On Indo–European /a/

Xaverio Ballester *

DOI: https://doi.org/10.54392/ijll2342
Received: 24-11-2023; Revised: 06-12-2023; Accepted: 10-12-2023; Published: 22-12-2023

Abstract: Although the original phonemic vowel pattern of the founding era of Indo–European Linguistics (/a i u/) was largely abandoned by the end of the 19th century, there are good reasons to rethink its recovery and functionality. The main reason that led to the rejection of the presence of /a/ in the original vowel pattern can no longer be maintained today. The so–called “new sound of Indo–European” is indeed too new and too unsound. What we need is a real old Indo–European sound. Old and sound too.

Keywords: Proto-Indo-European, Phonology, Vocalism, Vowel /a/

The Aryan Law that Changed All, the French Law that Changed Nothing

One of the very first descriptions of the Indo–European vowel pattern, very worthy of mention, is the proposal of a phonemic inventory with only three vowel sounds: /a i u/. It was a proposal, however, that was perhaps too soon abandoned. Indeed, for official Indo–European Linguistics the absence of /a/ became in practice an axiom, so that, in almost any later proposal on Indo–European vocalism, most scholars accepted the idea that the vowel /a/ would never have existed. Consequently, all the obvious cases of presence of /a/ in Indo–European material were explained by means of varied and bizarre arguments of a kind of derogatory, childish or popular vocalism and finally rejected as quantitatively not significant. One cannot hide that the presence of /a/ in Proto–Indo–European would be barely compatible with the then emerging academically dominant laryngeal theory. Beekes proclaimed (1991: 238): «I consider as one of the most important insights provided by the laryngeal theory that PIE had no phoneme *a=. On the contrary, for us, as we shall try to demonstrate, precisely the absence of *a/ in Proto–Indo–European would be one of the best indications of the fallacies and fallibility of the laryngeal theory. Anyway, the slogan «Against a Proto–Indo–European phoneme *a= (so litteratim Lubotsky 1989) has become one of the most frequent proclamations of the laryngealist dogma. So, everyone against *a!

However, if considered today without prejudice, the argument that motivated the rejection of the primitive Indo–European /a/ presents, at least from a current phono–typological perspective, no validity at all. In order to demonstrate this assertion and to briefly go over the matter, we will invoke here an objective testimony on this subject. Szemerényi (1999: 134) wrote:

«Influenced by the antiquity of Sanskrit, the founders of the IE linguistics and the immediate successors assumed that the Old Indic triangular system i–a–u represented the original situation. In 1864 G. Curtius drew attention to the fact that in many cases European languages opposed e to the Sanskrit a; cf. Gr. Ækta, Lat. demec, Goth. taihun, Lith. dešint but Skt. daša. He supposed, however, that in this respect all the European languages had innovated as a closed group, i.e. they had split the original a into e and a. It was not until 1871 that Arthur Amelung¹ came to realize that the European e as opposed to Sanskrit a represented the original situation, though this view did not win general acceptance until later, with Brugmann’s famous article of 1876². The originality of the (European) e was then proved within Old Indic also by the discovery of the Aryan law of palatalization». 

¹ In reference to Die Bildung der Tempusstämme durch Vocalsteigerung im Deutschen.
² The reference is to Karl BRUGMANN, «Zur Geschichte der stammabstufenden Deklinationen», Curtius’ Studien 9 (1876) 367–8 and 380–1.
This decisive “law” of Aryan palatalization would consist, again in the words of Szemerényi (1990: 63), in that «In Aryan \(k\ g\ g\ h\) were palatalized to \(č\ j\ h\) before \(e\) (which later became \(a\)) and \(y\)». It is the so-called “law of palatalals” or “law of Collitz–Saussure” (Cavazza 2004: 148), conjecturing the existence of an */e/* in the prehistoric phase of Sanskrit. Thus, for example, from the enclitic ‘and’ \([t\]a\] in Sanskrit (–\(ca\)) or Avestan (–\(ča\)) as opposed to Celtiberian –\(QVE\) or –\(CuE\), Greek \(τε\) or Latin –\(que\), a previous form *–\(kue\) was deduced and not *–\(kua\), since — so it was said— otherwise the consonant could not have been palatalized. It was assumed, therefore, that after the regular loss of [\(w\)] {{\([kwa > ka]\)}} the dorsal consonant [\(k\)] should have remained unchanged in the Indo–Iranian languages, but the existence of a palatalization in [\(t\]a\] could only mean, according to those neogrammarians, an original *–\(k^*e\) sequence. Similarly, the existence of Avestan and Sanskrit \(j\an\i\)– with a palatalized consonant in contrast to Greek \\(\upsilon\upsilon\upsilon\) or Gothic \\(qin\o\) would invite one, under the effects of this Aryan law, to reconstruct a root *\(g^*\text{en}–\text{woman}\) etc. etc.

