Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian

Warries, A.R.

2 Relative chronology of Tocharian sound changes

To be able to properly compare the phonological systems of pre-Proto-Tocharian and pre-Proto-Samoyed, we need to have an intricate understanding of their developments through time. Otherwise, the comparison remains impressionistic. In this chapter, I will present the most important sound changes that occurred from Proto-Indo-European to Proto-Tocharian. There are still many points of disagreement among specialists, so where space allows, I will treat the differing interpretations of the data and argue which I find to be best supported. In a few cases, I will offer an alternative solution.

In chapter 3, I will treat the developments of Samoyed in the same way. For both languages, the aim is to obtain a relative chronology of phonological developments that is as secure as possible with the data from presently available etymologies and to reconstruct the developments of the phonological system through time. Only then can a comparison be attempted between pre-Proto-Tocharian and pre-Proto-Samoyed, which will be the topic of chapter 4. In the present chapter on Tocharian phonological changes, the possible indications of contact between Tocharian and Samoyed will not be considered, as the internal relative chronologies should each stand on their own.

2.1 Introduction

A large number of the developments from Proto-Indo-European to Tocharian are well-established and have been studied in considerable detail. Ringe (1996) gives a useful overview of most sound changes and places them in a relative chronology. However, not all the information in Ringe's book is still representative of the more recent research in the field, and the relative chronology there cannot be used as a starting point for our purposes without re-evaluation. Pinault (2008: 413–460) also discusses various aspects of Tocharian historical phonology, and many other sources contain treatments of the general picture (e.g., Malzahn 2010: 1–22), provide a dedicated overview (Hackstein 2017), or treat the developments of specific sounds or sequences (e.g., Hilmarsson 1993; Kim 1999; Pronk 2009), so that there are various sources of information to draw on. The purpose of this chapter on Tocharian sound changes is to synthesise an improved relative chronology of the Tocharian phonological system. In order to do this within the confines of this dissertation, those changes that are thought to have impacted the overall phonological system will be given a preferential treatment, while a few minor developments that are less informative and would unnecessarily overcomplicate matters are left aside.

The general Tocharian phonological system and the ways in which it differed from that of Proto-Indo-European will be considered in subsection 2.2. Then, various

phonological changes will be covered in more detail by topic, starting with general changes to the Proto-Indo-European consonant system in subsection 2.3, followed by a section dedicated to palatalization and its relationship to developments in both the consonants and the vowels in subsection 2.4. Subsequently, some late changes that took place in the vowel system are treated in subsection 2.5. A number of sound laws will have to be mentioned in the discussion of others, and therefore may need to be brought up earlier than they themselves appear with a dedicated discussion; cross-references will be provided where appropriate. At the end of each subsection, there is a brief summary of the changes discussed. The numbers that accompany the cited etyma in this chapter correspond to etymological reference list 1, which contains more extensive etymological information and references. It can be found at the end of this dissertation.

2.2 Tocharian and Proto-Indo-European phonology

The phoneme inventories of Tocharian A and B are very similar, especially as regards the consonants. The differences between Tocharian and Proto-Indo-European, on the other hand, are significant. The reconstructed Proto-Indo-European consonant inventory is presented here in Table 2.1.

Table 2.1: The Proto-Indo-European consonant system.

PIE	labial	dental	palatal	palatovelar	velar	labiovelar	"laryngeal"
stop	*p (b) bʰ*	*t d dʰ*		*ḱ ǵ ǵʰ*	*k g gʰ*	*kʷ gʷ gʷʰ*	
							h₁ h₂ h₃
sibilant		*s*					
nasal	*m*	*n*					
liquid		*l r*					
glide			*i̯*			*u̯*	

Proto-Indo-European had a large number of obstruent phonemes, with a three-way opposition in the stops between, e.g., **t* vs. **d* vs. **dʰ*. In the labial series, the evidence for PIE **b* is extremely limited, so that its proper existence as a member of the PIE consonant system is doubtful (see, e.g., Olander 2020, with literature).

A single sibilant **s* can be reconstructed, as well as four resonants **m*, **n*, **l*, **r*, and the glides **u̯* and **i̯*. There are also the three laryngeals **h₁ h₂ h₃*, whose original phonetic realization is a matter of debate (e.g., Kümmel 2012; Kloekhorst 2018). They often behave alike in their development in the historical phonology of the various Indo-European languages, including Tocharian. Taken as a class of consonants, the laryngeals can collectively be written with a capital **H* in those cases where the comparative material

does not allow us to determine which of the three was present, or where differentiation is not necessary for the issue at hand.

The traditional notation of stops as voiceless, voiced and voiced aspirate is adhered to throughout this dissertation in reconstructed forms, without taking a definite stance in the longstanding discussions about the phonetic interpretation of these series. A main point of contention is the interpretation of the traditionally simple voiced series with **d*, **ǵ*, **g*, **gʷ* as glottalized **ʔd*, **ʔǵ*, **ʔg*, **ʔgʷ* instead of plain voiced (Gamkrelidze & Ivanov 1972, Hopper 1973). This interpretation would make the almost complete absence of PIE **b* in the consonant system more understandable: the absence of a glottalized bilabial in a glottalized series is cross-linguistically more common than the absence of a voiced bilabial in a voiced series (ibid., and, e.g., Kortlandt 1985). The idea that the traditional series of voiced stops was originally glottalized is known as the Glottalic Theory.[1]

Additionally, combining the Glottalic Theory with evidence from Anatolian, Kloekhorst has proposed an interpretation of the original stop system as consisting only of voiceless stops, with geminates **pː*, **tː*, **ḱː*, **kː*, **kːʷ* instead of the traditional series **p*, **t*, **ḱ*, **k*, **kʷ* and glottalized voiceless stops **ʔt*, **ʔḱ*, **ʔk*, **ʔkʷ* instead of the traditional **d*, **ǵ*, **g*, **gʷ* (Kloekhorst 2016). Due to the uncertain position of Tocharian within Indo-European, the potential relevance of this alternative reconstruction for Tocharian is still unclear, as will be discussed in 2.3.4.4. The three competing interpretations are summarized in Table 2.2 for reference.

Table 2.2: Three alternative reconstructions of the Proto-Indo-European (or Proto-Indo-Anatolian) stop system.

	labial	dental	palatovelar	velar	labiovelar
Traditional	*p (b) bʰ*	*t d dʰ*	*ḱ ǵ ǵʰ*	*k g gʰ*	*kʷ gʷ gʷʰ*
Glottalic	*p (ʔb) b*	*t ʔd d*	*ḱ ʔǵ ǵ*	*k ʔg g*	*kʷ ʔgʷ gʷ*
Kloekhorst	*pː (ʔp) p*	*tː ʔt t*	*ḱː ʔḱ ḱ*	*kː ʔk k*	*kːʷ ʔkʷ kʷ*

The consonant system of Tocharian A and B is in any case rather different from that of Proto-Indo-European, as can been seen below in Table 2.3. Additional reconstructed Proto-Tocharian phonemes are marked with an asterisk. The velar nasal *ṅ* is dependent on a following velar stop.

[1] If the traditionally plain voiced series was characterized by a glottalic pronunciation, the aspiration (or breathy voice) traditionally assumed for the series **bʰ*, **dʰ*, **ǵʰ*, **gʰ*, **gʷʰ* would become redundant (see, e.g., Kümmel 2012 for an overview with further references).

Table 2.3: The Proto-Tocharian, Tocharian A and Tocharian B consonant inventories. Consonants exclusive to Proto-Tocharian marked with an asterisk.

TAB~PT	labial	dental	retroflex	palatal	velar	labiovelar
stop	*p* **p*^{y}	*t*			*k*	*k*^{w}
affricate		*ts*		*c*		
sibilant		*s*	*ṣ*	*ś*		
nasal	*m* **m*^{y}	*n*		*ñ*	(*ṅ* =[ŋ])	
liquid		*l r*		*ly*		
glide	*w* **w*^{y}			*y*		

Tocharian phonology does not make distinctions based on such features as voicing (**t* vs. **d*) or aspiration (**d* vs. **d*^{h} or **t* vs. **t*^{h}) at all, and no traces of potential glottalization (**d* vs. *ˀ*d* or **t* vs. *ˀ*t*) are found either.[2] Instead, Tocharian extensively employs palatalization in a great number of oppositions, so that almost every consonant has a palatal or palatalized counterpart. The only exceptions are *r*, *y* and the labials *p*, *m* and *w*; the palatalized counterpart to both *k* and *k*^{w} is the *ś*. The labials do need to be reconstructed with palatalized counterparts **p*^{y}, **m*^{y} and **w*^{y} for Proto-Tocharian on account of some special developments. For example, the correspondence between TA *w* and TB *y* in certain words like TA *want*, TB *yente* 'wind' are understood as distinct developments of a palatalized labial glide **w*^{y} to TA *w* and TB *y*; thus, PT **w*^{y}*ente* 'wind' (e.g., Hackstein 2017: 1326).

The vowel systems of Proto-Indo-European and Tocharian are also different, as shown in Table 2.4. The great changes that took place cannot be easily seen from looking at a general overview of the systems, and most vowels will need to be discussed individually.

Table 2.4: Proto-Indo-European and Proto-Tocharian vowel inventories.

	front	back		front	central	back
high	*i*	*u*	PIE > PT	*i*	*ə*	*u*
mid	*e ē*	*o ō*		*e*		*o*
low					*a*	*å*

General innovative features of the Tocharian vowel system include the reduced central vowel /ə/ and a lack of length distinctions (i.e., no **ē* or **ō*). The length mark in Tocharian orthography marks either a qualitative opposition (Tocharian A), or the placement of

[2] Geminate stops in Tocharian do not correspond to PIE **t*, and are thus independent from Kloekhorst's reconstruction of this phoneme as an original geminate **t*:.

the accent (Tocharian B). At some point in Tocharian prehistory, front vowels palatalized preceding consonants, creating the pervasive palatalization opposition in the Tocharian consonant system. In the process, some front and back vowels, such as PIE *$\bar{e}$ and *o, merged phonologically (see 2.4).

In Proto-Indo-European, the high vowels *i and *u were the vocalic allophones of the consonants *$i̯$ and *$u̯$. The diacritic "$\underset{\wedge}{\ }$" indicates non-syllabicity of phonemes that may otherwise occur as syllabic nuclei. The opposite is marked with the diacritic "$\underset{\circ}{\ }$", as in the syllabic resonants *$m̥$, *$n̥$, *$l̥$ and *$r̥$, and the syllabic laryngeals *$h_1̥$, *$h_2̥$, *$h_3̥$, which could all function as syllabic nuclei as well as more canonical consonants. While syllabification was non-phonemic and automatic in Proto-Indo-European itself, making these diacritics strictly speaking anachronistic (e.g., Beekes 2011: 120), they are a useful tool for elucidating the developments in the later languages and will be extensively employed for this purpose throughout the present chapter.

This brief overview illustrates that Tocharian phonology underwent considerable modifications from its Proto-Indo-European ancestor. A number of developments are also unusual within Indo-European, such as the complete loss of distinctive manners of articulation in the stop system, which is not paralleled in any other branch of Indo-European. Some of the changes in the vowel system are also striking, like the shift from PIE *o to PT *e and the complete loss of distinctive vowel length. In the rest of this chapter, the general picture sketched here will become more detailed, and a number of important remaining problems will be discussed.

2.3 Early consonant developments

The developments of the Tocharian consonants are unique within Indo-European, and they provide an important point of comparison with Uralic/Samoyed. It is therefore important to gain a proper understanding of the various changes and the many difficulties involved. Overviews of the developments of laryngeals (subsection 2.3.1) and syllabic resonants (subsection 2.3.2) will be given first in this account of early pre-Proto-Tocharian phonology. Both led to the creation of new vowel phonemes and belong to some of the first changes in the pre-Proto-Tocharian phonological system. The developments of palatovelars and labiovelars constitute an interesting aspect of Tocharian in relation to its geographical position in the east, where satemization and its associated changes are more commonly observed in other Indo-European branches (see subsection 2.3.3). Possible interactions between labiovelars and various vowel phonemes could furthermore provide insight into the relative chronology of Tocharian sound changes. The discussion of the developments of the stop system as a whole (subsection 2.3.4) addresses some of the most characteristic and simultaneously most puzzling aspects of Tocharian historical phonology, chief among them the development of PIE *d.

2.3.1 Laryngeals

The three laryngeals *h_1, *h_2 and *h_3 constitute a famous part of the Proto-Indo-European consonant inventory. Some Anatolian languages, such as Hittite, preserve them in certain phonological positions, but in most other branches, the Proto-Indo-European laryngeals were eventually lost as distinct consonants. In Tocharian, the general tendency is for any laryngeal *H that stood between two consonants to become PT *a. This can be schematically represented as a development from *$CH̥C$ to *CaC (Ringe 1996: 22–25, 32–37, Hackstein 2017: 1316–1318); cf. the development of PIE *$ph_2tēr$ to PT *$pat^{y}er$ 'father' (TB *pācer* /pácer/, TA *pācar*). In a similar fashion, *$Cih_{2/3}C$ became *$CyaC$. With *Cih_1C the result is generally assumed to have been *$CīC$ instead (Ringe 1996: 25; Hackstein 2017: 1314), but there are not many examples to confirm this development. In word-final position, at least *Cih_2 also became *Cya, but the developments of *Cih_1 and *Cih_3 in this position are uncertain.

