation as a straightforward reduplication to be 'plausible' and Cooper (2009, 58 fn. 14) judges the item to be '… likely [a product] of reduplication'. *EIEC*'s explanation, on the other hand, seems unnecessarily complicated to us. But either alternative would eliminate this item as a clear-cut counterexample: With *LIV*'s, it would not be unequivocally monomorphemic; with *EIEC*'s, it would not have a repeated consonant.[11]

Item 2 *h_1reh_1- 'fragen', with reflexes in Anatolian and Greek

This item will be discussed in a separate section below.

Item 3 *h_1reh_1 'rudern'

LIV lists verbal reflexes in Baltic, Celtic, Germanic, and Greek. In addition, the entry is footnoted (p. 251) as follows: 'Vgl. lat. *rēmus* m. 'Ruder'; daneben schwebeablautend *h_1erh_1- in ved. *aritár-* m. 'Ruderer', *arítra-* 'Ruder'.' The Greek form is given as '?gr myk. Inf. *e-re-e /ereen/* 'rudern',' which is footnoted with 'Könnte auch thema-tisches **ere-* < **$h_1érh_1$-e-* sein.'

EIEC (490 *s.v.* ROW) gives the entry as *h_1erh_1- 'row' and lists Baltic, Celtic, and Germanic verbal reflexes, continuing with 'Cf. the derivative *$h_1erh_1tér$- 'rower': Grk ἐρέτης 'rower', OInd *aritár-* 'rower'.' The entry concludes with, 'Geographical distribution makes this word a sure candidate for PIE status.'

This item will be discussed in a separate section below.

Item 4 *h_2elh_2-

LIV's entry gives only Greek and Italic reflexes, namely Gk. *ἀλάομαι* 'irre umher', Umbr. *amb-oltu* 'soll herumgehen', and Lat. *amb-ulō*, *-are* 'umher-gehen'. However, as pointed out to us by Michael Weiss (pers. comm.), there is no firm evidence for a final *h_2 in this root: *ἀλάομαι* could well be denominal from *ἄλη* 'wandering', and the Italic forms are entirely uninformative. We note furthermore that *EIEC* (629 s.v. WANDER) lists the root as having no final laryngeal: *h_ael- 'wander'. In any case, it cannot be

[11] And see section 8 below for yet another possible reason, which is fundamentally phonetics-based, that this item may not in fact be an exception to the RCC.

shown to have a repeated consonant and is thus not a valid counter-example.

Item 5 **h₂seu̯s-*

LIV gives reflexes in Balto-Slavic, Greek,[12] and Indo-Iranian; thus the validity of the item appears strong, dating back at least to Ringe's 'West IE'. However, *LIV*'s entry is footnoted with a reference to Lubotsky (1985) '[z]um Wurzelansatz'.

Lubotsky (1985)'s detailed analysis is convincing. The final point in his paper (p. 9) reads:

> **H₂sus-* is not a verbal root, but an adjective 'dry'. The verbs 'to be, become dry' were derived from it in the separate languages. The adjective **H₂sus-* should be considered an original perfect participle of the root **H₂es-*, which probably was a *perfectum tantum*.

Apparently following Lubotsky, *EIEC* (170 *s.v.* DRY) gives two relevant PIE forms: **h₂es-* 'be/become dry' and **h₂sus-* ~ **h₂sousos* 'dry'. Obviously, the second of these and *LIV*'s **h₂seu̯s-* correspond. In the subentry for **h₂sus-* ~ **h₂sousos* 'dry', *EIEC* writes:

> The underlying meaning seems to be adjectival. Perhaps **h₂sus-* is from the perfect participle of [**h₂es-* 'be/become dry']. Both words are widespread and assignable to PIE.[13]

Thus the form may well have been recognized by PIE speakers (and even West IE speakers) as not monomorphemic, and so cannot function as a counterexample.

Item 6 **prep-*

This looks to be a legitimate root. However, it is unclear whether it dates back to PIE or merely to a later dialectal stage (presumably Ringe's 'West IE'). *LIV* gives Greek and Armenian verbal reflexes, with this footnote:

> Nur arm., gr. und wohl in air. Richt 'Form, Gestalt' < **pr̥p-tu-*, ganz unsicher Zugehörigheit von ahd. *furben* 'reinigen, putzen, fegen' (aus 'in die Augen fallend machen, ein Ansehen geben'?)….