It seems to us, however, that exchanging a stable vowel model for such Aryan minutiae has really been a methodological abuse, the consequences of which for Indo–European phonology have been, in our opinion, catastrophic. Thus, according to such flamboyant argument, the antiquity of */e/* and not of */a/* should also be seen in, for example, the Italian, Portuguese or Spanish campo ‘field’, since French also presents the palatalized velar in champ \([f\]a\]... and in charbon ‘coal’ but Latin carbo, chante’s/he sings’ but Latin cantat, etc. etc. So, according to some law of Frankian palatalization that original Latin root should be reconstructed as ‘\(cemp\)’ and not as it surely was: campo. Fortunately, today we have enough documentation to admit that a consonant followed by [\(a\)] or before [\(\text{ae}\)] and, naturally, by [\(k\ e\)] and similar vowels— may very well be palatalized. This is, for instance, the case of the modern Persian, where «velars tend to be fronted before front vowels, including » (Windfuhr 1997: 681).

The fact is that, at least before [\(a\)] or before [\(\text{ae}\)], palatalizations also occur in many languages, among which, as we see, there are some that are very close and well known. Thus, in French we have char ‘car’ (Latin carrus) or chose ‘thing’ (Latin causa), words that stemmed always from a root with [\(a\)], and where a palatalizing [\(e\)] very probably never existed. Alternations as for ‘leg’ Italian gamba – French jambe or for ‘shank’ Italian garetto – French jarette would show the same palatalization before [\(a\)]. Also in Romansh, possibly influenced by a Celtic superstratum (see Alnei 2000: 742), we have palatalization of velars followed by [\(a\)]/\(\text{Ian}\)– ‘dog’ (Latin canis), ciastel ‘castle’ (Latin castellum), or gialina ‘hen’ (Latin gallina). Likewise, to quote now a non–Indo–European example, in the Turkish of the Oghuz group «\(k\) and \(g\) are palatalized in front of anterior vowels and an \(â\) of Arabic–Persian origin» (Manuzelli 1993: 558). All this is not surprising, if compared to what happens in some Lak dialects: «Curiously, Turči and Shuni govory palatalize /\(k\)/, etc. to [\(č\)] only preceding [\(a\)] and not [\(I\)]» (Anderson 1997: 991).

Out of the Vocalic Eden

Thus, a rough and rigid topic of traditionalist Indo–European Linguistics is the denial of the presence of the proto–phoneme [\(a\)]. This active opposition to [\(a\)] should not surprise us, since the recognition of this phoneme would show the same palatalization before [\(a\)] or before [\(\text{ae}\)] —or, in the most extravagant versions, of [\(e\)] or [\(o\)] (\(\text{si}\))— in the primitive phonemic pattern.

The truth is that the vowel [\(a\)] is the champion of the vowels, because statistically it is present in almost all the languages of the world and its absence from a vowel phonemic pattern is coincidentally highly improbable from a statistical point of view. Therefore, before denying the existence of salt in the reconstructed Indo–European ocean and conjecturing an otherwise unknown chemical component, it is necessary to try to reconstruct that impressive mass of oceanic water as salty as well, because almost all other known oceans have salt. And since the high frequency of a compact vowel [\(a\)] ([\(\text{ae} \text{ ...}\)]) is also well documented in so many languages, it would be completely necessary to explain how the habitual Scipio of the phonemes —the first to enter battle and the last to leave— would have disappeared precisely at the critical, crucial and decisive moment of the disintegration of the ancestral mother tongue, as «the most common vowel in nearly all languages is \(a\)» (Ladefoged 2001: 160). Certainly, synchronic Typology has shown long ago that the phoneme [\(a\)] is the most resistant of vowels (Jakobson & Waugh 1980: 133) besides [\(i\)]. Phonogenetically, [\(a\)] is probably the first vowel to appear, because its maximum opening and its scarce

3 «\(k\) e \(g\) sono palatalizzati davanti a vocali anteriori ed a \(â\) di origine arabo–persiana».  

characterization —not anterior or posterior, not labialized, not nasal, not...— makes of this open sound an ideal candidate for the embryo of a vocalic inventory. Ego sum alpha et omega, i.e. the beginning (/a/) and the end (/oː/).

And not only that. If the non–existence of the most allophonic —but also the most frequent and tenacious— of the vowels in a linguistic reconstruction is aprioristically very improbable, it is also very difficult not to take it into account in the reconstruction processes for the oldest phases of the Indo–European languages, where it would be like walking through a forest (vowels) without finding any tree (/a/).

As Many Vowels as Stars in the Sky

Probably, a mistake of traditional Indo–European Linguistics has been to fail to understand that vowels generally present greater allophonic variability —sometimes much greater— and consequently a potentially greater phonemic mutability —sometimes much greater— than consonants. Greppin (1997: 790): expressively says «The vowels of Armenian dialects are as varied as the constellations of the heavens. The original short vowel system of a e i o u can become almost anything». The explanation of this phenomenon does not seem difficult or arcane. We know that in a communicative model an element becomes more significant when it is less frequent, for example, the presence in Latin or English of the grapheme <x> or the phoneme /h/ tells us more than the presence of the grapheme <s> or the phoneme /m/. Now, it is so that all known languages always have a number of consonants not less than the number of vowels —and usually a number at least double— ergo the consonants will be, under equal conditions, always more significant, more important and, therefore, less likely to change than vowels.