When a laryngeal occurred next to a PIE vowel *e or *o, it is normally assumed that a number of fairly standard changes occurred as also seen in other branches of Indo-European (cf., e.g., Ringe 1996: 7–9; Beekes 2011: 147), at least as an initial step. Further changes to these vowels are given in more detail below in other (sub)sections. PIE *e could be coloured by *h_2 to become *a, either simple *a from the sequence *h_2e, or long *$ā$ from the reverse sequence *eh_2. On the other hand, *h_1 only caused lengthening in *eh_1, but had no colouring effects, thus yielding *$ē$. The vowel PIE *o was not coloured by any laryngeal.

The developments of *h_3e and *eh_3 are uncertain. There are only very few examples with these sequences reflected in Tocharian, and they are contradictory and depend on different interpretations. Based on the effects of *h_1 and *h_2 as well as parallel developments in other Indo-European languages, it is normally assumed that *h_3e became *o and that *eh_3 became *$ō$ (e.g., Ringe 1996: 7–9, Hackstein 2017: 1313–1314); cf. subsections 2.4.5.1 and 2.4.5.2. Since instances of apparent *h_3e to *o could in principle reflect *h_3o with an *o*-grade, the involvement of the laryngeal in colouring the vowel is uncertain, and alternative interpretations are possible. Thus, while the developments of these sequences as presented here are often assumed, some of them remain uncertain.[3]

- PIE *Ch_1C > *a (e.g., 78. PIE *$-mh_1no-$ > *$-mano-$, the mid.prs.ptc. suffix)

[3] Kim has proposed some different developments of sequences with a final laryngeal specifically at the end of a word: there, the laryngeal might be lost under certain conditions. The data is difficult to unambiguously interpret, and both word-final position and the number of syllables may play a role (Kim 2018: 98–112). Kim's conclusions, if accurate, would affect the precise development of particular words and suffixes, but they have no bearing on the further general vowel developments of Tocharian, and will be left out of considerations here.

- PIE *Ch_2C > *a (e.g., 91. PIE *$p\underset{\circ}{h}_2t\bar{e}r$ > *$patēr$ ‘father’)
- PIE *Ch_3C > *a (e.g., 23. PIE *$ǵnh_3$- > *$ǵna$- ‘know’)

- ? PIE *Cih_1 > *$\bar{i}$ (e.g., 13. PIE *$d\underset{\circ}{u}idḱ\underset{\circ}{n}t$ → *$d\underset{\circ}{u}ih_1ḱ\underset{\circ}{n}t$ > *$wīkəmt$ ‘twenty’)[4]
- PIE *Cih_2 > *ya (e.g., 18. PIE *$d^hi\underset{\circ}{h}_2g^w$- > *$d^hyag^w$- ‘pierce, bite’)
- PIE *Cih_3 > *ya (e.g., 28. PIE *$g^wi\underset{\circ}{h}_3u$- > *$g^wyaw$- ‘live’)
- PIE *Cuh_2 > *wa (e.g., 106. PIE *$s\underset{\circ}{u}h_2dro$- > *$swadro$- ‘sweet’)[5]

- PIE *$\#h_1e$ > *e (e.g., 35. PIE *$h_1ek\underset{\circ}{u}o$- > *$ekwo$- ‘horse’)
- PIE *$\#h_2e$ > *a (e.g., 45. PIE *$h_2eǵ$- > *ag- ‘drive’)
- ? PIE *$\#h_3e$ > *a (e.g., 54. PIE ?*$h_3em(e)so$- > *$amso$- ‘shoulder’)
- ? PIE *$\#h_3e$ > *o (e.g., 56. PIE ?*h_3eros- > *$oros$- ‘appearance’)
- PIE *$\#Ho$ > *o (e.g., 58. PIE *h_3ok^ws- > *ok^ws- ‘eye’)
- PIE *$\#h_1i$ > *i (e.g., 37. PIE *h_1itr- → ?*$h_1itōr$ > *$itōr$ ‘path’)

- PIE *eh_1 > *$\bar{e}$ (e.g., 77. PIE *meh_1n-s- > *$mēn$-$\bar{e}(n)$- ‘moon’)
- PIE *eh_2 > *$\bar{a}$ (e.g., 2. PIE *$b^hreh_2tēr$ > *$b^hrātēr$ ‘brother’)
- ? PIE *eh_3 > ??? (no reliable examples, see subsection 2.4.5.1)
- PIE *oh_1 > *$\bar{o}$ (e.g., 15. PIE *$duoh_1$ > $dwō$ ‘two (2)’)
- ? PIE *oh_2 > ??? (no reliable examples, see subsection 2.4.5.1)
- ? PIE *oh_3 > ??? (no known examples)

The development and effects of especially *h_3 are difficult to establish due to a general lack of reliable examples. However, we can tentatively summarize the following trends in the developments of the three laryngeals:

- *Ch_1C = *Ch_2C = *Ch_3C, namely > *CaC
- *Cih_1C ≠ *Cih_2C = *Cih_3C, namely > $CīC$ and $CyaC$ respectively
- *h_1e ≠ *h_2e, namely > *e- and *a- respectively; *h_3e is uncertain
- *eh_1 ≠ *eh_2, namely > *$\bar{e}$ and *$\bar{a}$ respectively; *eh_3 is uncertain

[4] This change has very few examples, but this seems to be the best way to get the TAB *i* in this word (Ringe 1996: 25; Hackstein 2017: 1314). A possible parallel is the optative suffix PT *-*i*- from PIE *-ih_1- potentially via *$\bar{i}$ (Ringe 1996: 25; Hackstein 2017: 1314), and perhaps TA *wir* ‘young, juvenile’ if correctly etymologized and reconstructed as from PIE *uih_1ro- (Hackstein 2017: 1314). The problem with the word for ‘twenty’ itself is that it did not have a real *h_1 originally. On the other hand, there is no parallel for a development from PIE *id to pre-PT *$\bar{i}$ either.

[5] I do not know any reliable examples with PIE *uh_1 or *uh_3. We might expect the former to become *$\bar{u}$ if the development from *ih_1 to *$\bar{i}$ is correct, and the latter *wa on the basis of comparison with *$ih_{2/3}$ and *uh_2, but this is by no means assured.

We see one point of general agreement between all three: between two consonants, the laryngeals behave the same way and become **a*. In other environments there is mostly disagreement. After **i*, only **h_2* and **h_3* certainly became **a*. **h_1* rather seems to have caused lengthening, although this is based on few, partially uncertain examples. For the sequences of the type **HeH* and **eHH*, only clear information on **h_1* and **h_2* is reliable, and they yield different results: **h_1* gives an *e*-vowel, and **h_2* causes a change to a long *a*-vowel.

Laryngeals in other positions were often simply lost, for instance in most cases in word-initial position (in **$h_1$rud^{h}ro-*, **$h_2$ster-*, **$h_2$ueh$_1$nto-* below), in several consonant clusters with resonants (in **plh$_1$no-*, **prHwo-*, **solh$_2$uo-* below), and between vowels (in **tuHom* below). In intervocalic position the resulting hiatus could be filled up with a glide. Only PIE **$h_2$ntb^{h}oh$_1$* 'both' shows vocalization of a word-initial laryngeal as **a* (Ringe 1966: 15).

44. PIE **$h_1$rud^{h}ro-* > early pre-PT **rud^{h}ro-* 'red'
52. PIE **$h_2$ster-* > early pre-PT **ster-* 'star'
53. PIE **$h_2$ueh$_1$nto-* > early pre-PT **wēnto-* 'wind'
92. PIE **pl̥h$_1$no-* > early pre-PT **pəlno-* 'full'
93. PIE **pr̥Hu̯o-* > early pre-PT **pərwo-* 'first'
102. PIE **solh$_2$u̯o-* > early pre-PT **solwo-* 'whole'
113. PIE **tuHom* > early pre-PT **tuwom* 'you'

50. PIE **h$_2$ntb^{h}oh$_1$* > **antb^{h}ō* 'both' (or perhaps **$h_2$entb^{h}oh$_1$* > *antb^{h}ō*)

While the laryngeals were eventually lost in Tocharian, they had not yet disappeared in all positions when the syllabic resonants vocalized. This can be seen in the development of PIE **u̯l̥h$_2$ōnts*, the ancestor of PT **wəlo* 'king', which became post-PIE **wəlōnts* rather than ***wlōnts* (see Lubotsky 1994 on the reconstruction of this word). The latter pre-form would have resulted from ***u̯lōnts* with loss of the laryngeal before vocalization of **l̥* as **əl*, and this should have yielded PT ***lo* rather than **wəlo*; cf. the development of **u̯l̥h$_2$ntos* > **wlantos* > PT **lante*, the genitive of the same word. The syllabic resonants are the topic of the next subsection.

The development of PIE **eh$_1$* and **oh$_1$* seems to be identical to that of the corresponding long vowels **ē* and **ō*, so that an early merger of **eh$_1$* with **ē* and of **oh$_1$* with **ō* can be assumed, before any of the subsequent vowel changes. The elimination of the laryngeals as separate segments in other environments is difficult to date. The PIE sequence **eh$_2$* is the only source of pre-PT **ā*, so that no comparison with primary "**ā*" is possible. The situation with pre-PT **a* from PIE **h$_2$e* and from vocalized laryngeals is similar. However, after colouring (**eh$_2$* to **ah$_2$*, **h$_2$e* to **h$_2$a*, etc.), the three laryngeals at least cannot be phonologically differentiated anymore, so that an interpretation with a

single laryngeal phoneme **H* becomes possible. The laryngeal disappeared in cases where it lengthened the (coloured or uncoloured) preceding vowel before any other vowel changes had taken place. In principle, a phoneme **H* could have persisted in other environments for an indeterminate time.

2.3.2 Syllabic resonants

In Proto-Indo-European, the resonants **m*, **n*, **l* and **r* (collectively, **R*) could act both as consonants and as vowels. That is to say, they could function as the onset or coda of a syllable, but also as the nucleus. When they function as the nucleus of a syllable, they are known as syllabic resonants. Syllabic resonants, collectively written as **R̥*, stopped being syllabic at some point in pre-Proto-Tocharian. The general pattern is that a short vowel **ə* was inserted before the resonant (Ringe 1996: 67–69; Hackstein 2017: 1319–1320).[6] This means that syllabic resonants constitute one of the sources for the important innovative Proto-Tocharian vowel **ə*. The phonetic interpretation of this vowel as it developed initially will be further discussed in 2.4.4. The list below contains some representative examples of this development.

3. PIE **bʰr̥ǵʰros* > early pre-PT **bərgros* 'tall'
4. PIE **dekm̥* > early pre-PT **dekəm* 'ten (10)'
9. PIE **dn̥ǵʰueh₂* > early pre-PT **dəngʰwā*/*gʰəndwā* [metathesis] 'tongue'
13. PIE **duih₁km̥t* > early pre-PT **wīkəmt* 'twenty (20)'
60. PIE **ḱm̥tom* > early pre-PT **kəmtom* 'hundred (100)'
92. PIE **pl̥h₁nos* > early pre-PT **pəlnos* 'full'
93. PIE **pr̥Huos* > early pre-PT **pərwos* 'first'
98. PIE **septm̥* > early pre-PT **septəm* 'seven (7)'
118. PIE **u̯l̥h₂ōnts* > early pre-PT **wəlōnts* 'king'

In word-initial position, the development was different: instead of **ə*, the basic prop vowel was (pre-)PT **e* (Ringe 1996: 99–100; Hackstein 2017: 1319–1320). This (pre-)PT **e* was different from PIE **e*, which became PT **yə-*. Instead, it merged with PIE **o*, which also became PT **e*. It is difficult to determine whether this means that word-initial

[6] It has been suggested that the syllabic resonant **R̥* at first vocalized as **uR*, which later became PT **əR* with the general change from **u* to **ə* (e.g., Burlak 2000: 124). However, examples like PIE **ḱm̥tom* to PT **kənte* 'hundred (100)' show that this is impossible at least for **m̥*, since an intermediate ***kumtom* should have yielded PT ***kʷənte* with an initial labiovelar; this is not the case (cf. subsection 2.4.3). Since the difference between **u* and **ə* can only be seen in this type of context, next to a velar, the other examples do not provide any indication either way. See also Ringe (1991) for a discussion of this issue.

syllabic resonants were first vocalized as *o, and alternative interpretations are possible. The examples below show these developments.

- PIE *$\underset{\circ}{n}C$ > PT *e(n)- 'negative prefix' (also *o(n)- and *a(n)- by umlaut)
- PIE *$h_1\underset{\circ}{n}C$ > PT *e(n)- 'in; adverbializing prefix' (also *o(n)- and *a(n)- due to umlaut); next to full grade *h_1enC > PT *yə(n)- 'id.'