---

[12] With some reservations concerning the existence of an initial laryngeal. However, as Corbeau (2013, 15) points out, irrespective of the status of the proposed laryngeal, the form would still have a repeated consonant.

[13] The *LIV* root (p. 257) that corresponds to *EIEC*'s **h₂es-* 'be/become dry' is **h₂eh₁s-* '(durch Hitze) vertrocknen'. *LIV* does not cross-reference **h₂eh₁s-* with **h₂seu̯s-*.

EIEC (25 *s.v.* APPEAR) lists **prep*- 'appear' as having Celtic, Greek, and Armenian reflexes, and adds, 'Perhaps OHG *furben* belongs here as well.' The subentry for the root is then characterized as follows: 'A word of the west and center of the IE world.'[14]

Corbeau (2013, 17) writes:

> Attestations only from Greek and Armenian, as LIV provides, would be too limited to posit PIE origin. Old Irish makes that more plausible; as would Old High German, but note that this form is semantically far off. In any case, without any Eastern forms, care must be taken.

In Corbeau's Table 2.1 (2013, 20), which summarizes his assessment of all the roots he categorizes as (potentially) having the same consonant in both onset and coda, he finds the PIE status of **prep*- to be doubtful, based on its geographical distribution (a position that is consistent with *EIEC*'s view). We agree.

Item 7 **ses*-, with reflexes in Anatolian and Indo-Iranian (and, possibly, Armenian and Celtic)

Weiss (2020, 49), e.g., writes: 'This root is perhaps onomatopoetic or in some other way iconic. Cf. PDE [i.e., Present Day English] *shush!*', and footnotes this as follows:

> Alternatively, **ses*- may represent an affective reduplication of the initial consonant of **su̯ep*- 'sleep'. Cf. Fr. *faire dodo* 'go to sleep', where *dodo* is from *dormir* 'sleep'.

Similarly, *EIEC* (526 *s.v.* SLEEP) states that **ses*- 'rest, sleep, keep quiet'

> … is almost surely onomatopoeic in origin — derived from the sound of gentle snoring (cf. NE *counting z's*) or from an interjection similar to NE *sh!*

Cohen (2017, 120f.) offers as an alternative a morphophonological derivation via a telic *s*-extension of PIE **sed*- 'sich setzen' (*LIV* [513]).

**ses*- is surely a PIE form, but it is either onomatopoetic or not (originally) monomorphemic; it is therefore not a valid counterexample.

---

[14] *EIEC* (xiv), regarding dialectal provenance, states: '[W]e use a number of impressionistic labels to localize words which may never have extended throughout the Proto-Indo-European speech community.'

Item 8 **(s)g^{u}esh$_2$-*[15]

This item may not be an *s*-mobile form. The only reflexes in *LIV* showing an underlying initial *s* are Greek: e.g., aor. *ἔσβη* 'erlosch' and pres. (Hesych.) *ζείνομεν· σβέννυμεν*, (Att.-Ion.) *σβέννῡμι* 'lösche (aus)'.

EIEC (188 *s.v.* EXTINGUISH) lists the PIE item as **g^{w}es-*, and along with Anatolian (Hitt. *kist-*), Balto-Slavic, Indic, and Tocharian reflexes, gives *σβέννῡμι* 'extinguish' for the Greek reflex. The subentry continues, tellingly, however:

> Semantically this set fits very well but, unfortunately, the Greek argues for a labio-velar **g^{w}-* while Hittite supposes a simple velar **g-* (the rest of the set are ambiguous between these forms). The **prefix *σ-*** [emphasis, ours] in Greek (< **sg^{w}es-nu-*) is also unexplained.[16] Aside from these difficulties, the root can be reconstructed with a moderate degree of confidence.

It is noteworthy that *EIEC* treats this item quite differently than it does an obvious *s*-mobile etymon such as *(s)meld-* 'to melt' (378 *s.v.* MELT), which straightforwardly exhibits the optional *s*.