Consequently, in any case, in the reconstruction of Indo–European or any other linguistic group, one must never forget the principle that we cannot operate with vowels and consonants as parallel entities in the number of changes. The Indo–European series for ‘moon — month’: Albanian muay, Armenian amis, Gothic mēno-, Greek μύη, Irish mī, Latin mēnis, Lithuanian mēnuo, Russian месец, Sanskrit mās-, Old High German māno, Eastern Tocharian mañ, would always better preserve the consonants of an ancient Indo–European root *man/aʃ— than the vowels (māno, μύη, mi, Mond, moon, muay...). Likewise, in ‘mother’, from a probable root *matar— or *mātar—, we can, according to the various Indo–European languages, draw a vast vocalic arch with at least five vowel timbres in the first syllable: Latin māter, Greek Modern ἐμήτερα/, Prussian mūṭa o German Mutter, French mère and Lithuanian motina. Indeed, the consonantal structure /m t r/ has generally been much more stable than the old /a/. Also likely from a *pāntas ‘ford – jetty’ we have Vedic pāṁthā— ‘road’, Prussian pintis ‘road, Armenian hun ‘ford’ or Serbo–Croatian put ‘road – travel’ and Greek ντος ‘see’ or Latin pont– ‘footbridge – bridge’. Similarly, from a base *ugnis or *agnis ‘fire’ we would have Old Indic agnih, Latin ignis, Lithuanian ugnis and Old Slavic ognis, etc. That this greater vowel variety is not only a cross–linguistic phenomenon or a result uniquely proper to vast territories is also demonstrated by the existence of the same trait in the dialectal sphere and in much smaller territories. Thus, for example, in Romansh we find according to the dialects for ‘fire’ [fek], [fok], [fs], [fwa],[fwo] or [fyk], and for ‘time’ [temp], [temp], [tamp] or [tomp] (Dell’Aquila 2006: 254), presenting the old Latin /a/ as a somewhat special problem (Dell’Aquila 2006: 264).

Logically, as a very common vowel in so many languages, /a/ tends to present a great allophony, especially in languages where there are only three vowel phonemes or timbres such as /a i u/, and where /a/, perhaps the most basic of the vowels, that undergoes more changes. In fact, in a trivocalic model /a i u/ the vowel /a/ usually enjoys a great virtual allophony: [a æ ə a ɔ ã e ø ə e u ɔ u] and, although in competition with /i/ and /u/ and in competition with /u/, also [e e] and [o o] tend to be direct allophones of /a/, respectively. Thus, the /a/ in Harari knows the variants [a a æ e] (Wagner 1997: 488). In the Berber Fiquig dialect /a/ can be realized [æ ə o] depending on the context (Kossmann 1997: 45). In Shilha or Tashelhiti Berber, whose vocalic pattern is /a i u/, «/i/ tends to becomes /i, e, e..., /u/ becomes respectively /u, o, o, ò, oe.../ and [...] /a/ is perceived as [æ, a, o...]>» (Okukrim 1995: 33). In Arabic, depending on the context, /a/ is realized [a a æ æ] (Kaye 1997: 197). In Ubykh /a/ «may be realized as [a, æ, o]» (Campbell 2000: 1703). In Tadjik /a/ according to phonetic context is realized [a æ æ æ] (Rastorgueva 1992: 5). In Burushaski, likewise depending on the context, /a/ is pronounced [æ æ æ] (Anderson 1997: 1029). In Marathi /a/ can be realized [] (Campbell 2000: 1089). In Sindhi short /a/ is realized [a] (Campbell 2000: 1500). In the non–final syllables of Belait «/a/ varies freely between [a æ æ]» (Clynes 2005: 431). Again Greppin

4 In Prussian we have *ā > ā after labial and velar consonants.
5 «/a/ tiende a convertirse en [i, e, e...], /u/ hace lo propio hacia [u, o, o, ø, o... y [...] /a/ se percibe como [æ, a, o...]».
(1997: 790) for the Armenian Van dialects: «/a/ can become a ā i ī y eo eo u īw e a and zero». In Scottish Gaelic «the phoneme /a/ can yield a wide variety of realizations in the [ae – a – a] sector» (Gillies 2002: 152). Yagua /a/ knows the allophones [a ae e] (Wise 1999: 315) and Arabela /a/ is realized [a ae e o a] (Wise 1999: 317). It will be noted that, logically, the lower a phonetic inventory is, the higher its allophony will regularly be, so that «in a three vowel system the areas are larger and more vague than in a more complex system [...] The vowels of three vowel systems often show considerable subphonic variation», so Crothers (1978: 109), who gives the example of Greenland Eskimo with [æ a] for /a/, [i e a] for /i/ and [u y o e] for /u/ and recalls that in the pattern /a i u/ the last vowel often presents the widest variation ([u o ...]).

### The Major Vocalic Triad: /A I U/

As Ladefoged (2001: 159) points out: «Probably every language uses at least three distinct vowels [...] Languages that have only three vowels usually have sounds that can be symbolized I, a, o or i, a, u [...] Languages use these three vowels extensively because [...] they are far apart in the vowel space». According to Crothers (1978: 105) the pattern /a/[...] i[...] u[...]/ is the third most common vowel model and is represented in 23 of the 209 languages studied by this author. Indeed, it would be a linguistic universal the fact that «All languages have /i a u/» (Crothers 1987: 115, 134, 136). Furthermore, /a i u/ are the most frequent oral vowels and /ā ū ū/ the most frequent nasal vowels (Ladefoged & Maddieson 1996: 298). Certainly, from the phonetic point of view /a i u/ represent the optimal contrast —being the most extreme and distinctive cardinal units— amongst vowels (Fant 1973: 11, 187, 189).