41. PIE *$h_1\underset{\circ}{n}\acute{k}$- > PT *enk- > TB eṅk-, TA eṃts- 'seize, grip'
43. PIE *h_3rg^w- > PT der. *erkent > TB erkent, TA arkant 'black, dark'
79. PIE *$\underset{\circ}{n}\acute{k}\underset{\breve{}}{u}o$- 'mortal' > PT *enk^{w}e > TB eṅkwe, TA oṅk 'man'

The negative prefix *$\underset{\circ}{n}$- and the local prefix *$h_1\underset{\circ}{n}$- 'in' are the most common sources for word-initial syllabic resonants. The developments of PIE *$\underset{\circ}{n}\acute{k}\underset{\breve{}}{u}o$- and *$h_1\underset{\circ}{n}\acute{k}$- suggest that the presence or absence of an initial *h_1 had no influence on the vocalization of the nasal. This contrasts with the development of *$\underset{\circ}{h_2}ntb^hoh_1$ 'both' with a vocalized *h_2 in the previous subsection, unless this reflects a less likely full-grade form *$h_2entb^hoh_1$ instead. The only instances of a non-nasal "initial" syllabic resonant are found in reflexes of the root *h_3rg^w- 'black, dark', which also vocalized with a vowel *e as *er-, similar to the *$(h_1)\underset{\circ}{n}$ in *$\underset{\circ}{n}\acute{k}\underset{\breve{}}{u}o$- 'mortal → man' and *$h_1\underset{\circ}{n}\acute{k}$- 'seize, grip'.

Ringe points out that it is difficult to be sure when the prop vowel *ə became a proper phonemic vowel in its own right, since a sequence *[əR] could originally simply have been the automatic phonetic representation of */$\underset{\circ}{R}$/. However, when some final nasals were lost, *ə from word-final *$\underset{\circ}{n}$ and *$\underset{\circ}{m}$ would have found itself alone and fully phonemic (Ringe 1996: 67). It is impossible to tell whether initial syllabic resonants immediately yielded a separate vowel *ë (non-front, non-rounded), but this may have remained non-phonemic until *o became *ë as well, as it would have been in allophonic distribution with *ə until then.

However, the phonemic vocalization of the syllabic resonants, and thus the appearance of *ə in the vowel system, can also be connected with another development. In PT *wəlo 'king' from PIE *$\underset{\breve{}}{u}\underset{\circ}{l}h_2\bar{o}nts$, we see that a syllabic resonant could occur in front of regular vowels where a laryngeal was lost; this is not the usual situation. Therefore, instead of positing a stage like hypothetical *$w\underset{\circ}{l}\bar{o}nts$, with a vocalic resonant directly followed by a vowel, it is preferable to think of this form as *$wəl\bar{o}nts$ with the new vowel *ə in the first syllable as soon as the laryngeal disappeared in that position. It remains difficult to accurately determine when exactly consonantal laryngeals disappeared, however, even if all of their most salient interactions with the vowels took place at an early stage (see 2.3.1).

2.3.3 Palatovelars and labiovelars

There are three series of velars commonly reconstructed for Proto-Indo-European: palatovelars **ḱ*, **ǵ*, **ǵʰ*, plain velars **k*, **g*, **gʰ* and labiovelars **kʷ*, **gʷ*, **gʷʰ*. The palatovelars developed into palatal affricates in especially a number of eastern branches (e.g., in Indo-Iranian and Balto-Slavic), whereas they merged with the plain velars in other, mostly western branches (e.g., in Italic and Germanic). Languages belonging to the latter group are known conventionally as *centum*-languages, based on Latin *centum* /kentum/ from PIE **(d)ḱmtom* 'hundred'. The first group contains the so-called *satem*-languages, based on Avestan *satəm* from the same PIE **(d)ḱmtom* 'hundred'.

In Tocharian the development of the palatovelars is straightforward, but the labiovelars are disputed. They are preserved in some cases and lost in others, without an uncontroversial set of conditioning factors. Proposed interpretations involve the surrounding vowels, and thus potentially influence our understanding of the relative chronology in both the stop system and the vowel system. In this subsection, the early Tocharian developments of the palatovelars will be summarized and the complicated data of the labiovelars will receive a more detailed treatment.

2.3.3.1 Developments of PIE palatovelars

The Tocharian words for 'hundred' are TA *känt* and TB *kante*, going back to PT **kənte* from PIE **(d)ḱmtom* (Krause & Thomas 1960: 64-65). This is just one of the examples that show that Tocharian depalatalized its palatovelars, and is to be counted among the *centum*-languages, unlike its geographically closest Indo-European neighbours of the Indo-Iranian branch. The status of Tocharian as a *centum*-language is further apparent from words like PIE **ḱu̯ō* > PT **ku* > TAB *ku* 'dog'; cf. Skt. *śvā́* 'id.', PIE **h₁eḱu̯os* 'horse' > PT **yəkʷe* > TB *yakwe*; TA *yuk* 'horse'; cf. Skt. *áśva-* 'id.', PIE **bʰr̥ǵʰro-* > PT **pərkre* 'tall, high'; cf. Skt. *bṛhánt-* (Ringe 1996: 39–41 cites many more examples).

The depalatalization of the palatovelars and their merger with plain velars in Tocharian may have occurred very early, since they do not show any palatal effect, nor any other distinctive behaviour (Ringe 1996: 39–42). Unless we can make a phonemic distinction between palatovelars as in PIE **ǵombʰo-* 'tooth' and secondarily palatalized velars, **ǵombʰo-* should probably have depalatalized to **gombʰo-* by the time that PIE **gʷih₃u-* 'live' was palatalized via **gyaw-* to PT **kʸaw-*.

2.3.3.2 Developments of PIE labiovelars

While the development of palatovelars to plain velars is entirely straightforward, the Tocharian development of the Proto-Indo-European labiovelars is quite complicated, as is their representation in the Tocharian script (see, e.g., Hilmarsson 1993; Kim 1999; Pinault 2008: 445–447; Fellner 2005). The majority of original Proto-Indo-European

labiovelars lost their labialization in Tocharian, but a few did remain. The proposed conditions for the delabialization of labiovelars intersect with the developments of PIE **o*, **e*, **i* and **u* and must thus be discussed in order to gain a more complete understanding of the chronology of some developments in the vowel system as well. The followings list contains the words where a labiovelar was preserved and attested in either Tocharian language.[7]

29. PIE **gʷm̥-sk-* 'come' > PT **kʷəmnəsk-* > TB *kənməsk-*[8], TA *kumnäs-* 'come'
64. PIE **kʷekʷlo-* 'wheel; circle' > PT **kʷəkʷle* > TB *kokale*, TA *kukäl* 'wagon'
67. PIE **kʷi-so* 'who' > PT **kʷəse* > TB *kᵤse*, TA *kus* 'who'
68. PIE **kʷr̥i̯h₂-* 'buy' → **kʷr̥i̯h₂-wor* > **kʷəryawer* > PT **kʷəryor* > TB *karyor*, TA *kuryar* 'trade, commerce' (the labiovelar must have been delabialized analogically in TB)
88. PIE **pekʷ-* 'ripen' → *pekʷ-l̥* > PT **pʸəkʷəl*, pl. *pʸəkʷəla* > TB *pikul*, pl. *pikwala*, TA *pikᵤl*, pl. *puklā* 'year'
101. PIE **sokʷo-* 'sap'> **sokʷo-* > PT **sekʷe* > TB *sekwe*, TA *saku* 'pus'
125. PIE **u̯l̥kʷo-* 'wolf' > **wəlkʷo-* > PT **wəlkʷe* > TB *walkwe* 'wolf (?)'

These examples display a variety of phonological contexts in which a labiovelar **Kʷ* was apparently retained: **Kʷm̥*, **Kʷl*, **Kʷi*, **Kʷr̥*, **Kʷl̥*, and **Kʷo*. The initial labiovelar of PIE **kʷekʷlo-* 'wheel' must have been restored, however, since other evidence shows that the sequence **Kʷe* merged with **Ke* at some point: both ended up as palatalized TAB *śə*. The merger between these sequences is clear from the three examples below, but the exact mechanism and intermediary steps remains uncertain.

27. PIE **gʷenh₂* > PT **kʸəna* > TB *śana* /śə́na/, TA *śäṃ* 'wife'
45. PIE **h₂eǵ-e-* > **age-* > PT **akʸə-tər* > TAB *āśtär* 'lead' prs.mid.3sg.

---

[7] Ringe and Kim also list TA *kᵤñaś* 'quarrel, strife', connected with the root **gʷʰn-* 'hit' (Ringe 1996: 107; Kim 1999: 150). Dragoni has instead proposed this to be borrowing from Khotanese *gūrāś-* 'to quarrel' or the infinitive *gūrāśä*, although the shift from **r* to **ñ* is unclear (Dragoni 2023: 92–93). It is better to disregard this word in the current discussion.

[8] Delabialization in the sequence PIE **gʷm-* is posited by Pinault to account for the non-labialized reflex of 'come' in TB *kəmnəsk-*; this would have been regular in the forms such as the PIE **gʷmont*, which yielded prt.3pl. TB *kameṃ* /kə́men/ via **kmen* through a secondary restoration of the root syllable with an additional **ə*. The resulting root shape **kəm-* could have been generalized to the derivative *kəmnəsk-* in the prehistory of Tocharian B specifically; Tocharian A *kumnäs-* retains the regular labiovelar reflex of PIE **gʷm̥-sk-*, and no reflex corresponding to the Tocharian B preterite is known (Pinault 2008: 446). In any case, TA *kumnäs-* shows that the labiovelar was retained in this word.

65. PIE **kʷetuores* > PT **kʸətwerə* > TB *śtwer*, TA *śtwar* 'four (4)'

The first labiovelar of PIE **kʷekʷlo-* should thus regularly have been palatalized, yielding something like TB ***śk(w)ale*, TA ***śukäl*. Since this did not happen, the **e* must have been affected by the flanking labiovelars. An intermediate stage **kukʷlos* could be posited, similar to Greek κύκλος 'circle, ring, wheel' (Hilmarsson 1986: 61), or a secondary re-labialization due to assimilation of from expected **kekʷlos*.[9] Since the Proto-Tocharian reconstruction required for the vowel development in TB *kokale* includes two labiovelars, PT **kʷəkʷle*, it follows that a sequence PIE **kʷl* was preserved. Many other examples show a loss of labiovelars in additional contexts, without palatalization to **ś*, as listed below.

16. PIE **dʰegʷʰ-* > PT **tˢək-* > TB *tsək-*, TA *tsäk-* 'burn'
18. PIE **dʰi̯ə̯h₂gʷ-* > PT **tˢak+a-* > TB *tsaka-*, TA *tsākā-* 'pierce'
26. PIE **gʷeh₃u-* > **kow* > PT **kew* > TB *keᵤ*, TA *ko* 'cow'
30. PIE **gʷreh₂-* → **gʷrə̯h₂-mn̥-* > PT **kramər* > TB *krāmär* 'weight, heaviness', der. *kramartstse* 'heavy', TA der. *krāmarts* 'heavy'
31. PIE **gʷr̥h₂u̯o-* > PT **kərwenʸe* > TB *kärweñe*, TA *kärwaṃs-* 'stone'
32. PIE **gʷʰen-* 'hit' → **gʷʰn̥-sk-* > PT **kəska-* > TB *kəska-* 'scatter, strike apart'
33. PIE **hₓeh₃gʷʰ-* > PT **yok-* > TB *yok-*, TA *yok-* 'drink'
38. PIE **hₓleng*ʷʰ- > PT **lənk-* > TB *lənk-*, TA *länk-* 'hang, dangle'
43. PIE **hₓregʷ-* → **hₓr̥gʷ-ont-* > **erkont-* > PT **erkent-* > TB *erkeṃt*, TA nom.pl. *arkaś* 'black'
58. PIE **h₃okʷ-s-* > **ok(s)* > PT **ek* > TB *ek*, TA *ak* 'eye'
65. PIE **kʷel-* → **kʷolo-* > **kolo* > PT **kele* > TB *kele* 'navel'
68. PIE **kʷo-* → **kʷo-tos* > **koto* > PT **kete* > TB *kete* 'whose'
70. PIE **kʷrinh₂-* > **krəna-* > PT **kərna-* > TB *kərna-* 'buy' (present stem)
82. PIE **nokʷtu-* > PT **noktə* > TA obl. *nokte* 'at night', adv. *noktiṃ* 'last night'
83. PIE **nokʷteu̯i̯o* > PT **nektʸəwʸe* > TB *nekciye*, TA *nakcu* 'at night, last night'
87. PIE **pekʷ-* > PT **pək-* > TB *pak-*, TA *päk-* 'cook, ripen'
90. PIE **penkʷto-* > PT **pʸənkte* > TB *piṅkte*, TA *pänt* 'fifth'
108. PIE **tekʷ-* 'run, flow' → **tekʷos-* > **tekos* > PT **tʸəke* > TB *cake* 'river'
127. PIE **u̯ōkʷs*, acc.sg. *u̯okʷm̥* > **wok* > PT **wek* > TB *wek*, TA *wak* 'voice'

[9] The labiovelars in this word were only finally lost in Tocharian B *kokale* in the process of a minor sound change converting PT **ə* to TB *o* between two labial consonants (see subsection 2.5.6). TA *kukäl* reflects the original PT **ə* affected differently by at least one preserved labiovelar in that language.