So we see that the initial *s* of the item given by *LIV* need not — and probably should not — be accepted as beginning an *s*-mobile root. Obviously, without this *s*, the root has no repeated consonants, and thus can be eliminated as a counterexample

Item 9 **skek-*

LIV gives only Germanic, Slavic, and Celtic reflexes. On that basis, Corbeau (2013, 18) states, 'Only European attestations, as LIV provides for **skek-*, are not enough to be certain that this root was PIE'.

Similarly, Cooper (2009, 58 fn. 14) writes, '[A]s for **skek-*, its contiguous distribution only in Germanic, Slavic and Celtic may make its PIE status questionable.'

Like **prep-*, this appears to be a legitimate root, but one whose PIE status is dubious — perhaps dating back only to 'West IE'.

---

[15] Omitted by Corbeau (2013).

[16] *LIV* recognizes relevant anomalies as well; the PIE item itself is annotated in part (p. 542 fn.1): 'Ein Problem ist der vom Gr. vorausgesetzte Labiovelar: die toch. Evidenz (otoch. nicht †*kus-*) scheint damit kaum vereinbar, … desgleichen die attraktive Verbindung mit heth. *kāst-/kist* c. 'Hunger' ….' In the same vein, Melchert (1994 , 120) writes, 'The problem of */g/ vs */g^{w}/ in the PIE root 'be extinguished' is quite real but is not confined to Hittite …. TochA *käs-* … also requires a plain velar */g/ like Anatolian *kišt* ….'; (since this item is also attested in Palaic, its etymon goes back to at least Proto-Anatolian; see, e.g., Melchert [1994, 210]). And Kloekhorst (2008, 461) in his entry for *kāšt- / kišt* gives only Tocharian cognates and reconstructs the PIE etymon with no initial *s* and with a palatal or velar (i.e., not labialized) *g^{h}*.

Item 10 **(s)kers-*, with reflexes in Baltic (Lithuanian), Germanic (Old High German), and Italic (Latin)

The evidence for an *s*-mobile in this item is, at best, sketchy. The only *s*-initial reflex in the *LIV* entry is '?[ahd. (+) *skerran* 'scharren, schaben''. The question mark, when occurring in this position, is defined (*LIV* [44]) as follows: 'Vor einer einzelsprachlichen Bildung: Fraglich ist, ob diese sprachliche Bildung tatsächlich die uridg. Stammbildung fortsetzt.' Moreover, the gloss given for the putative OHG reflex is a questionable semantic match for the one given for the other two reflexes — the Lat. and Lith. forms are both glossed as 'Wolle krempeln'. Thus, we believe the OHG form has no probative value here. Eliminating it leaves us with a root that does not evince a repeated consonant and therefore cannot serve as a counterexample.

Item 11 **tetk̂-*, with reflexes in Balto-Slavic and Indo-Iranian

This has long been thought to be a reduplication. Indeed, *LIV* footnotes the entry as follows: '… abstrahiert aus einem uridg. zum Wurzelaor. umgedeuteten Primärstamm **tétk̂-*, der auf einen voruridg. von **tek̂-* 'zeugen, gebären' gebildeten redupl. Aor. ***te-tk̂-* zurückgeht …'. Like Items 1 and 5 (and possibly Item 7), the form may still have been recognized at the relevant stage of PIE as not monomorphemic, and can therefore be eliminated as a counterexample.

Item 12 **h₁erh₁-*

EIEC (474 *s.v.* QUIET) gives two stems:[17] **h₁erh₁-m-* 'to rest, support' (with reflexes in Baltic, Celtic, Germanic, Greek, and Indo-Iranian); and **h₁erh₁-u̯eh₁-* 'quiet, calm' (with reflexes in Baltic, Germanic, and Greek). This item will be discussed in a separate section below.

Item 13 **h₁i̯eh₁-*

This item will be discussed in a separate section below.