In fact, the pattern /a i u/ is documented in languages such as Amorite (Buccellati 1997: 23), Berber (Kossmann & Stroomer 1997: 463, 468), Aleutian (Greenberg 2000: 50), the Caddoan dialects (Del Moral 2002: 114), Cheuqua (Campbell 2000: 1380) or Aymara (Sala 1998: 97), but is also at the base of so many other languages, thus */*a i u/ looms behind models like the /a aː iː uː/ of so many Afroasiatic languages, while still being apparent in Arabic (Kaye 1997: 196), Eblaite (Gordon 1997: 42), Ancient Hebrew (Rendsburg 1977: 76), Old Egyptian (Loprieno 1977: 440), Phoenician (Segert 1997: 60), Sayhadic (Kogan & Korotayev 1997: 223) or Ugaritic (Gordon 1997: 51). The Berber from Figuig <i>distinguishes the three full vowels a, i, u and the “neutral” vowel e ([a])</i> (Kossmann 1997: 45). Also for the Western Sudanic subgroup of the Niger–Kordofanian languages, some authors have proposed a primitive model /a i u/ (<i>cfr. Williamson & Blench 2000: 37</i>). «Proto–Oceanic had five vowels (as does modern Fijian) —the three proto–Austronesian vowels *i, *u and *a, and also e and o, which developed from proto–Austronesian *ay and *aw respectively» (Dixon 1988: 9). In Old Malay there were long and short /a i u/ with /e/ and /o/ only in words copied from Sanskrit (Mahdi 2005: 188, 189). Austronesian Cham also has /a aː iː uː/ (Del Moral 2002: 128). Among the Australian languages we find the simple pattern /a i u/ or similar in Arabana–Wanganura (Maddieson 1984: 332), Aranda (Maddieson 1984: 330), Diyari (Maddieson 1984: 332), Gugu–Yalanji (Maddieson 1984: 331), Kariera–Ngarluma (Maddieson 1984: 330), Nyangumarta (Maddieson 1984: 329), Nunggubuyu (Maddieson 1984: 325) and Western Desert (Maddieson 1984: 329). Indeed, the Australian vocalic pattern «is notably uniform: there is one scheme with three vowels /a, i, u/, while /e/ and /o/ are less common and may be no phonologic»7 (De Meo 1998: 197). In Europe «An early form of Gothic may have had a short vowel system consisting of three members: /a/, plus /i, u/ with allophonic variants» (Lehmann 2002: 23). In America one can mention Kootenai with /a aː iː uː/ (Mithun 2001: 452), Ifiupiaq with /a aː iː uː/ (Rice 2004: 340), the Muskogean languages with «i, a, and o, with contrastive length and nasalization» (Mithun 2001: 464) or Aymara with /a aː iː uː/ (Campbell 2000: 158). Also, Papantla Totonac and Zapotitlán Totonac would have only three timbres /a i u/ but twelve vowel phonemes (Levy 1987: 9, 15), the Coatepec Totonac and the Ahuacatlán Totonac have three timbres and six phonemes: /a aː iː uː/ (Levy 1987: 14, 16) as well. The 23 /a i u/ languages listed by Crothers (1978: 138) in his survey are as follows:

6 «distingue les trois voyelles pleines a, i, u et la voyelle “neutre” e ([a])».
7 «è notevolmente uniforme: esiste uno schema a tre vocali /a, i, u/, mentre /e/ ed /o/ sono meno comuni e possono non essere fonologiche».
Also, situations like /aː iː uː eː oː/ in, for example, Balochi (Elfenbein 1997: 765) or Brahui (Elfenbein 1997: 799) would be better explained by interpreting /aː iː uː eː oː/ as later amplifications. The Chorezmian knew /aː iː uː eː oː/ (Reczek 1986: 152) and Yaghnobi presents /aː iː uː eː oː yː/ (see also Skalmowski 1986: 207). The Gaulish inventory probably contained the vowels /aː iː uː eː oː/ (Lambert 1997: 41).

In Karuk «Apex vowels /i, a, and u/ occur both short and long, but the mid vowels are always long: /e, o/» (Mithun 2001: 435). In Kamera «The cardinal vowels /a, i, u/ may occur in both stressed and unstressed syllables of the root, while /e, o/ may only occur in the first, stressed, syllable of the root» (Klamer 1998: 16). In addition, in the languages mentioned so far and in many other languages, such as Swahili or Uyghur, /a i u/, unlike the other vowels, often undergo joint treatment, that is, as a series.

The Triumvirate Again but Disguised

In fact, an original model consisting of three vowel bases with the three cardinal timbres can be glimpsed with some ease behind almost all the major linguistic groups in the world: V a V i V u, that is, /a i u/ in the most common phonetic realizations.