The precise conditioning of this loss of labialization has been interpreted in various ways, and clearly no single rule can cover all of the material. For a number of words, it has been assumed that a delabialization took place next to PIE **o*. This is interpreted as a change from **K^{w}o* to **ko* that took place before PIE **o* became unrounded to **ë* and eventual PT **e* in the relative chronology of sound changes (Ringe 1996: 42; Kim 1999: 149–150). This development would not only have taken place word-initially, but also word-internally (Ringe 1996: 42; Kim 1999: 159-164), and, according to Hackstein, between *o* and a following consonant (Hackstein 2017: 1325).[10] The relevant examples are as follows:

26. PIE **g^{w}eh$_{3}$u-* > **kow* > PT **kew* > TB *ke$_{u}$*, TA *ko* 'cow'
43. PIE **h$_{x}$rg^{w}-* → **h$_{x}$rg^{w}-ont-* > **erkont-* > PT **erkent-* > TB *erkeṃt*, TA nom.pl. *arkaś* 'black'
65. PIE **k^{w}el-* → **k^{w}olo-* > **kolo* > PT **kele* > TB *kele* 'navel'
68. PIE **k^{w}o-* → **k^{w}o-tos* > **koto* > PT **kete* > TB *kete* 'whose'
108. PIE **tek^{w}-* 'run, flow' → **tek^{w}os-* > **tekos* > PT **t^{y}əke* > TB *cake* 'river'

However, we have already seen above that there are two examples where the sequence **K^{w}o* did seem to preserve the labiovelar, namely in TB *walkwe* /wə́lk^{w}e/ 'wolf (?)' and TB *sekwe* 'pus'. Kim tentatively ascribes the retained labiovelar in *walkwe* to influence from the neighbouring **l* (Kim 1999: 166-167), but since the rationale would be that *l* is cross-linguistically a relatively more consonantal and that *k^{w}* in consonant clusters was preserved, which is not demonstrably the case (see below), this does not seem to be a fully cogent explanation. An assimilation to the initial **w-* might theoretically be possible, but cannot be demonstrated. TB *walkwe* thus remains puzzling, and provides an argument against a sound change from **K^{w}o* to **ko*. TB *sekwe*, TA *saku* 'pus', if properly derived from PIE **sok^{w}o-* 'juice, resin', would also show an unexpected retention of the labiovelar in the sequence **K^{w}o*.[11]

[10] Potentially, TB obl. *kolmai*, TA *koläm* 'basin' belongs here as well. Huard (2022: 411-416), with his new interpretation 'basin' rather than 'boat', derives this word from a PIE **g^{w}ol-u-* with an added suffix *-me*. Earlier etymologies depart from **kolmeh$_{2}$* or **kolmō-*. I have decided to exclude this word from consideration on account of the uncertainties involved in its etymological interpretation.

[11] Pinault assumes that the words where labiovelars were preserved in front of apparent **o* originally had a different nominative in **-ku* from earlier **-kwō(n)*, which was restructured, but (re)introduced **k^{w}* into the entire paradigm (Pinault 2008: 425). However, as there is no actual trace of an original nominative in **-ku*, this explanation is unconvincing.

Of the words that appear to show a development from **K^{w}o* to PT **ke*, the only example that is not a formation specific to Tocharian is TB *ke$_{u}$* 'cow' from PIE **g^{w}eh$_{3}$u-*. Here, it could also be assumed that delabialization occurred due to the following PIE **e* before laryngeal colouring, so that this example would align with the general merger of **K* and **K^{w}* before **e* instead (discussed above).

Four of the examples where **o* appears to have caused delabialization are formations specific to Tocharian, so that they could be considered less reliable and informative than TB *walkwe* 'wolf (?)' and *sekwe* 'pus'. The stem **k^{w}el-* from which TB *kele* 'navel' is derived could have lost the labiovelar in front of the **e*, before the formation of **k$^{(w)}$olo-* 'navel', and the expected labiovelar in **tek^{w}os-* 'river' may have similarly become delabialized in case forms such as gen.g. **tek^{w}es-os*, as suggested to me by Alexander Lubotsky (p.c., October 2024).

The missing labiovelar in TB *erkent* 'black' may not be directly understandable in the same way, but the related TB *orkamo*, TA *orkäm* 'dark; darkness' also lacks the labiovelar, so that this could be considered a feature of the root in Tocharian independent of the following **o* in the antecedent of the formation *erkent*. In root-final position, a number of other labiovelars were lost as well, without a clear single identifiable cause due to the inherent variation in following elements (see below). The lack of a labiovelar in TB *kete* 'whose' is unclear to me, but as a new formation, it may have been taken over from other forms of the paradigm as well. The original gen.sg. PIE **k^{w}eso* may have lost the labiovelar as well before it was replaced if **K^{w}e* was merged with **Ke*. Labiovelars also disappeared in a number of consonant clusters, as seen in the examples below.

58. PIE **h$_{3}$ok^{w}-s-* > **ok(s)* > PT **ek* > TB *ek*, TA *ak* 'eye'
82. PIE **nok^{w}tu-* > PT **noktə* > TA obl. *nokte* 'at night', adv. *noktiṃ* 'last night'
83. PIE **nok^{w}te̯u̯i̯o* > PT **nekt^{y}əw^{y}e* > TB *nekciye*, TA *nakcu* 'at night, last night'
90. PIE **penk^{w}to-* > PT **p^{y}ənkte* > TB *piṅkte*, TA *pänt* 'fifth'
127. PIE **u̯ōk^{w}s*, acc.sg. *u̯ok^{w}m̥* > **wok* > PT **wek* > TB *wek*, TA *wak* 'voice'

This environment is deemed preserverative by Kim instead, as he assumes that the trace of an earlier labiovelar in descendants of PIE **nok^{w}t-* 'night' is seen in the rounded vowel in TA *nokte* 'at night', and *noktiṃ* 'last night' (Kim 1999: 163–165). However, on account of the lack of a similar rounding in TA *nakcu* 'at night, last night' and the absence of a labiovelar in the TB cognate *nekciye* 'id.', the vocalism in Tocharian A should be explained differently (Pinault 2008: 424). Early umlaut due to a following **u* in an original nominative PIE **nok^{w}tu-* may be assumed instead of a labializing effect of an alleged preserved labiovelar (Pinault 1990: 181–190).

In the majority of examples, either the preceding **o* or the cluster could be regarded as the reason for delabialization (Hackstein 2017: 1325). However, it is still uncertain if **o*

caused delabialization at all (see above on **K^{w}o*). Furthermore there is at least one example where a labiovelar disappeared in a cluster without a neighbouring **o*, namely in the ordinal numeral PIE **penk^{w}to-* 'fifth'. This word does not retain any labialization in TB *piṅkte* (not ***piṅk$_{u}$te*), let alone in TA *pänt* where the velar stop is lost altogether due to a cluster simplification. However, an intermediate delabialization in the corresponding cardinal PIE **penk^{w}e* could also be assumed to have had an analogical effect.

At first sight, TB *laṅk$_{u}$tse* 'easy, light (not heavy)', derived from a PIE root **h$_{1}$leng^{wh}-* 'move lightly', would appear to indicate a retention before a consonant. However, TB *laṅk$_{u}$tse* should preferably be derived from a *u*-stem **h$_{1}$l̥ŋg$^{(w)h}$u-* also seen in Gr. ἐλαχύς 'small' and Skt. *laghú-* 'light', *raghú-* 'swift'. Consequently, the cluster TB /-k^{w}t^{s}-/ in this word probably arose through syncope from pre-PT **lənkut^{s}e*. The syncope could have taken place after PIE **penk^{w}to*, with a primary cluster, had already been simplified to pre-PT **penkto*.

We must furthermore assume that TB *akwatse* 'sharp' was secondarily derived from or affected by an unattested **ak^{w}ə* (PIE **h$_{2}$eḱu-*, see Pinault 2008: 445) after syncope had become inactive, since otherwise a quasi-PIE formation **h$_{2}$eḱut̥io* should have yielded TB ***āk$_{u}$tse* like TB *laṅk$_{u}$tse* from earlier **h$_{1}$lngʰu-tio-*. The labiovelar in the interrogative pronoun TB *mäkte* 'how', from **mo-k^{w}i-tod* 'what' (Pinault 2010; Peyrot 2018a: 80-81) may have been lost irregularly; TB *mäkte* 'how' rather than ***makte* /mə́kte/ is in any case not stressed in the same way as normal content words.

Delabialization may also be assumed in the cluster **K^{w}r*, *pace* Kim (1999: 151), who posits a continuation of **K^{w}r* as such until the Proto-Tocharian period on account of TA *kuryar* 'trade, commerce'. I think that this form is better interpreted as the result of **k^{w}* before a syllabic **r̥*, which was vocalized to **k^{w}ər* before the simplification of **k^{w}r* took place; cf. **k^{w}r̥i̯h$_{2}$wer* in the first list of this subsection. The lack of a labiovelar reflex in TB *karyor* 'trade, commerce' is likely due to analogy with the present stem PT **kərna-*, via **krəna-* from PIE **k^{w}rinh$_{2}$-* 'buy', which is one of the two examples showing a delabialization in a cluster **K^{w}r*. In that form, the initial labiovelar would necessarily have been in contact with the non-syllabic **r* until relatively late Tocharian liquid metathesis. The other example, PT **kramər* 'weight, heaviness' is dubious in its formation, as it is without exact parallels in the other Indo-European languages, but if it is correctly derived from the same PIE root **g^{w}reh$_{2}$-* 'heavy' as in Gr. βαρύς 'id.' and Lat. *gravis* 'id.', the labialization on the initial **g^{w}r-* was clearly lost, pointing to delabialization in this type of cluster.

70. PIE **k^{w}rinh$_{2}$-* > **krəna-* > PT **kərna-* > TB *kərna-* 'buy' (present stem)
30. PIE **g^{w}reh$_{2}$-* → **g^{w}rh$_{2}$-mr̥-* > PT **kramər* > TB *krāmär* 'weight, heaviness', der. *kramartstse* 'heavy', TA der. *krāmarts* 'heavy'

In one contrasting word, TB *kärweñe*; TA *kärwaṃs-* 'stone', a labiovelar before an original syllabic **r̥* was delabialized, but this was probably the result of dissimilation with the **w* in the following syllable instead (so Ringe *apud* Kim 1999: 151 fn. 26; Pinault 2008: 425).

31. PIE **gʷr̥h₂u̯o-* > PT **kərwenʸe* > TB *kärweñe*, TA *kärwaṃs-* 'stone'

The lack of a labiovelar in TB *kəska-* 'scatter, strike apart' is difficult to explain departing simply from **gʷʰn̥-sk-*. This would be expected to develop as **gʷm̥-sk-*, where Tocharian A shows a retention of the labiovelar in *kumnäs-* 'come'.

32. PIE **gʷʰen-* 'hit' → **gʷʰn̥-sk-* > PT **kəska-* > TB *kəska-* 'scatter, strike apart'

However, it may be imagined that the *sk*-present was influenced by the original root present, as reflected in Skt. 3sg. *hánti*, 3pl. *ghnánti* and Hitt. 3sg. *kuenzi*, 3pl. *kunanzi* from PIE **gʷʰenti*, **gʷʰnenti*.[12] Either **gʷʰe* or **gʷʰne* —or perhaps both— may have undergone regular delabialization and influenced the *sk*-present.

The remaining examples of delabialization occur in root-final position. There is no single rule that can capture all instances, since the following elements were variable, but understandable causes may be outlined.

16. PIE **dʰegʷʰ-* > PT **tˢək-* > TB *tsək-*, TA *tsäk-* 'burn'
18. PIE **dʰi̯h₂gʷ-* > PT **tˢak+a-* > TB *tsaka-*, TA *tsākā-* 'pierce'
33. PIE **hₑeh₁gʷʰ-* > PT **yok-* > TB *yok-*, TA *yok-* 'drink'
38. PIE **hₑleng*ʷʰ- > PT **lənk-* > TB *lənk-*, TA *länk-* 'hang, dangle'
43. PIE **hₑregʷ-* → **h₂u̯gʷ-ont-* > **erkont-* > PT **erkent-* > TB *erkeṃt*, TA nom.pl. *arkaś* 'black'
87. PIE **pekʷ-* > PT **pək-* > TB *pək-*, TA *päk-* 'cook, ripen'
108. PIE **tekʷ-* 'run, flow' → **tekʷos-* > **tekos* > PT **tʸəke* > TB *cake* 'river'

Pinault proposes that some of the labiovelars in question were lost in consonant clusters, as in the cases discussed above, for example in the *s*-presents of PT **pək-* 'cook, let ripen (act.); boil, ripen (mid.)' from PIE **pekʷ-* and PT **tˢək-* 'burn' from PIE **dʰegʷʰ-* (Pinault 2008: 424). This would also work for the root PIE **dʰih₂gʷ-*, which must have been in

[12] The parallel *sk*-present found in Hittite *kuu̯aške/a-* ipfv. 'kill, slay' may reflect a secondary formation from the present on account of the vocalism (see Hackstein 1995: 200-201), although Kloekhorst (2008: 485-486) has argued that it can be regularly derived from PIE **gʷʰn̥-ske/o-*. If the Hittite and Tocharian formations are both old, analogical influence from the earlier root present in Tocharian is still a possibility.

direct contact with suffixal consonants both as an original root present (LIV: 142) and in its eventual nasal present inflection in TB prs. *tˢakna-* 'bite, pierce'. PT **lənk-* 'hang' from PIE **h₁leng^{wh}-* forms a root present as well in Tocharian, and can thus have lost the labiovelar in contact with a following consonant.