Item 14 **mēms-*

For this item, we advocate the etymology expounded in Cohen (2015) as a nominal reduplication based on an *s*-extended form of the root **meh₁-* '(ab)messen' (*LIV* [424]). We note that, even if his etymology is not accepted, in a discussion particularly relevant for our present concerns, he points out (2015, 58) not only the dubious identical consonant on

[17] Joe Salmons *apud EIEC* (p. 474) writes, in explanation: 'While the precise form and details are not yet clear, some form of this root (with two different suffixes) can be posited for PIE.'

both syllable margins of this questionable root, but also the consistent lengthened-grade vowel and the impossibility of the long vowel to have been generated by Szemerényi's Law. **mēms*- can thus be eliminated as a viable counterexample.

In sum, all but four of the potential counterexamples to the RCC can be eliminated by one or both of the following considerations:

a. The evidence does not support reconstructing the root for PIE per se, but, rather, only for a later node.
b. The evidence suggests that the item does not reflect a phonological constraint on a PIE root, but, rather, that the root is embedded in a more complex form resulting from either grammatical (or, possibly, onomato-poetic) reduplication or some other grammatical process.

6 The remaining counterexamples and statistical analysis

Still unaccounted for are the four counterexamples: Items 2, 3, 12, and 13 — respectively **h₁reh₁*- 'fragen', **h₁reh₁* 'rudern', **h₁erh₁*-, and **h₁i̯eh₁*-. Remarkably, these items are all of the form (using Weiss's notation) *h₁*CE*h₁* / *h₁*EC*h₁*, which, we argue, is not coincidental.[18] The large number of unquestioned PIE verb roots contained in *LIV* allows us to test whether that position is statistically supported.

Even though there are only two PIE verbs in *LIV* that have an unquestionably repeated consonant, in both cases that consonant is **h₁*. We ask whether this number makes **h₁* significantly different from all the other consonants, since the latter invariably fail to co-occur simultaneously in both syllable margins (i.e., both the onset and the coda). The appropriate tool for deciding is Fisher's Exact Test (FET), a standard statistical test where one of the factors (repetition of a consonant, in the present case) has a low number of occurrences. FET tests the null hypothesis that the two sets of consonants being compared are indistinguishable in their behavior. By convention, the null hypothesis is refuted by a probability of at most 1 in 20 chances that the difference between the compared sets is simply an accident of the sample. When the null hypothesis is refuted, the conclusion is that the compared sets are indeed different with respect to the behavior being tested — in our case, with respect to the possibility of the same consonant simultaneously occurring in both margins of a lexical root.

[18] Obviously, in all four of these items the 'C' is in fact a resonant, and we will offer some thoughts in that regard in section 9 below. In any case, we will give an explanation in section 8 below for the allowability of the combination of any (non-guttural) C with **h₁* in PIE roots.

In total, *LIV* lists 675 unproblematic PIE verb roots. Of these, 589 contain no instance of *h_1. Thus 86 contain at least one instance of *h_1 in one or both of the syllable margins. Since there are more than six times as many roots lacking *h_1 as containing it, if *h_1 were no different from the rest of the PIE consonants, we would expect at least six times as many roots lacking *h_1 to contain a repeated consonant as we observe for roots containing *h_1. In other words, we would expect between ten and twelve non-*h_1 roots to contain repeats compared to the two *h_1 roots that contain repeats. But we observe no non-*h_1 repeats among the unproblematic roots. According to FET, the probability that this is an accident of the available sample is p=0.0161, well within the conventionally accepted range for a statistically significant result in, *i.a.*, linguistic studies, $p \leq 0.05$.

While this result supports the position that *h_1 is indeed different from all the other PIE consonants with respect to the RCC, it leaves a number of more specific problems unresolved. In particular, the preceding FET offers no clue as to which and/or how many of the non-*h_1 consonants should be expected to repeat if they behaved similarly to *h_1. It is not expected that each of the consonants would repeat, because there are more than twelve PIE consonants to begin with, where only twelve consonants with a single repetition would satisfy expectations. Similarly, it is unclear just how many consonants would be expected to repeat — only one, or more? Only two examples of repetitions of a consonant other than *h_1 would deprive *h_1 of its significance according to the FET, whether by a single repetition for two consonants other than *h_1 or by two repetitions of a single consonant other than *h_1 — but the test says nothing about which consonant(s) to expect in such a case (in either case the FET distinguishing *h_1 from the sum of all other consonants would be 0.0811, a conventionally insignificant probability). We underscore the fact that there are limitations on the amount of data available for testing each non-*h_1 consonant individually against *h_1 for statistical significance. This is particularly obvious for *b, the lowest frequency consonant (n=4 according to Cooper [2009, 63 Table 5]), where a repetition in 2% of the examples, corresponding to 2 of the 86 examples of *h_1, would be a small fraction of a single example (< 0.1), and therefore completely unexpected.[19] Later discussion will give evidence indicating that *h_1 by itself is not the exception to the RCC, but rather the repetition of *h_1 in opposite syllable margins.