---

8 And also nasalized the corresponding long ones.
and in correspondence with the three great vertices of the vocal apparatus and with the «three major places of articulation in consonants» (Clements 2000: 127) and —let's add— in vowels: dorsal (/a/), coronal (/i/) and labial (/u/). The basic nature of /a i u/ can be verified not only by the existence of numerous languages that directly show this vowel model, but also in the emergence of various phonetic phenomena, where it can be deduced without much difficulty that such a model is latent. One of these phenomena would be the common epenthesis —or *hanza* in some grammatical traditions— since as a function of such an epenthetic consonant, only the assyllabic variants of /a i u/ are presented, that is, [ʔh j w] in the corresponding allophones according to the reference language. Thus, when a language does not present only the apex vowels /a i u/, its essential or cardinal character can still be deduced from various phonetic phenomena. For instance, in so many African languages with nasalized vowels, the presence of [ʌ i û] is more frequent than the presence of the other nasalized vowels (Clements 2000: 139). In Kisi [a i u] «occur more frequently than do the mid vowels [e õ o]». The peripheral vowels, for example, are common and the central vowels rare in both ideophones and affixes, and at the end of stems» (Childs 1995: 36), so that «verb extensions, primarily suffixes, feature only the peripheral vowels [i u a]» (Childs 1995: 305 n15). Furthermore, the vowel triumvirate is generally the one adopted epenthetically in Kisi to adapt consonant clusters from other languages, thus we have French *drapeau ‘flag’ > dálapô*, or English *school > sükû* (Childs 1995: 64). In Ainu, that has /a i u e o/ as vocalic pattern, final consonants, especially /r/, often develop a final vowel [a i u], usually according to the preceding vowel (Refsing 1986: 69). In Central Yiddish the long vowel series contained only /a: i: u:/ (Jacobs &. et al. 2002: 391). In Seediq: «The vowel in the final syllable is /i/, /u/, or /a/» (Tsukida 2005: 292). In Old Icelandic there was only /a i u/ in unstressed position (Faarlund 2002: 42) and still nowadays «the only vowels that occur in completely unstressed syllables in native Icelandic words are /i, u, a/» (Thráinsson 2002: 149). After a stressed syllable in Faroese there is only /a i u/ (Barnes & Weihe 2002: 191). In fact, in those languages that suffer vocalic reductions to three unstressed vowel timbres, such as Russian or Catalan, the phonemes that correspond to the timbres /a i u/, though usually made less tense ([a i v]), are the most resistant. In the Kalapuyan group, albeit the vowel inventory contains /a aː iː uː/ only the diphthongs /au au iu/ are admitted (Mithun 2001: 432). Although Lacota has /a i u e o/ as oral vowels, it only has /ã ũ ù/ as nasal vowels (Mithun 2001: 506).

As we saw (cfr. Ladefoed 2001: 159: «only three vowels [...] symbolized i, a, o or i, a, u»), the phonemic model with /a i o/, that is, the one where /u/ and /o/ do not contrast, may in many cases very well derive from an older model with /a i u/. We find /a a: i: o o:/ in Blackfoot and /a i o/ in Tsuut’ina (Rice 2004: 340).

A pattern just like /a i u/ —for example, /a i u a: i: u:/ in Old Persian— does offer great balance and simplicity, but has the inconvenience of its high degree of allophony, and is consequently conducive to enlargement. Therefore, it is not surprising that diachronically many languages have added phonemes with other vowel timbres, usually adding /e/ and later /o/, and this seems to have been exactly the process in the Indo–European linguistic *continuum*.

### From Three to Infinity

Omitting now the here non–determining issue of the phonemic status of vowel quantity, the fact is that there are Indo–European languages with inventories of three vowel timbres /a i u/ and five /a i u e o/. It is thus quite likely that the Indo–European languages with four timbres /a i u e o/ occupy a diachronically intermediate position. For traditional Indo–European Linguistics, the meaning of the process can hardly be other than a preposterous route from five to three. Unfortunately, we are already too used to hearing without being astonished that the ancient Indians have simply [con]fused a more complex vowel —but without /a/, without /i/ and without /u/— pattern into a simpler one with three *new* vowels /a i u/. Obviously, all this sounds rather *contra naturam* and quite uncogent. It seems that these authentic *Aryans* —perhaps due to partial deafness (†) or other generalized hearing disease (‡)— had enormous problems in no longer being able to capture the different acoustic personalities of such *difficult* vowels as /e/ and /o/, since here we certainly cannot blame contact with non–Indo–European Dravidians —whose languages tenaciously retain the /a i u e o/ of the common Dravidian phase (Steever 1998: 13)— for this merging of /e/ and /o/ with /a/.

In any case, the typological —synchronic and diachronic— testimony —and one could add: common sense as well— points again in the opposite direction. Typologically, it therefore seems much more plausible that Indo–European vocalism has gone through the three successive banal phases of development well documented in other historical languages, i.e. going through the following timbres and corresponding phonemes, these perhaps also in their double quantitative version of long and short:

i. /a i u/
The First Two Steps

The first phase (= /a i u/) would be documented directly in Luvian, which «has only three vowels, /a/, /i/, and /u/, in contrasting short and long varieties» (Melchert 2008: 35), and in Vedic, Sanskrit and Old Persian, but it would also be clearly reconstructable for the entire Indo–Iranian group. So, Lubotsky (2018: 1875): «PIIr had only two vowels; a and ā. Most probably, they were distinguished [(a) or (ā)] than ā ([(êːː)]) [...] there also were [i] and [u] [...] allophones of the phonemes /j/ and /y/».