Kim rather ascribes the disappearance of the labiovelars to a delabialization in the sequence **-K^{w}o*, which occurred in part of the paradigm of the original thematic presents of some of these verbs. This **-K^{w}o-* alternated with **-K^{w}e-*, where the labialization was lost due to palatalization to **-k^{y}e-* and, eventually, TAB *-śä-*. Due to these two changes, the root-final labiovelars would have been completely eliminated from the paradigm (Kim 1999: 160). This may work for *d^{h}eg^{wh}-* and *pek^{w}-*, which can be reconstructed with a thematic present in Proto-Indo-European (LIV: 133, 468), but the explanations put forward by Pinault and Kim are not mutually exclusive. It is also possible that the labiovelar was delabialized only in **K^{w}e*, and that the delabialization was analogically spread from there (cf. above).

While attested without a labiovelar, the verb TAB *yok-* 'drink' seems to have initially escaped the effects of these processes. Though complicated in terms of its diachronic derivation (see Peyrot 2013: 448-449), the development of the verb PIE **h₁eh₁g^{wh}-* 'drink' to TAB *yok-* is most easily explained if there was still a labiovelar present at a relatively advanced stage to cause a rounding of the vowel. Otherwise, the expected outcome is ***yek-*.[13] The root-final labiovelar in **h₁eh₁g^{wh}-* should thus probably have avoided delabialization in at least part of the paradigm.

It is clear that the developments of original labiovelars from Proto-Indo-European in Tocharian is a complicated matter. At a shallow Tocharian level, it is also necessary to account of the verb TB *k^{w}əĺpa-*, TA *k^{w}älĺpa-* 'desire'. On account of the palatalized lateral, this must derive via metathesis from earlier **k^{w}l^{y}əpa-*. The viability of the sequence **k^{w}l* in this word accords with the retention of the labiovelar in **k^{w}ek^{w}lo-* 'wheel' (see above).

The relative chronology of changes implied by the varying reflexes of original labiovelars in these different environments is summarized in Table 2.5. Syncope, epenthesis, stress shift and liquid metathesis will be more extensively discussed in Chapter 5.

[13] According to Kim (1999: 165), the verb PT **yok-* 'drink' retained its original labiovelar before suffixes starting with a consonant (contrary to the delabialization in that context supported by Pinault and Hackstein and discussed above).

Table 2.5: A relative chronology of sound changes involving labiovelar stops.

relative chronology	example	instead of	reconstruction
nkʷt > *nkt*	*piṅkte*	***piṅk$_u$te*	**penkʷto-*
? *kʷr*, *kʷm*, *kʷn* > *kR*	*kärnā-*	***kurnā-*	**kʷrinḫ$_2$-*
syncope 2nd syllable	*laṅk$_u$tse*	***laṅktse*, ***läṅkwatse*	**h$_1$lo̯gʰu+ti̯o*
epenthesis/stress shift	*kokale*	***kokle*, ***kwakle*	**kʷekʷlo-*
liquid metathesis	*kulypa-*	***kälypa-*	**kʷlʸəpa-*

I have argued above, against Kim (1999) and in accordance with Pinault (2008) and Hackstein (2017), that there is at present no convincing example of preserved labiovelars before other consonants, apart from **kʷl*, as necessitated by TB *kokale* 'wagon' from **kʷekʷlo-* 'wheel'. The labiovelar in TB *laṅk$_u$tse* only secondarily came to be followed by **t*ˢ after syncope, as the stem ended in a vowel: **h$_1$lng*ʷʰ*u-*. Furthermore, the labiovelar in the stem PT **kʷərya-* as in TA *kuryar* 'trade, commerce' was not followed by the **r* from the moment that syllabic resonants had become vocalized: **kʷr̥i̯ḫ$_2$-* became **kʷəry*a-. The velar in the present stem TB *kärnā-*, on the other hand, was directly followed by the **r* until liquid metathesis: **kʷrinḫ$_2$-*. Whether liquid metathesis had already occurred in Proto-Tocharian is unclear, so that either PT **krəna-* or PT **kərna-* could be reconstructed (see also 5.2.2.3).

Because the Tocharian labiovelars are so complicated, perhaps in some respects more complicated than has been assumed, it is difficult to fit their developments into the relative chronology. Importantly, the development from **Kʷo* to **Ko* cannot be taken for granted on account of TB *walkwe* 'wolf (?)' and *sekwe* 'pus'. If **Kʷo* did remain labialized with delabialization occurring only in clusters and before original PIE **e* or **ē*, the latter change can still be dated before the merger of PIE **o* and PIE **ē* that occurred with palatalization (2.4), but it cannot be placed before the unrounding of PIE **o* to pre-PT **ë* that took place earlier.

2.3.3.3 Secondary labiovelars

Sequences of a PIE velar plus consonantal **u̯*, i.e., of the type **Ku̯o-*, end up as labiovelars in Tocharian (Kim 1999: 158, 163, 167). This is seen in the following examples:

35. PIE **h$_1$eḱu̯o-* > **ekwo* > PT **yəkʷe* > TB *yakwe*, TA *yuk* 'horse'
49. PIE **h$_2$erǵu-* → PT **arkʷi* > TB *ārkwi*, TA *ārki* (**kʷi* > TA *ki*) 'white'
61. PIE **ḱu̯onm̥* > **kwonəm* > PT **kʷenə* > TB *kweṃ*, TA *koṃ* 'dog (obl.sg.)'
79. PIE **n̥ḱu̯o-* > **enkwo* > PT **enkʷe* > TB *eṅkwe*, TA *oṅk* 'man'

119. PIE ±*$\underset{\cup}{}$*erǵ-unt-* → **werkwənt-* > PT **wʸərkʷənt-* > TB obl. *yerkwantai* /yerkʷə́ntay/, TA *wärkänt* 'wheel'[14]

As shown by the examples **h₁ek̑$\underset{\cup}{}$uo-* 'horse' and **n̥k̑$\underset{\cup}{}$uo-* 'mortal → man', these new labiovelars did not disappear before PIE **o* (cf. Ringe 1996: 42), and thus accord with the evidence from PIE **ulkʷo-*, TB *walkwe* 'wolf (?)' and **sokʷo-* 'sap', TB *sekwe* 'pus' with their retained labiovelars before **o*.

Additional secondary labiovelars came into existence in contact with original vocalic PIE **u* as well, but this development is treated in subsection 2.4.3 instead, together with a more general overview of the development of PIE **u*.

2.3.4 The stop system

Tocharian is notorious among the Indo-European languages for its lack of distinctive manners of articulation in the stops. The Tocharian stops are simply /p t k kʷ/ with the additional affricates /tˢ/ and /c/ (±[tɕ], derived from palatalized PIE **t*/*dʰ*). This is a clear departure from the Proto-Indo-European system, which had up to three distinct manners of articulation at each place, cf. the traditional reconstruction of the dental series as PIE */t dʰ d/. As mentioned briefly in section 2.2, the exact phonetic values of the Proto-Indo-European stops are still a matter of debate (e.g., Kümmel 2007: 299–309), but it is clear that there were three distinct series. At some point in pre-Proto-Tocharian, these distinctions were lost. The examples below show how voiceless stops *p*, *t*, *k*, *kʷ* result in Tocharian from PIE "voiced" and "voiced aspirated" stops (cf. Krause & Thomas 1960: 64–68; Ringe 1996: 152–154, Pinault 2008: 422–424, etc.).

1. PIE **bʰeh₂ǵʰu-* > PT **pok-* > TB obl. *pokai*, TA *poke* 'arm'
2. PIE **bʰreh₂tēr* > PT **protʸer* > TB *procer*, TA *pracar* 'brother'
3. PIE **bʰr̥ǵʰro-* > PT **pərkre* > TB *pärkare* /pərkə́re/, TA *pärkär* 'tall'
21. PIE **dʰugh₂tēr* > PT **təkatʸer* > TB *tkācer* /tkácer/, TA *ckācar* 'daughter'
29. PIE **gʷm̥-sk-* > PT **kʷəmnəsk* > TB *kənməsk-*, TA *kumnäsk-* 'come'
32. PIE **gʷʰn-sk-* > PT **kəska-* > TB *kəska-* 'scatter, strike apart'
35. PIE **h₁ek̑$\underset{\cup}{}$uo-* > PT **yəkʷe* > TB *yakwe* /yə́kʷe/, TA *yuk* 'horse'
43. PIE **h₁regʷ-* → **h₂r̥gʷont-* > PT **erkent-* > TB *erkent*, TA pl. *arkaś* 'black'
44. PIE **h₁rudʰro-* > PT **rətre* > TB *ratre* /rə́tre/, TA *rtär* 'red'
76. PIE **medʰu-* > PT **mʸətə* > TB *mit* 'honey'

[14] The development of this word is unexpected in both Tocharian A and B. In TA the labiovelar was apparently delabialized due to dissimilation with the word-initial *w-*. In TB there is an unexplained development from **ə* to *e* in the first syllable; ***yärkwantai* /yərkʷə́ntai/ would have been the expected counterpart of TA *wärkänt* (Kim 1999: 168; Adams 2013: 547–548).

It is difficult to establish any intermediate stages in which for instance three manners of articulation may have been first reduced to two, before finally becoming one. It is clear that not all stops simply merged into a voiceless series, however. There are a few different developments that affected PIE *b^h and *d, and not their respective counterparts *p and *t/d^h produced at the same point of articulation. In the following subsections, we will take a look at the developments of these stops and consider what they tell us about the evolution of the Tocharian consonant system through time.

2.3.4.1 Assimilation of *mb^h to *m

There is a discrepancy in the reflex of the Proto-Indo-European labials *p and *b^h when they followed a nasal *m. Namely, a PIE *b^h was lost in that position, while *p was retained.[15] There are only very few examples of these clusters:

24. PIE *$ǵomb^ho$- > PT *$keme$ > TB *keme*; TA *kam* 'tooth'
104. PIE *$stemb^hH$- > PT *$s^ytʸəma$- > TB prt. *ścəma*- 'stand'
110. PIE *$temp$- 'stretch, extend' > PT *$tʸəmp$- > TB *cəmp*-, TA *cämp*- 'be able'

This change is important for the phonetic interpretation of the stop series in the earliest pre-Proto-Tocharian, as the difference in the behaviour of *b^h and *p in this environment should be phonetically understandable. On the basis of the standard interpretation of *b^h as breathy voiced, Ringe conceptualized this development from PIE *mb^h to PT *m as either an extension of breathy articulation of *b^h to the preceding *m followed by a deletion of the stop element, or as a progressive assimilation via an intermediary stage *mb, i.e., *m + simple voiced *[b] (Ringe 1996: 42-44).[16]

Peyrot has considered what this sound change might mean for the interpretation of the early pre-Proto-Tocharian stop system, also from the perspective of the Glottalic Theory, and concluded that the series *b^h (and by extension potentially also the other voiced aspirates) would have had a weaker articulation than *p (etc.) in one way or another (Peyrot 2022a: 91–92).

---

[15] One single potential etymology containing *b in the cluster *mb shows a development like *b^h in this position: PIE *$lembH$- 'hang' > PT *$ləma$- > TB *ləma*-, TA *lämā*- 'sit' (LIV: 411). If *mb indeed developed as *mb^h, that could imply that *b^h and *b were more similar to one another in some respect than to *p.

[16] Ringe also discusses the possibility that voiced aspirates at other places of articulation could have been lost after nasals as well, but are no good examples (Ringe 1996: 42-44). The development of *$h₁lund^h$- to PT *$lən$- suggested by Ringe does not hold, since the Tocharian stem in question is rather *$lənn$- from *$ləntn$- (Peyrot 2013: 446–447), so that the type of nasal cluster simplification seen here is really restricted to the labial place of articulation.

However, this can be understood in all three models for the Proto-Indo-European stop system given above in 2.2. From the traditional point of view, the voicedness of *b^h could have facilitated full assimilation from *mb^h to *m. In the glottalic model, *mb^h is interpreted as *"mb", so that assimilation can have taken place in the same way. Kloekhorst's model, with an interpretation of *b^h as *"p" may be interpreted with an intermediary voiced stage *[mb] before eventual full assimilation. The changes involved in all three models are relatively minor, but they should probably precede the general devoicing of Tocharian stops that applies to all stops in non-labial nasal clusters as well.