---

[19] On a similar basis, Cooper (2009, 63f.) questions the categorical status of the PIE **DVD constraint, i.e., the constraint against voiced unaspirated stop co-occurrence in both syllable margins of a lexical root. He points out, in particular, that the extreme paucity of roots with *b (n=4) depresses the frequency of roots with unaspirated voiced consonants (n=150) below the level of voiced aspirated stops (n=167) in contrast to the numerical superiority of all other voiced unaspirated stops over their aspirated counterparts. The low frequency of *b is particularly noteworthy compared to its aspirated counterpart *b^h (n=53).

7 Phonological systems with phonetic faucals and laryngeals

We will argue below that the permissibility of roots having *h_1 in both margins of PIE roots is a straightforward consequence of *h_1's being, phonetically, a laryngeal. In that regard, it will be helpful to begin with a short cross-linguistic discussion of phonological systems having both types of guttural consonants, namely faucals and laryngeals — of which PIE is an example.[20]

As delineated by Rose (1996, 2000), who analyzes gutturals in a wide variety of currently-spoken languages — including members of the Semitic, Cushitic, Salish, and Caucasian families — there are, generally speaking, two major subtypes of languages having both laryngeals and faucals: those in which they function as a single phonological grouping (e.g., Syrian Arabic) and those in which they function as separate phonological groupings (e.g., Interior Salish).[21]

PIE, as posited by the great majority of scholars, had three guttural consonants, the so-called laryngeals: *h_1, *h_2, and *h_3. Of these, only *h_1 was, phonetically, a laryngeal.[22] The other two, as we have stated, are usually reconstructed as faucals: *h_2 as a voiceless fricative ([χ] or [ħ], or the like); *h_3 as a voiced[23] labialized fricative ([ʁʷ] or [ʕʷ], or the like).[24] PIE, it turns out, aligns most closely with the second subtype (i.e., with Interior Salish). We know this because of the phonetic/phonological

---

[20] Using *faucal* to denote postvelar (uvular, epiglottal, or pharyngeal), but not laryngeal. Faucals are also referred to as *supralaryngeals*.

[21] This distinction is probably best viewed as the (near-)endpoints of a continuum. Thus, for example, Tigrinya and Tigre, two closely-related North-Ethiopian Semitic languages, have some morphophonological processes that treat all gutturals alike and others that distinguish laryngeals and faucals, but differ as to specifics in this regard (see Rose [2000, 88–98] for details).

[22] *h_1 is most commonly reconstructed as [h], though a significant minority of scholars believe it to have been [ʔ] or are undecided between the two phones.

[23] *h_3 is generally presumed to have been voiced because, unlike *h_1 and h_2, it voiced preceding voiceless stops; thus PIE *$pi\text{-}ph_3\text{-}e\text{-}ti$ 'drinks' > Ved. *pibati* 'drinks' (see, e.g., *LIV* [462f. fn. 4 *s.v.* *$peh_3(i)$- 'trinken']).