The second phase /a i u e/ would be —more directly or indirectly— documented in non–Luvian Anatolian, Albanian, Baltic, Germanic, Slavic and Thracian languages and dialects. Indeed, in most languages with four vowel timbres, the fourth, next to the optimal /a i u/, is an intermediate, unlabeled opening vowel, like /e /a/. Normally, [o ɔ] and the like are rather considered as allophones of /u/ in this phase. In Akkadian the vowel pattern appears to be composed of /a i u e/ with [ə] as an allophone of /u/ (Buccellati 1997: 22–23). It is rarer to find inventories with a greater number of labial vowels (V^2) than coronal vowels (V), and, by the way, the same happens with consonants. Bissa, one of the two languages of the Mande group, has /a i u e o i/ and the respective long ones, while Ligbi, the other Mande language, has the much more common symmetric pattern with /a i u e o s/ and the respective nasalized vowels (Kropp & Nadén 1988: 158, 161). We find /a i u e/, for example, in Sumerian with /a aː i iː u uː e eː/ (Jiménez 1998: 23), Iloko (Rubino 2005: 328), Malagasy (Rasoloson & Rubino 2005: 460), Proto–Eskimo (Greenberg 2000: 50), Pawnee with /a i u e/ and distinctive vowel quantity (Mithun 2001: 371), Molala with /a aː i iː u uː e eː/ (Mithun 2001: 459), the Shastan languages with /a i u e/ (Mithun 2001: 498) or Pipil with /a aː i iː u uː e eː/ (Campbell 1985: 26).

Naturally, in many cases the series /a i u a/ can also very well represent an old series */a i u a/; since in many languages [a] is the historical result of an unstressed, weak or reduced vowel, especially /a/. In these cases, the secondary character of [a] is evident in various phenomena, such as the non–existence of /a aː/ when the rest of the vowels do have their corresponding long version. Thus, [a] does not have a long correlate, unlike almost all other vowels (/a i u e o/) in Berber (Kossmann & Stroemer 1997: 472) nor in Coos (Mithun 2001: 397), or in phenomena such as the non–existence of a stressed [a] or its epenthetic vowel character in many languages. Thus, in Amharic [a] is added in poetry to a final consonant for reasons of rhyme or prosody (Leslau 1990: 421, 426). Also [a] is used as an epenthetic vowel in Ge‘ez (Gragg 1997: 177). In late Egyptian the unstressed vowels, especially the posttonic ones, merged with /a/ (Loprieno 1997: 444). In Malagasy «The vowel /a/ may be reduced to [a] in unstressed environments» (Rasoloson & Rubino 2005: 460). In Pashto [a] comes from vowel reduction (Skalmowski 1986: 185). In Kilivila «In word final position all vowels are frequently reduced to [a]» (Senft 1986: 13). In Taüya «A single unstressed vowel is optionally reduced to [a] if it is non–initial» (MacDonald 1990: 52). In Javanese /a/ is always unstressed (Campbell 2000: 815). In the literary language of Mansi /a/ appears only in unstressed syllables (Keresztes 1998: 393). In Romanian all unstressed vowels show the tendency to become [a] and in Swedish the unstressed vowels in a final position (Campbell 2000: 1569). In Albanian dialects [a] is stressed only in Tosk (Sanz 1996: 34). We also frequently find [a] as a result of the weakening of unstressed vowels in the Germanic continuity, as in Afrikaans (Donaldson 2002: 483), Danish (Haberland 2002: 319), Frisian (Hoekstra & Tiersma 2002: 509), Middle English (Van Kemenade 2002: 116), Norwegian (Askedal 2002: 221), Middle Dutch (Van der Wal & Quak 2002: 74), Middle High and Middle Low German (Van der Wal & Quak 2002: 92), and contemporary German (Eisenberg 2002: 350, 353). The same happens in the Celtic continuity, as in Scottish Gaelic (Gillies 2002: 148), Irish (Mac Eoin 2002: 107) or Manx (Broderick 2002: 232). In Eastern Catalan [a] emerges only as a realization of unstressed /a/. We also find epenthetic [a] in Modern French (le cheval ‘the horse’ [la ʃəvəl / l ʃəval]) and often in Armenian.