2.3.4.2 Early assibilation of dental stops

The dental series shows different developments as well. PIE *t and *d^h, but not *d, undergo an assibilation process when followed by *y (= *$i̯$), ending up as PT *t^s (Winter 1962: 20-22; Ringe 1996: 79).[17] The sequence PIE *dy yielded PT *y instead, with complete loss of the *d (Ringe 1996: 65). The assibilation of *ty and *d^hy and their contrast to *dy are seen in the following examples.

18. PIE *$d^hi̯Hg^w$- > *d^hyag^w- > PT *t^saka- > TB *tsaka*- 'pierce, bite'
86. PIE *$peh_2nti̯h_2$ > *$pāntya$ > PT *$pont^sa$ > TA *pontsāṃ* 'all'
117. PIE *$u̯lh_2nti̯h_2$ > *$wlantya$ > PT *$lantsa$ > TB *lāntsa*; TA *lānts* 'queen'

107. PIE *$su̯id-i̯e$- > *$swidye$- → PT *$səya$- > TB *səya*- 'sweat'

This indicates that *d had in some way changed or disappeared in the sequence *dy so as to no longer be a stop at the time that assibilation took place: if there was no stop element present in *dy anymore, it logically could not have become an affricate.

There is also another proposed assibilation of only word-final *-*ti* and *d^hi to TAB *ṣ*, potentially via *-*si*. This change was proposed by Jasanoff (1987: 106–112), based on the development of three suffixes with this shape.

34. PIE *h_1id^hi > PT *$pə$-$yəṣə$ > TB *paṣ*, TA *piṣ* 'go!'
111. PIE *-*ti* > PT *-$ṣə$ > TA -(*ä*)*ṣ* (prs.3sg. verbal ending)
112. PIE *-*ti* > PT *-$ṣə$ > TA -(*ä*)*ṣ* (ablative ending)

The first example is the irregular imperative of 'go', PIE *h_1id^hi > PT *$pə$-$yəṣə$ > TA *piṣ*, TB *paṣ* 'go!' (*pə*- is the Tocharian imperative marker). Another verbal ending, 3sg. primary

[17] For *$d^hi̯$ to PT *t^s this is not as firmly established due to a dearth of reliable example; only PIE *$d^hi̯Hg^w$- to PT *t^saka- 'pierce, bite' supports this development; cf. Hackstein (2017: 1324), who omits mention of *$d^hi̯$ altogether.

PIE *-*ti*, might have regularly yielded TA *-(*ä*)*ṣ* in, e.g., *lkāṣ* 'sees' from PT **ləka-ti*. Jasanoff's third example is the Tocharian A ablative in -(*ä*)*ṣ*, which he compares to the Hittite ablative in -(*a*)*z* (/(a)t^{z}/) from **-ti*, possibly connected with the adverb **h₁éti* 'from this' (Jasanoff 1987: 109–110). This development is accepted by Ringe (1996: 79–80) and Hackstein (2017: 1325), but not by Pinault (2008: 620).

There are a few factors that cast doubt on this sound change. If the change of word final *-*ti* to TAB *ṣ* occurred as posited by Jasanoff, it appears that assibilation of *-*ti* to an intermediate *-*si* needs to have happened before palatalization, since palatalized **s* is the usual source of TAB *ṣ*. This development would then have to be separated from that of word-final *-*nti*, because that became TAB *-ñc* instead, rather than **-*nṣ* vel sim. It is furthermore noteworthy that the development of **ty* and **ti* would not be entirely parallel, since the former remained affricated TAB *ts*, while the latter would have become totally assibilated in TAB *-ṣ*. A special word-final development on the consonant is also difficult to envision, since it is the vowel *-*i* that was word-final until relatively late apocope.

Whether or not **ty*/*d*^{h}*y*- initially remained distinct from initial **d*- is also difficult to tell, since the respective sequences did eventually merge as TAB *t*^{s}- (compare the examples above and those given for **d* in the next subsection). In order to maintain this potential distinction, I will use the symbol "*ċ/ż*^{h}" to denote the product of assibilated **t* and **d*^{h}. Especially in the case of "*ċ*", the dot is intended as a neutral diacritic to avoid confusion with TAB *c* from (later) palatalized **t*/*d*^{h}, which resulted from a different change, and to avoid using the same symbol as either [tʂ] <č> or [tɕ] <ć> in Uralic phonological transcription. It is not meant to represent any specific phonetic value, but may be interpreted as *[ts] or *[tʃ] or *[tɕ] at the earliest stages, all typologically reasonable affricate outcomes of sequences of the type *-*tj*- (see Kümmel 2007: 258–261).

2.3.4.3 The aberrant development of PIE **d*

Just as the cluster **dy* developed differently from **ty* and **d*^{h}*y*, the development of PIE **d* in Tocharian in general is a more complicated matter than that of **t* and **d*^{h}. As a result, it may also be more informative. PIE **d* essentially has four commonly accepted reflexes, namely *t*^{s}, *ś*, *t*, or ∅. These will each in turn be discussed in this subsection.

It is well-known that PIE **d* disappeared before **w*, **y* and **r*, as seen in the following examples (see Ringe 1996: 64-66, Hackstein 2017: 1324):

10. PIE **dor-u-*, **dr-eu̯-* > PT **orə* > TAB *or* 'wood'
11. PIE **drem-* 'run' > der. PT **rəmer* > TB *ramer* 'quickly'
12. PIE **du̯eiˌ-* 'fear' > PT **wəy-* > TB *wəy-*, TA *wäy-* 'be frightened'
14. PIE **du̯ito-* > PT **wəte* > TB *wate*, TA *wät* 'second (2nd)'
15. PIE **du̯oh₁* > PT **wəw* > TAB *wu* 'two (2)'

106. PIE **sṷuh̥₂dro-* > PT **sware* > TB *swāre*, TA *swār* 'sweet'
107. PIE **sṷid-i̯e-* > **swidye-* → PT **səya-* > TB *səya-* 'sweat'
121. PIE **u̯id-u̯on-* > PT **wəwe* > TB *ūwe* 'wise'
126. PIE **udr-* > PT **wər* > TB *war*, TA *wär* 'water'
126. PIE **udr-io-* > PT **wərəye-* > TB der. *wriyeṣṣe*, TA *wri* 'watery'

In the first example PIE **dor-u-*, the disappearance of **d* must originally have taken place in particular forms like the genitive **dre-u-*, with a zero-grade in the root. Thus, apparently, **doru* : **dreu* became **doru* : **reu* and then analogically **oru* : **reu* (on this development, see also 2.3.4.4).

In a few cases, PIE **d* is reflected as PT **t*, merged with non-palatalized PIE **t* and **dʰ*. This is found in some clusters, especially after PIE **n* in the cluster PIE **nd* (see Malzahn 2010: 742; Hackstein 2017: 1324). The main examples are PIE **spend-* to PT **spənta* 'trust' and PIE *splend-* to PT **plənta-* 'be pleased'. The development of the root PIE **kedh̥₂-* to **kəta-* 'strew, scatter' is less clear. The nasal present (an old formation on account of cognates like Gr. σκίδνημι 'scatter', see LIV: 550) shows a metathesised variant TB *kənta-* next to renewed *kətna-*, but it is unclear whether the metathesis would have been early enough for this verb to develop like **spend-* and **splend-*.

62. PIE **kedh̥₂-*, prs. **kədnh̥₂-* or **kn̥dh̥₂-* ? > PT **kəta-*, **kənta-* > TB *kəta-*, *kənta~kətna-*, TA *kätā-*, *känā-* 'strew'
80. PIE **neud-*, prs. **nud-sk-* > **nutsk* > PT **nətk-* 'push away'
103. PIE **spend-* > PT **spənta-* 'trust' (secondary *-a*)
104. PIE **splend-* > PT **planta-* 'be pleased' (secondary *-a*)

Outside of these clusters, PIE **d* in word-initial position generally became TAB *tˢ* (Ringe 1996: Hackstein 2017: 1323–1324). When palatalized, the result was TAB *ś*. There are not many examples available. The palatalized *ś* is most clearly seen in the numeral 'ten (10)' TB *śak*, TA *śäk* from PIE **deḱm̥* (Ringe 1996: 146–147; Hackstein 2017: 1324). The reflex TAB *tˢ* is found most clearly in a few verbal roots, where palatalization was analogically removed in certain forms.

5. PIE **demh̥₂-* > PT **tˢəma-* > TB *tˢəma-*, TA sbj. *tˢämā-* 'grow'
6. PIE **der-* 'split' > PT **tˢər-* > TB *tˢər-*, TA *tˢär-* 'separate'
7. PIE **deu̯h̥₂-* > PT **tˢəwa-* > TB *tˢəwa-* 'attach oneself to', TA *tˢäwā-* 'fit; obey'
8. PIE **deu̯k-* 'pull' > PT **tˢəwka-* > TB *tˢawka-*, TA *tˢäwkā-* 'suck (out)'

It is notable that there are no secure intervocalic examples of **-d-*. It is therefore difficult to say for sure what happened to it in that position (cf. already Winter 1962). The only

instance adduced by Ringe is TB *witsako* /witˢəko/ 'root', which he derives from a pre-form **wẹdəga-* from pre-Proto-Ossetic **wēdāga* (Ringe 1996: 146). However, as Bernard shows, it is not possible to derive this word from any known Iranian source (Bernard 2023: 222–227), so that its value for determining of the details of a Tocharian sound law is extremely questionable. This word is not inherited from Proto-Indo-European in any case, meaning that it would not show a development of PIE **d*. In old borrowings from Indo-Iranian, IIr. **d* is rendered as PT **t* instead: e.g., **spaldaka-* ⇒ PT **speltke* > TB *speltke*, TA *spaltäk* 'zeal, effort' (Bernard 2023: 75–77).

Unfortunately, the few potential reflexes of PIE word-internal **d* are not fully probative. One example of Indo-European origin could be in the dual of 'foot', **pod-h₁e*, which, according to Winter, regularly became pre-PT **peyə*. This was rather untransparent dual was then enlarged with the productive dual suffix *-ne* to create PT **peyəne*, attested as TA *peṃ*, TB *peyne*, *paine* (Winter 1962: 31). This would point to a *y*-reflex in intervocalic position (or ∅ with *y* arising in the resulting hiatus). The nominative singular *paiyye* may be derived from the dual via **peyə-ye*, so that it is of little help in determining what happened. With a parallel loss of **d*, it is assumed that PT **sʸəm-* 'sit' derives from PIE **sed-* 'sit' in some way; the final *-m* would be taken from *ləma-*, with which it forms a suppletive paradigm, although clearly the final *-a* was not taken over (Adams 2013: 717, Pinault *apud* Malzahn 2010: 927).[18]

2.3.4.4 Interpreting the development of PIE **d* in Tocharian

The development of PIE **d* in Tocharian can be viewed in a number of different ways. In order to understand the development of PIE **d*, a coherent account is needed of all its different reflexes, including a plausible phonetic interpretation. Key points that need to be explained by such an account are:

1. The loss of PIE **d* in various clusters (and likely in intervocalic position), especially in contrast with preservation of **t* and **dʰ* in similar environments
2. The merger of PIE **d* with the outcomes of PIE **t* and **dʰ* in the cluster **-nd-*
3. The development of PIE **d* to TAB *tˢ* (with palatalized counterpart TAB *ś*)

[18] A new word-internal reflex of *ts* has been proposed for *tənts-* 'scatter', derived by Hackstein from PIE **tund-* 'strike, beat'. The *ts* would in this case have spread by analogy from the non-nasal allomorph of the root, **tud-* > **təts-* (Hackstein 2017: 1324), so that it would not be a direct reflex as such. Since no allomorph without the nasal is attested, this etymology as a whole remains rather speculative, so that this word cannot give us reliable information on the development of intervocalic **-d-*.

The difficulty lies in plausibly combining these three developments in a single phonetic interpretation while also taking into account the continued separate existence of PIE **t* in Tocharian. Ringe provides a rather straightforward reconstruction of the development from PIE **d* to TAB *t*^{s} via an intermediate **“d^{z}”* and explains the different behaviour of these dentals in clusters with **y* by positing an earlier devoicing of **d^{h}* and the other voiced aspirates to **“t^{h}”*, etc. (Ringe 1996: 47–48, see also 2.3.4.5). However, neither the putative development from **d* to **“d^{z}”* nor the development from **d^{h}* to **“t^{h}”* are otherwise observable in the Tocharian material.

Furthermore, while a change from **d* to **“d^{z}”* to TAB *t*^{s} is at first glance straightforward and economical, I do not think that this is particularly likely as a development for typological and phonetic reasons. Assibilation, or affrication in general, can be understood as a reinterpretation of the turbulent airflow that occurs right after a stop's release as a fricative or sibilant. Such turbulent airflow is generally more pronounced in voiceless and especially in aspirated stops than in voiced ones, so that one would least expect a voiced stop to become an affricate (Hall & Hamann 2006). For example, when followed by [i] or [j], two sounds that increase post-release turbulence with their restriction of the airflow through the mouth and thus often promote reinterpretation of a stop as an affricate, it appears that there is less turbulent airflow following [d] than following [t]. Based on this, [d] is only expected to assibilate when [t] does, too, not the other way around (see Hall, Hamann & Zygis 2006 on this topic).