[24] One other position should be mentioned here: Hartmann (2021) uses neural networks in a machine-learning experiment that attempts to ascertain the phonetic values of *h_1, *h_2, and *h_3. The experiment yields some generally-accepted feature values for those three items, but also a few that border on the bizarre. Focusing on the latter, we are informed in table 23 ('Summary of detected phonetic features') on p. 63 that, *i.a.*, *h_1, *h_2, and *h_3 were [+voice], *h_1 and *h_3 were [+velar], and *h_1 was [+labial]. But such values make no sense in light of what we know of the contextual effects of these phonemes. Hartmann realizes this of course, and in the discussion immediately following table 23 tries to ameliorate the problem. And in the opening sentence of his Conclusion section (2021, 67), he has backed off to the position that '… the most likely values of the three laryngeals are [ɣ(ʷ)]/[x(ʷ)] > [ɦʷ] : [ʕ] : [ɣʷ]/[ʁʷ].' This is an improvement, but, in agreement with the *communis opinio*, we cannot accept the purported labiality of *h_1 or inherent voicing of *h_2.

effects that were generated: In parallel with what occurs in Interior Salish, the PIE faucals, *h_2 and *h_3, retracted neighboring *e to *a and *o, respectively, and *h_1 had no effect whatever on the position and/or quality of neighboring vowels.[25] This is in close keeping with the findings of Walker & Rose (2015, 2), who conclude their paper as follows:

> In sum, gutturals can show semi-transparency effects, with the potential to affect vowel quality and for laryngeal and supralaryngeal gutturals to behave differently. The phonetics of gutturals sheds light on these patterns. A phonological analysis informed by the production of gutturals makes better-fitting typological predictions than previous accounts.

Simply put, we argue that, with respect to RCC, the faucal fricatives (*h_2 and *h_3) and the laryngeal obstruent (*h_1) operated as two distinct phonological classes.

8 Allowability of *h_1 repetitions

But why should repetitions of just *h_1 in the root have been licit? We propose that the reason is straightforwardly phonetic, based on the findings of Kehrein & Golston (2004), who conducted a wide-ranging and detailed examination of the world's languages. In their abstract (2004, 325), they write:

> [W]e show that voicing, aspiration and glottalisation occur at most once per onset, nucleus or coda in a given language, and that the order in which they are produced within onset, nucleus and coda is never contrastive. To account for these restrictions, we propose that laryngeal features are properties not of segments, but of the onsets, nuclei and codas that dominate them.

They go on to say (2004, 325f.):

> Following Ladefoged & Maddieson (1996, 2), our study focuses on the elements 'that are known to distinguish lexical items within a language', i.e. on laryngeal contrasts involving voicing, aspiration and glottalisation that account for minimal pairs…. The facts that we present here suggest that natural languages allow for at most a single unordered set of laryngeal features per margin or nucleus, *whatever* the number of segments in that domain. For this reason, we propose that … [a]n onset, nucleus or coda has a single unordered set of laryngeal features. The idea that laryngeal features may characterise prosodic levels above the segment is not new of course…. What is novel is our claim that

[25] The relevance of evidence from Salish for the differing effects of PIE *h_1, h_2, and h_3 on neighboring vowels is underscored by Vine (2002, 294ff.), whose analysis is based in good part on data and conclusions in Bessell (1992, 1998). Bessell (1992, 91) writes, *i.a.*, concerning Salish: 'Glottals are never reported to have effects comparable to those found in the context of uvulars and pharyngeals.'

> laryngeal features only characterise prosodic levels above the segment; segments never license these laryngeal features on their own.

So with this in mind, we see that, regardless of whether *h_1 was [h] or [ʔ] (or even in some contexts [ɦ]), the respective onsets and codas were different unitary items in *h_1reh_1- 'fragen', *h_1reh_1 'rudern', *h_1erh_1-, and *$h_1\underset{\frown}{i}eh_1$-, and thus these roots do not constitute counterexamples to RCC. Note that this explanation is unaffected by whether the 'row'-root is reconstructed as *h_1reh_1 or *h_1erh_1. Also, this formulation offers an alternative reason for eliminating *h_1eh_1s- (Item 1) as a counterexample, since it too would have a unitary onset and coda that were different.[26]

9 Why are combinations of *h_1 and *R prevalent?

We again raise the question here of why all four of *h_1reh_1- 'fragen', *h_1reh_1 'rudern', *h_1erh_1-, and *$h_1\underset{\frown}{i}eh_1$-, remarkably, involve margins having *h_1 and a resonant (i.e., glide or consonantal sonorant). Can we find a principled explanation for this? Perhaps. If we assume PIE *h_1 was, at the relevant stage, [h] and, following Kehrein & Golston (2004), we view these items as containing (perforce) unordered collections of the phonetic features of [h] and the resonant, and, further that these collections were interpreted as phonemic voiceless resonants, then the insights of Blevins (2018) can be brought to bear. She states in her abstract (2018, 31):

> One phonetic source of voiceless sonorants[27] is coarticulation in RH and HR and clusters, where R is a sonorant and H is a segment produced with a spread glottal gesture.