Of course, also the series /a i o e/ can also be considered a variant of /a i u e/. A pattern /a i o e a/: i: o: e:/, for example, is postulated for Proto–Algonquian (Mithun 2001: 337–338). Mutatis mutandis, the same goes for /i/,
Close Encounters of the Third Kind

Historically, the third phase /a i u e o/ would have been reached by Armenian, Celtic, Hellenic, Italic — including Latin— and Tocharian languages. Since the model /a i u e o/ is probably the most common —because it represents, an optimal model according to Crothers (1978: 104)— a natural tendency to go from /a i u/ to /a i u e o/ cannot be surprising. Certainly steps like /a i u > a i u e > a i u e o/ are well documented. Thus, /e/ seems the most likely first expansion of /a i u/. Proto–Austronesian likely had /a i u/, often /e/ and /o/ being results of /ai/ and /au/. Still in colloquial Indonesian [e] and [o] are frequent realizations of the diphthongs /ai/ and /au/ respectively (Ewing 2005: 229). An evident proof that the model /a i u/ can generate a much richer and more complex vowel inventory, is also provided by modern Aryan and Semitic languages, which, starting from a model with only three vowel timbres /a i u/, have developed a model so much richer, and, for example, «Many modern Arabic dialects have, however, developed other vowels such as /a/, /e/, /o/ etc.» (Kaye 1997: 196). Ge'ez, for instance, would have developed the vowels /e o æ a/ from the three ancient Afroasiatic timbres */a aː i iː u uː/, with /e/ derived from */[a]*/, /o/ from */[aw]*/, /æ/ from */[aim]*/ and /a/ from */[i]*/ (Gragg 1997: 177). Cypriot Arabic has /a i u e o/ (Borg 1997: 222). Tiberian Hebrew had /a i u e o⟩⟨/ (Khan 1997: 91). From three vowel timbres in Old Persian /a(ː) i(ː) u(ː)/ we have gone to /æ o i u/ in Modern Persian. Kurdish, probably due to Turkish contact, has expanded the old Iranian pattern to /a i u æ o i u/.

However, the reverse step (*[a]i u e o > a i u e > a i u/), nisi fallimur, is not documented except in the circumstance that it mediated the adaptation of foreign phonemes. Such, for example, would be the situation of Spanish words by Quechua, and this always with multiple nuances, because some Quechua speakers have also come to copy directly the /e/ and /o/ from Spanish.

In full coherence with all these data, we would hardly have any other possibility than to deny the Indo–Europeanism of the Indo–Iranian peoples, because it would be necessary to conjecture that an alloglottic population, non–Indo–European and speaking languages with just three vowel timbres /a i u/, would have adopted an Indo–European language with a different and richer vocalism. But against this perspective stands the resounding fact that in all these alloglottic populations that might be recipients of this Indo–Europeanization —that is, in populations with Burushaski, Dravidian, Munda or even Sino–Tibetan languages—we do not find that poor vocalism /a i u/ but precisely the contrary: a richer vocalism, so that, for example, the Dravidian languages, the best positioned as potential linguistic substratum, regularly present at least five vowel phonemes and, in fact, a model with five timbres and ten vocalic phonemes is postulated: /aː iː uː eː oː/ for Proto–Dravidian (Steever 1998: 13).

The model /a i u e o/ —note that we quote, à la tibétaine (Scharlipp & Back 1989: 23) in the probably glottogonic order—is quite well known, being the model, among other languages, of Spanish and Basque. According to Ladefoged (2001: 159): «About 20 percent of the world’s languages have five contrasting vowels».

Some /A/ Supporters, Some Arguments in Favour of /A/

It was to Hans Krahe’s credit throughout various works (1954, 1962, 1964a, 1964b, etc.) to have shown the probable ancient presence of /a/ in many forms of Old European (alteuropäische) hydronymy. In that same direction, Robert Schmitt–Brandt (1973: 112), following Anton Scherer («Der Ursprung der “alteuropäische” Hydronymie», Atti e memorie del vii Congressso internazionale di scienze onomastiche, Istituto di Gliottologia, Florence 1963, II 405–417, non vidimus) brilliantly notes the detail that «the “Old European” river names often show a–vocalism, where the corresponding Indo–European appellatives o–vocalism could be expected». This contrast evidently supposes that /o/ represents a more recent evolutionary stage of /a/ (see Schmitt–Brandt 1973: 113). Likewise, Francisco Villar has convincingly argued (1991: 164–169) the most recent character of the opposition /a ∼ o/ and the plausible antiquity of /a/ (/a > a o/). A solid argument for this is that «it does not occur on any occasion that the same

9 «die “alteuropäischen” Flüssenamen häufig a–Vokalismus aufweisen, wo die entsprechenden indogermanischen Appellativa o–Vokalismus erwarten ließen».
linguistic tradition is attested first as an a/o language and then as an /a/ language (Villar 1991: 165). Similarly, Berenguer (1997: 48): «the reconstruction of an o-grade as primitive versus an a-grade lacks positive justification altogether».

In turn, it is again a great merit of Schmitt–Brandt in having seen that the persistence of /a/ in clearly onomatopoeic words, typical of children's lexicon or expressive words is another good indication of the antiquity of the vowel, since this kind of words tend to escape the normal phonetic evolutions of common heritage terms. Schmitt–Brandt (1973: 100) highlights roots such as, among others, *bab- or *pap- 'swell', *baba- and *balbal- 'babble', *bata- 'chatter', *kuak- 'croak - caw - quack', *haha- 'laugh', *kak- 'crow', *nana- 'mother - grandmother', *pappa- 'father - food'... So, to quote just one clear example, we have Armenian xaxank ‘laugh[ter]’, Old High German kachazen ‘to laugh’, Greek κοξίκω ‘I laugh’, Latin cachinno ‘I laugh’, Sanskrit kákhati’s/he laughs’ or Old Slavic xoxotat ‘to laugh’. .. Pokorny’s Wörterbuch lemma is per se expressive enough: kha khad (1959: 634). However, laryngeal obedience scholars are usually very reluctant to consider this kind of words worthy of entering into a comparison and then into linguistic reconstruction. Pronk (2019: 130), for example, refuses to take into consideration the root *atta- ‘father’, giving as a pretext that this «word is a nursery term that cannot be used for the reconstruction of PIE phonology». It seems that the children of the Indo–European supermen would have been the only human population on this planet that never used tender, cute or funny nursery terms.