Similarly, on “spontaneous” affrication, Kümmel (2007: 173) states that it is almost exclusively found to happen with voiceless plosives. He does, however, give the Tocharian change as **d* to **“d^{z}”* and also palatalized **d^{y}* to **“d^{zy}”* without further comment, directly following Ringe.[19] However, there does not seem to be a good parallel for direct affrication of just **d* and not **t* —let alone a supposed **“t^{h}”* from PIE **d^{h}*, which should be the most likely to undergo affrication—, and this lack of phonetic naturalness casts doubt on the seemingly straightforward path from **d* to TAB *t*^{s} via **“d^{z}”* before general Tocharian devoicing.

[19] Kümmel's other example of spontaneous assibilation of **d* is found in Middle Cornish, where **d* changed to **d^{z}*, and eventually *z*, while **t* changed to **t^{s}*, and on to *s*. Here the environment in which **t* was changed, namely after **l* and **n*, seems more restricted than that of **d*, which assibilated after any vowel. However, this cannot be taken as an example of **d* being after all more prone to assibilation than **t*, since Cornish intervocalic **d* is the result of earlier **t* through lenition. At the point in time when the Cornish assibilation took place, **t* existed almost exclusively word-initially, where neither it nor **d* was affected. Between vowels, there was a contrast between geminate **tt*, voiced **d* and fricative **ð*. Singleton **t* did not occur intervocalically, a position where **d* assibilated to **d^{z}*/*z*, because it had turned into that self-same **d* at an earlier stage through a process of lenition.

The particular development of PIE **d* in clusters is not properly addressed by Ringe's hypothesis either, so that the lack of affrication of **d* in PIE **spend-* > PT **spənta-* 'trust' and PIE **splend-* > PT **planta-* 'be pleased' remains unexplained. On the basis of precisely the development in these verbs, Malzahn instead assumes that **d* first became a voiced fricative *[ð] before turning into TAB *t*^{s} (*ś* when palatalized). An intermediate *[ð] fits here, because in a position after a nasal, fricativization of obstruents is typologically less likely to occur (Malzahn 2010: 742). A lenited intermediary *[ð] could also be connected with the disappearance of PIE **d* in the clusters **dw*, **dy* and **dr* (i.e., via *[ðw], *[ðy], *[ðr]), as an intermediate stage in that development.

If we consider the disappearance of **d* assumed for PT **peyə* 'foot' (and perhaps somehow in PT **s^{y}əm-* 'sit'), lenition via a [ð] provides a plausible path for **d* to have yielded intervocalic **y* or ∅ as well. Especially intervocalic [ð] is often further lenited and lost (e.g., Kümmel 2007: 109–113). Since the loss of **d* before **w*, **y* and **r* also took place in word-initial position, lenition of **d* to *[ð] could be a general development, not restricted to word-internal position. In this way, one could reconstruct a general development of **d* to *[ð], blocked in at least the clusters **nd* (PIE **spend-* > PT **spənt+a-*, PIE **splend-* > PT **plant+a-*) and **dsk* (PIE **nud-sk-* > PT **nətk-*).

For the development to TAB *t*^{s} in initial position there are many paths that could be imagined, but none can be proven on the basis of the available data. If *[ð] is taken as a first step in the development, based on the reflexes ∅ and *-t-* after nasals, some possibilities for further developments to TAB *t*^{s} are:

- [ð] > [z] > [dz] > [ts]
- or [ð] > [θ] > [tθ] > [ts]
- or [ð] > [θ] > [t^{h}] > [ts]
- or [ð] > [dð] > [dz] > [ts]
- or even [ð] > [j̟] > [j̟j̟] > [cɕ] > [ts][20]

The fortition that seems to be necessary in all of these sequences could be tentatively interpreted as the result of the phonologically strong onset position in which TAB *t*^{s} is obtained from **d*. Again, none of these hypothetical intermediary stages are reflected in the Tocharian material, and they thus remain speculative.

Peyrot briefly considered the possibility of interpreting **d* as glottalized **ʔd* in early pre-Proto-Tocharian, but concludes that there is no direct evidence in favour of it to be found in Tocharian (Peyrot 2022a: 91–92). We may still explore if there is any indirect evidence. Disappearance of PIE **d* in the clusters **dw*, **dy* and **dr* could in principle be

[20] Compare the allophonic realization of initial /j̟/ in Spanish as [j]~[j̟]~[j̟j̟]~[ɟ] (Hualde 2005: 383–384) and the development from **ty*/*d^{h}y*, with a clear original palatal element, to TAB *t*^{s} (see 2.3.4.2).

interpreted in a glottalic framework as a change from glottalic *“*ʔd*” to *“*ʔ*”. The unusual analogical development seen in **doru* : **dreu* to **oru* : *reu* ‘wood’ seen above could then be understood as a development from glottalic **ʔdoru* : *ʔdreu* to **ʔdoru* : **ʔreu* and then **ʔoru* : **ʔreu* instead, as suggested by Alexander Lubotsky (p.c., March 2025). In that way, the element that was analogically spread was a consonant, **ʔ*, rather than a lack of a consonant, ∅. The Tocharian development of PIE **d* in clusters with resonants could then be considered as a conditioned debuccalization from *“*ʔd*”. In the cluster **-nd-*, where PIE **d* behaves like **t* and **dʰ*, a conditioned deglottalization could also be assumed, thus from glottalic *“*-nʔd-*” to * “*-nd-*”,[21] whence TAB *-nt-*. Within Kloekhorst’s hypothesis that the PIE **d* was originally an unvoiced *[ʔt], the interpretation of the Tocharian behaviour of **d* could be essentially the same as just outlined for *“*ʔd*”.

The behaviour of PIE **d* in clusters can thus be understood in two different ways depending on the assumed original qualities of this consonant in Proto-Indo-European (either voiced as per the traditional reconstruction of Proto-Indo-European, or glottalized as per the Glottalic Theory). The standard word-initial reflex of **d* as TAB *tˢ*, with its palatalized counterpart TAB *śˌ* still requires further discussion. Reasoning backwards from TAB *tˢ*, an aspirated *[tʰ] is a good candidate for a direct precursor to TAB *tˢ* from a typological perspective (see Kümmel 2007: 173–175). How this would have arisen from PIE **d* is unclear, but a partially parallel development is found in Germanic. PIE **d* became PGm. **t*. This is reflected as aspirated [tʰ] in most Germanic languages, with a further development to an affricate [tˢ] in High German and Danish. However, the other, concomitant changes to the stop system found in Germanic are not paralleled in Tocharian, and a reconstruction as *[tʰ] in itself does not neatly match the other observable developments of PIE **d* in Tocharian: it is difficult to understand how such a sound would be lost in clusters with a **r*, **y* or **w*, or in intervocalic position.[22]

In the interpretations of the development of **d* just discussed, several intermediary stages have to be posited, summarized in Table 2.6. On the one hand, various seemingly disparate developments of **d* in Tocharian can be connected with the hypothesis of a single initial change: lenition. On the other hand, if a glottalic stop is taken as a precursor to a *[tʰ], such a reconstruction is more parsimonious. Preference for one scenario over the other will probably depend on the interpretation of stop qualities in the Proto-Indo-European from which Tocharian split, which cannot be reliably established on the basis of Tocharian itself due to the general lack of distinctions. The geminates of Kloekhorst’s

[21] In the glottalic framework, the *“*d*” corresponds to the **dʰ* in the traditional reconstruction.

[22] One could imagine an alternative lenited intermediate [θ] from [tʰ] rather than [ð], as the latter is based on the assumed characteristics of PIE **d* rather than on the Tocharian reflexes, so that a further alternative scenario could also involve a general reconstruction of this phoneme as **tʰ* with a conditioned lenition to *[θ] and then further to ∅ in clusters with **r*, **y*, **w*.

reconstructed system are just as absent from the Tocharian reflexes as the traditional voicing distinctions, and even an originally glottalic voiced *“*ʔd*” could have developed via a fricative like [ð] at some intermediary stage.

Table 2.6: Special developments of PIE **d* in Tocharian interpreted from the traditional reconstruction of Proto-Indo-European stops, within the Glottalic Theory, and based on Kloekhorst's model of Proto-Indo-Anatolian stops.

Traditional	Glottalic	Kloekhorst
dw > *ð̯w* > *w*	*ʔdw* > *ʔw* > *w*	*ʔtw* > *ʔw* > *w*
dy > *ð̯y* > *y*	*ʔdy* > *ʔy* > *y*	*ʔty* > *ʔy* > *y*
dr > *ð̯r* > *r*	*ʔdr* > *ʔr* > *r*	*ʔtr* > *ʔr* > *r*
nd > *nt*	*nʔd* > *nd* > *nt*	*nʔt* > *nt*
-d- > *-ð̯-* > ∅	*-ʔd-* > *-ʔ-* > ∅	*-ʔt-* > *-ʔ-* > ∅
d- > *ð̯* > (e.g.) *ϑ* > *tϑ* > *ts*	*ʔd-* > *ʔt* > *tʰ* > *tˢ*	*ʔt-* > *tʰ* > *tˢ*

For the present purposes of establishing a workable relative chronology, the disappearance (or weakening) of PIE **d* in clusters with **r*, **w* and **y* can be placed before the assibilation of PIE **t* and **dʰ* before the glide **y*. The latter two both give TAB *tˢ* in combination with **y*, rather than disappearing like **d* (see subsection 2.3.4.2). We can interpret this as either (1) **t* and **dʰ* were from the outset more similar to one another in some way to the exclusion of **d*, so as to be affected by the glide in the same way, or (2) **d* had already disappeared in this environment or changed in some significant way so as to avoid participating in assibilation. The first explanation is difficult to substantiate without positing an otherwise unsupported change like the hypothetical development **dʰ* to **tʰ*, whereby original **t* and new **tʰ* would become united by the feature [-voice] (this is done by Ringe 1996 47–48). Such a development assumes aspiration as a feature of **dʰ* in pre-Proto-Tocharian, even though the evidence in favour of that is very slim (see the next subsection, 2.3.4.5). It is clear, however, that **d* was more liable to disappear than the other two dental stops, and that it changed in some way in several different positions. It is thus possible that **d* had already become *[ð] (or *[θ]) or *[ʔ] or even ∅ in clusters like **dy* before the assibilations of **ty* and **dʰy* took place. This change would therefore be an early one in the relative chronology.

The specific development of PIE **d* in Tocharian thus remains highly uncertain. In combination with the discussion of assibilation in the previous subsection, we can make a provisional notational difference between the two ancestors of TAB *tˢ* as pre-PT *“**ḋ*” on the one hand, from PIE **d*, and pre-PT *“**ċ/ż*ʰ” on the other, from assibilation of PIE **ty* and **dʰy*. The dot in *“**ḋ*” is intended to mark an indeterminate phonetic interpretation, similar to the dot in *“**ċ/ż*ʰ”. On balance, I think that this is a more appropriate representation than any symbol with a more definite phonetic meaning.

2.3.4.5 A Tocharian version of Grassmann's Law

PIE **d* also plays an important role in discussion of another sound change, namely the Tocharian version of Grassmann's Law (Winter 1962: 24–25; Ringe 1996: 47; Hackstein 2017: 1324). Grassmann's Law is a well-known sound change from Greek and Sanskrit, whereby the first in a sequence of two successive aspirates **C^hVC^h* was de-aspirated due to dissimilation, yielding a sequence **CVC^h* instead. It is clear that two different versions of Grassmann's Law happened in Greek and Sanskrit, independently of one another, and it is thus in principle not particularly far-fetched to assume that a similar sound change happened in Tocharian at some point as well. If this type of sound change occurred in early Tocharian, it would imply that the Proto-Indo-European voiced aspirates like **d^h* were indeed phonetically aspirated in this branch as well. This is not obvious from any other Tocharian evidence, so that it cannot be confirmed in any other way than by reference to this Tocharian Grassmann's Law. The small set of possible examples all involve a development from PIE **d^h-* to TAB *t^s-*, as expected for PIE **d-*, rather than to TAB *t-*. De-aspiration by Grassmann's Law of, e.g., **$ǵ^h$* in a root like **"$ǵ^hed^h$-"* to **"$ǵed^h$-"* would not be visible in Tocharian, as both **$ǵ^h$* and **$ǵ$* regularly yield PT voiceless **k*.