She goes on to say (2018, 31), '… [A] fair number of languages show voiceless sonorant glides, liquids and nasals phonologized as a consequence of RH/HR coarticulation …'. Based on data from Maddieson (1984), Blevins notes in her table 2 (2018, 35) that only 3.5% of the world's languages have phonemic voiceless resonants, but, significantly, states the following (2018, 35): 'While voiceless sonorant consonants may not be loud sounds, like /h/, they may be contextually salient, contrasting with surrounding voiced sounds…'. And on page 47: 'Known cases of contrastively voiceless sonorant consonants arise from RH or HR clusters. Voiceless sonorants with these origins are relatively stable….' This relative stability would explain why the relevant items survive.

---

[26] We note, moreover, that our solution predicts that *h_1eh_1- is not a possible PIE root, since it would violate RCC; and indeed, no such PIE root is posited in, e.g., *LIV* or *EIEC*.
[27] I.e., what we are calling *resonants* here.

10 Conclusion and general speculations

In sum, we conclude that there are no clear-cut viable counterexamples to the RCC.

For purposes of further research, we think it is worthwhile to end with a speculative remark concerning the relation of the RCC to the prehistory of PIE: While the RCC is one of several root constraints reconstructed for PIE (as discussed earlier in this paper), it is the one easiest to construe as resulting from a familiar process of dissimilation that alters one of two identical consonants across a syllable nucleus (see Bennett [2015] for a survey and characterization of types of dissimilation, and Ohala [1993] for possible motivations for dissimilation and their consequences for sound change).

Support for such a proposal would most likely depend on lexical items shared by PIE and other language families in which the RCC is not evident. We note that a possible motivation for the RCC as an innovation in PIE is the avoidance of identical consonants coming together in the zero-grade of roots, since PIE does not evidence intra-root geminates among its consonant clusters.

Comments invited

PiHPh relies on post-publication review of the papers that it publishes. If you have any comments on this piece, please add them to its comments site. You are encouraged to consult this site after reading the paper, as there may be comments from other readers there, and replies from the author. This paper's site is here:

https://doi.org/10.2218/pihph.7.2022.8959

Acknowledgements

We are deeply indebted to Michael Weiss for his astute judgments, useful references, and thought-provoking remarks. We are also grateful to Juliette Blevins, Michela Cresci, and Patricia Keating for their helpful comments. And in our previous work on this subject, important suggestions were made by anonymous reviewers. Naturally, it should not be assumed that any of the aforementioned agrees with our positions.

Author contact details

Paul S. Cohen
3271 Nutly Circle
Yorktown Heights, NY 10598
USA

pausyl@aol.com

Benji Wald
5452 Barton Avenue
Los Angeles, CA 90038
USA

bjwald@gmail.com

References

Bennett, William G. 2015. *The phonology of consonants: Harmony, dissimilation, and correspondence*. Cambridge: Cambridge University Press.

Benveniste, Émile. 1935. *Origines de la formation des noms en indo-européen*. Paris: Adrien-Maisonneuve.

Bessell, Nicola J. 1992. *Towards a phonetic and phonological typology of post-velar articulation.* Vancouver, Canada: The University of British Columbia Ph.D. dissertation.

Bessell, Nicola J. 1998. Phonetic aspects of retraction in Interior Salish. In Ewa Czaykowska-Higgins & M. Dale Kincaid (eds.), *Salish languages and linguistics: Theoretical and descriptive perspectives* (Trends in Linguistics, Studies and Monographs 107). Berlin & New York: Mouton de Gruyter.

Blevins, Juliette. 2018. Evolutionary phonology and the life cycle of voiceless sonorants. In Sonia Cristofaro & Fernando Zúñiga (eds.), *Typological hierarchies in synchrony and diachrony*, 31–58. Amsterdam: John Benjamins.