Recapitulation and Concluding Remarks
Explicitly or implicitly, the following 20 arguments have been put forward so far:

1. Typological documentation shows that /a i u/ constitutes one of the most common phonemic vowel patterns.
2. The /a i u/ pattern is very natural as it is based on the optimal contrast —quantitatively minimum and qualitatively maximum— between the three cardinal points of articulation: dorsal, coronal and labial.
3. The lack of /a/ would contravene the linguistic universal that all languages have /a i u/.
4. Typological documentation shows that the trio /a i u/ can be and often is at the origin of more complex phonemic patterns, such as /a i u e o/, which seems to be the most common pattern in the world.
5. In fact, historically /a i u/ has generated richer vowel patterns, as, for example, in the Indo–Iranian continuum within the Indo–European sphere and also in the continua of other linguistic groups (Semitic, for example) or languages (Quechua, for example).
6. In the Indo–European case, for a process from /a i u/ to /a i u e o/ the influence of large contiguous Dravidian, Uralic and Turkic linguistic ensembles —all of which historically feature /a i u e o/— could be used as examples.
7. Typological documentation shows that /e/ and /o/ are generated in many languages from /a/ or also from /i/ and /u/ respectively.
8. The fact that the historical Indo–European /e/ and /o/ come mainly from /a/ would comfortably explain the relatively low presence of preserved Indo–European roots with /a/.
9. The vowel /a/ is the most frequent one. Its absence is extremely rare and therefore it should only be proposed using a set of solid arguments.
10. The typological fact that /e/ precedes /o/ in the development of the basic pattern /a i u/ is consistent with the Indo–European situation, where there are no languages with the pattern /a i u o/ and historically many Indo–European groups do not have /o/.

10 «que una misma tradición lingüística esté atestiguada primero como lengua a/o y luego como lengua /a/, no se da en ninguna ocasión».
11 «la reconstrucción del grado o como primitivo frente a a carece por completo de una justificación positiva». 

Indian J. Lang. Linguist., 4(4) (2023) 9-21 | 17
If /e/ or /o/ were older than /a/, we would expect to find patterns like /i u e/ or /i u o/ or /i u e o/ within historical Indo–European languages, but there are none.

The alleged merger of /e/ and /o/ into /a/ in Indo–Iranian—or in Luvian—is totally unnatural and lacks typological parallels.

The alleged merging of /e/ and /o/ into /a/ in Indo–Iranian would create a totally exceptional instance of regular, complete and systematic merger, which would have affected the entire lexicon of that language.

It would be very difficult to explain the reduction of /e/ and /o/ to /a/ in Indo–Iranian or in Luvian just by internal development or patrimonial evolution.

For the alleged change of Indo–European /e/ and /o/ to /a/, a contact phenomenon with the contiguous Burushaski, Dravidian, Munda, Turkic or Uralic languages, all of which regularly have /e/ and /o/ cannot be postulated.

In no way would the Aryan law of palatalization demonstrate the absence of /a/ in ancient Aryan languages nor the presence of /e/.

The three basic stages of regular evolution of /a i u/ > /a i u e/ > /a i u e o/ would be documented in the historical Indo–European languages, but not the reverse or opposite processes.

There are no known parallels to the putative reduction or regression of a more common pattern /a i u e o/ or similar to a pattern of the type /a i u/.

The clear presence of /a/ in the Old European hydronymy record—and /e/ or /o/ as its modern results—supports the existence of /a/ in a very early or at least previous Indo–European phase.

The presence of /a/ in clearly expressive, infantile or onomatopoeic forms of Indo–European languages points to the existence of /a/ in an early Indo–European phase too.

Therefore, we can conclude with Schmitt–Brandt (1973: 112): «It can thus be established that Indo-European had the three vowels *i, *a, *u in an early epoch of its development.» Perhaps the abandonment of the primitive vowel model, that of Franz Bopp and Georg Curtius, by traditional Indo–European Linguistics has been a colossal mistake; perhaps, then, it is time to return to the realistic simplicity of the model of the founders of the discipline, to an Indo–European Linguistics with a human face.

References


12 «Es kann somit festgestalten werden, daß das Indogermanische in einer frühen Epoche seiner Entwicklung über die drei Vokale *i, *a, *u, verfügte».

13 I am deeply indebted to colleague Peter Dunphy–Hetherington for reviewing the English of this paper and for his many suggesting comments.


Van Kemenade Ans (2002). Old and Middle English, the Germanic Languages, 110–141.


**Does this article screened for similarity?**

Yes

**Conflict of interest**

The Author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**About The License**

© The Author 2023. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.