Cohen, Paul S. 2015. A new etymology for PIE **mēms-* 'meat, flesh'. *Linguistica Brunensia* 65. 57–70.

Cohen, Paul S. 2017. PIE telic *s*-extensions and their diachronic implications. In Bjarne Simmelkjær Sandgaard Hansen, Adam Hyllested, Anders Richardt Jørgensen, Guus Kroonen, Jenny Helena Larsson, Benedicte Nielsen Whitehead, Thomas Olander & Tobias Mosbæk Søborg (eds.), *Usque ad radices: Indo-European studies in honour of Birgit Anette Olsen*, 118–133. Copenhagen: Museum Tusculanum.

Cooper, Adam I. 2009. Similarity avoidance in the Proto-Indo-European roots. *University of Pennsylvania Working Papers* (= *Proceedings of the 32nd Annual Penn Linguistics Colloquium*) 15(1). 55–64.

Corbeau, Alain Louis. 2013. *Phonological root structure constraints in Proto-Indo-European*. Leiden: Leiden University MA thesis.

EIEC = J. P. Mallory & D. Q. Adams (eds.). 1997. *Encyclopedia of Indo-European culture*. Chicago & London: Fitzroy Dearborn.

Hartmann, Frederik. 2021. The phonetic value of the Proto-Indo-European laryngeals: A computational study using deep neural networks. *Indo-European Linguistics* 9. 26–84.

Kehrein, Wolfgang & Chris Golston. 2004. A prosodic theory of laryngeal contrasts. *Phonology* 21(3). 325–357.

Kloekhorst, Alvin. 2008. *Etymological dictionary of the Hittite inherited lexicon*. Leiden & Boston: Brill.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford & Malden, MA: Blackwell.

LIV = Helmut Rix & Martin Kümmel (eds.). 2001. *Lexikon der indogermanischen Verben*. 2nd edition. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lubotsky, Alexander. 1985. The PIE word for 'dry'. *Historische Sprachforschung* 98. 1–10.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.

Melchert, H. Craig. 1994. *Anatolian historical phonology*. Amsterdam & Atlanta: Rodopi.

Ohala, John. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237–278. London: Longman.

Pinault, Georges-Jean. 2018. La linguistique historique et la tentation de l'Indo-Européen idéal. In Bernard Colombat, Bernard Combettes, Valerie Raby & Gilles Siouffi (eds.), *Histoire des langues et histoire des représentations linguistiques*, 23–54. Paris: Honoré Champion.

Ringe, Don. 2006. *A linguistic history of English, volume I: From Proto-Indo-European to Proto-Germanic*. Oxford & New York: Oxford University Press.

Rose, Sharon. 1996. Variable laryngeals and vowel lowering. *Phonology* 13(1). 73–117.

Rose, Sharon. 2000. Rethinking geminates, long-distance geminates, and the OCP. *Linguistic Inquiry* 31(1). 85–122.

Stifter, David. 2017. Varia II: The origin of time. *Ériu* 67. 219–226.

Vine, Brent. 1998. Indo-European and Nostratic: Some further comments (a response to 'Exploring the Nostratic hypothesis'). In Joseph C. Salmons & Brian D. Joseph (eds.), *Nostratic: Sifting the evidence*, 85–105. Amsterdam & Philadelphia: John Benjamins.

Vine, Brent. 2002. Gk. *σφήν*, Eng. *spoon*: A note on 'Eichner's Law'. *Münchener Studien zur Sprachwissenschaft* 62. 289–299.

Walker, Rachel & Sharon Rose. 2015. Guttural semi-transparency. Handout for paper given at the 2015 Annual Meeting on Phonology (AMP 2015). Vancouver, Canada, October 10, 2015.

Watkins, Calvert. 2000. *The American Heritage dictionary of Indo-European roots*. 2nd edition. Boston & New York: Houghton Mifflin.

Weiss, Michael. 2020. *Outline of the historical and comparative grammar of Latin*. 2nd edition. Ann Arbor, MI & New York: Beech Stave.