16. PIE **d^heg^{wh}-* > **"deg^{wh}-"* > TB *tsək-*, TA *tsäk-* 'burn'
17. PIE **d^heig^h-* > **"$deig^h$-"* > TB *tsəyk-*, TA *tsāykā-* 'form tr.'
19. PIE ?**d^heng^h-* 'reach' > **"$deng^h$-"* > TB *tsənk-* 'rise'

There is also a counterexample against Grassmann's Law, namely PIE **d^hub^hro-* 'deep', which became TA *tpär*, TB *tapre* 'high' (20). With Grassmann's Law, the expected reflex would have been TA ***tspär*, TB ***tsapre*. Ringe does not cite this form as a counterexample, because he reconstructs it as **d^hubro-* instead, with a non-aspirated **b*. The reconstruction of this root was amended by Kroonen; according to him, a verbal derivative from the same root in Germanic, ON *deyfa* 'dip', from **d^houb^h-éie-*, while the adjective ON *djúpr* 'deep', otherwise thought to reflect PIE **b*, has a secondary *p* that arose from PIE **b^h* first through gemination to **pp* before a nasal in **d^heub^h-no-* (Kluge's Law), and was subsequently shortened after the long vowel, i.e. **d^heub^h-no-* > **deuppa-* > **deupa-* 'deep' (Kroonen 2011: 253).

The root for 'deep' is therefore not an example of PIE **b* as previously thought, and an exception to a Tocharian version of Grassmann's Law. On the basis of this, Peyrot casts doubt on Grassmann's Law in Tocharian, suggesting that the **i* might have played a role in the example **d^heig^h-* by causing assibilation in the zero-grade **d^hig^h-* (Peyrot 2022a: 92 fn. 10). Unfortunately, this development is not directly paralleled elsewhere either. It furthermore might require differing developments of **d^hi/ti* word-initially and word-finally, with a potential full assibilation to **-si* and eventual **-ṣ* taking place word-finally (see subsection 2.3.4.2). This is not impossible, but it should be borne in mind. A

development from *d^he to *$t^sə$ in some way due to the *e might seem attractive on the face of it based on of the examples above, but this encounters similar problems: TB *lac*, TA *läc* 'he/she went out' is derived from PIE *h_1lud^het and show a development to TAB *c* instead, in line with palatalized *t (see 2.4 on palatalization).

These issues make it practically impossible to properly decide what actually happened, but it seems clear that the exact phonetic properties of *d^h and the other PIE voiced aspirates in early pre-Proto-Tocharian cannot be established based on such limited evidence (cf. Peyrot 2022a: 91–92). Only PT *$t^sək$- from *d^heg^{wh}- remains as a relatively secure example. If it does truly reflect an early Grassmann's Law that did not operate in PT *$təpre$ from *d^hub^hro, this change should have taken place very early. This is because it appears from other sound changes that PIE *d changed its pronunciation in some way early on as well (cf. subsection 2.3.4.3), which could have untethered it from its connection to *d^h as an unaspirated counterpart in a direct phonetic sense, at least within the traditional framework necessary for the operation of a Grassmann's Law. Any hypothetical Tocharian Grassmann's Law changing *d^h into *d by means of simple de-aspiration should therefore also have taken place at an early stage.[23]

2.3.4.6 Tocharian stop system developments and reconstruction

The general development of the Proto-Indo-European stops in Tocharian is a loss of distinctions with regards to features such as voicing, aspiration, glottalization or length (depending on the phonetics assumed for Proto-Indo-European). The labial and velar series each collapsed into a single stop, yielding /p k k^w/. However, PIE *b^h developed differently from *p in clusters with an *m and disappeared completely. This likely happened via an intermediary stage *[mb] before stops were generally devoiced in Tocharian, although the precise situation preceding *[mb] remains unclear (see 2.3.4.1).

In the dental series, PIE *d developed markedly differently from its counterparts *t and *d^h, as discussed in some detail in 2.3.4.3 and 2.3.4.4. The disappearance of *d in clusters with *r, *w or *y can be understood as a lenition (perhaps via *[ð]) and eventual elision, or as a debuccalization (to intermediary *[ʔ], with a glottalic interpretation of PIE *d). The intermediary steps that have to be assumed depend on the reconstruction of the original, Proto-Indo-European value of *d, and cannot be recovered directly from the Tocharian data. The development from PIE *d to TAB t^s may have happened via an aspirated *[t^h] in some way, but the possible path towards such a *[t^h] from PIE *d, once again, depends on the phonetic interpretation of PIE *d itself. It is unattractive to simply assume *[t^h] was the value of *d in all positions originally (without positing further

[23] Within a glottalic framework, the apparent change in *d^heg^{wh}- is more difficult to interpret, as it would entail an irregular secondary glottalization from *"deg^w-" to *"$d^ʔeg^w$-" (*"tek^w-" to ±*"$t^ʔek^w$-" in the framework of Kloekhorst's Indo-Anatolian stop system).

intermediate steps), as such a phonetic interpretation does not accord with its disappearance in clusters with *r, *w and *y, or its merger with PIE *t and *d^h as TAB t after the nasal *n.

Considering both the discussions surrounding the interpretation of the Proto-Indo-European stop system as a whole and in particular the hypothesis that Tocharian was the second branch to split off, a short note on the relevance of the Tocharian data in this perspective is in order. There are essentially three competing proposals, summarized in Table 2.7. If Tocharian is considered to be closer to either the traditional model or the "classic" glottalic model, there is no additional argument for the Tocharian second hypothesis. If Tocharian is considered to be more similar to Kloekhorst's interpretation of Proto-Anatolian and Proto-Indo-Anatolian, that might support the Tocharian second hypothesis and/or Kloekhorst's reconstruction of the Anatolian system for earlier nodes of Proto-Indo-European.

Table 2.7: Three competing interpretations of the Proto-Indo-European (or Proto-Indo-Anatolian) stops, with the corresponding Tocharian system on the right.

Traditional	Glottalic	Kloekhorst	Tocharian
*t *d *d^h	*t *$ˀd$ *d	*t: *$ˀt$ *t	t t^s t

The traditional system requires a number of intermediate steps to get from the *d to the Tocharian t^s (see 2.3.4.4), and the traditionally assumed aspiration of *d^h is not greatly supported within Tocharian, as it relies on the interpretation of a sound change with only two examples and one counterexample (see 2.3.4.5). Devoicing of *d and both devoicing and de-aspiration of *d^h are required to derive the Tocharian system from the traditional reconstruction of Proto-Indo-European.

The "classic" glottalic interpretation also requires a devoicing of the same two consonants, but they are instead interpreted as *"$ˀd$" and *"d" respectively. Some kind of shift from gottalization to affrication is also needed, perhaps with an intermediate step of aspiration (*[t^h]). Kloekhorst's model requires a degemination from *"t:" to Tocharian single t and a shift from glottalization to affrication, again possibly via aspiration.

In trying to determine the preferable solution, the traditional model is, at least on paper, less economical than the others, unless a Tocharian version of Grassmann's Law is assumed, since that appears to require aspiration of the phoneme *d^h (see 2.3.4.5). The "classic" glottalic model and Kloekhorst's model are on equal footing, since both require two systematic changes: devoicing and affrication in the "classic" glottalic model, and degemination and affrication in Kloekhorst's model. The position of Tocharian closer to either the Anatolian system as reconstructed by Kloekhorst or to the other Indo-

European languages therefore cannot be determined on the basis of the Tocharian stop system without additional evidence.

2.3.5 Summary

Before moving on through Tocharian historical phonology, it may be beneficial to briefly review the changes discussed so far. For reference, the Proto-Indo-European consonant inventory is reproduced again below. This is the system that the earliest stage of pre-Proto-Tocharian inherited, and the starting point for all subsequent developments.

Table 2.8: The Proto-Indo-European consonant inventory.

PIE	labial	alveolar	palatal	palatovelar	velar	labiovelar	"laryngeal"
stop	*p (b) bʰ*	*t d dʰ*		*ḱ ǵ ǵʰ*	*k g gʰ*	*kʷ gʷ gʷʰ*	
							h₁ h₂ h₃
sibilant		*s*					
nasal	*m*	*n*					
liquid		*l r*					
glide	*u̯*		*i̯*				

In the preceding subsections, we have seen some significant changes to this system already. As in most Indo-European languages, the laryngeals disappeared from the consonant system to become integrated into the vowel system instead, creating the additional vowel phonemes **a* (from **H̥* and **h₂e*) and **ā* (from **eh₂*), and generally increasing the frequency of long vowels like **ē* (from **eh₁*) and probably also **ō* (from **oH*) in the language. Perhaps consonantal laryngeals **H* were retained in some positions until a later stage, but after vowel colouring it seems that they cannot be differentiated anymore. Syllabic resonants were vocalized and added another vowel to the system, namely **ə*, which probably became phonemic at the latest when consonantal laryngeals disappeared. The impact of these changes on the vowel system of early pre-Proto-Tocharian is illustrated in Table 2.9 below.

Table 2.9: The Proto-Indo-European vowel inventory with initial developments to the early stages of pre-Proto-Tocharian

PIE	front	back	> pre-PT	front	central	back
high	*i*	*u*		*i ī*	*ə*	*u (ū)*
mid	*e ē*	*o ō*		*e ē*		*o ō*
low					*a ā*	

In the consonant system, the palatovelars as a group became plain velars, and many labiovelars also became plain velars. In contact with another consonant (except perhaps in intervocalic *-*kʷl*-) labiovelars also seems to have been delabialized. Intervocalic **Kṷo* developed to PT **kʷe*, preserving the (originally separate) labial element. Thus, while the palatovelars were eliminated from the consonant system, and the labiovelars suffered some casualties, the latter did survive in a more limited distribution and remained part of the consonant system.

With the assibilation of **t* and **dʰ* before **y* and **i*, an additional consonant phoneme (or phonemes) came into existence, provisionally written as *"*ċ*" (and perhaps also *"*żʰ*"). The assibilation likely happened after PIE **d* in clusters had changed in some way to lose the stop element of its original pronunciation, potentially becoming a fricative like *[ð] in all positions, or by debuccalization to *[ʔ] clusters with a resonant. I use the provisional symbol *"*ḋ*" to mark that some change had occurred to original PIE **d*. The consonants provisionally written as *"*ċ*/*żʰ*" and *"*ḋ*" merged at some unknown point in time, so that their reconstruction as separate entities is hypothetical at this stage of the discussion. Some justification will arise in the discussion of palatalization below (2.4.1). Finally, at some point in time, all (remaining) oppositions of voicing and/or aspiration became neutralized, so that the system ended up containing only voiceless stops. The resulting consonant system is given in the table below. The rewriting of PIE **i̯* and **u̯* as **y* and **w* is merely a notational change. Several word-final consonants and consonant clusters disappeared, but this did not affect the overall system, which is presented here in Table 2.10.

Table 2.10: The pre-Proto-Tocharian consonant inventory after loss of stop contrasts and special developments of PIE **ti̯*/*dʰi̯* and **d*.

pre-PT	labial	dental/alveolar	palatal	velar	labiovelar	"laryngeal"
stop	*p*	*t*		*k*	*kʷ*	
		"*ḋ*"				*H* (?)
affricate		"*ċ*"	~ "*ċ*"			
sibilant		*s*				
nasal	*m*	*n*				
liquid		*l r*				
glide	*w*		*y*			

This consonant system represents the minimal system that must have existed at the point in time that the next great change occurred: palatalization. Palatalization further altered the profile of the Tocharian phonological system, and it is the main topic of the next section.

2.4 Palatalization

A pervasive opposition between palatal and non-palatal consonants is one of the most salient features of synchronic Tocharian phonology. It characterizes the general consonant system and moreover plays an important role in the morphophonology of the language (cf. Krause & Thomas 1960: 61–64). Almost every consonant /C/ had a palatal(ized) counterpart /Cʸ/ in Proto-Tocharian and in both Tocharian A and B. It is important for our view on Tocharian relative chronology to understand how palatalization came about, since it was a pivotal development affecting both the consonant and the vowel systems. In the following subsections we will discuss the process of palatalization and the ways in which it transformed the phonological system of pre-Proto-Tocharian.

2.4.1 Palatalization as it affected the consonant system

Before palatalization took place, the pre-Proto-Tocharian consonant system was probably much like in Table 2.11 below (also given in 2.3.5). Since there are no developments with regards to palatalization that depend on voicing or aspiration, the system can be reconstructed without those distinctions. On this point I differ from Ringe's relative chronology, in which general devoicing and de-aspiration is one of the latest changes (Ringe 1996: 152–154). It has to be admitted that there is no way to be certain (cf. also the considerations in section 2.6), but this is the minimal and most economical system that is required going into palatalization.

Table 2.11: The early pre-Proto-Tocharian consonant inventory; repeat of Table 2.10.

pre-PT	labial	dental/alveolar	palatal	velar	labiovelar	"laryngeal"
stop	*p*	*t*		*k*	*kʷ*	
		"*ḋ*"				*H*?
affricate		"*ć*"				
sibilant		*s*				
nasal	*m*	*n*				
liquid		*l r*				
glide	*w*		*y*			

As a process, palatalization can be understood as first a coarticulation and then transference of frontness (or palatality) from a vowel to a neighbouring (in this case, preceding) consonant. Vowel mergers made the consonantal palatality phonemic. For example, an original opposition between a sequence /Ce/ with a front vowel /e/ and a sequence /Cë/ with a mid or back vowel /ë/ can be transformed into an opposition

between the sequences /C^{y}e/ and /Ce/ instead, with the same phonological vowel, but a contrast in palatality of the consonants. In this way, front-back distinctions in the vowel system were transferred to the consonants, and the consonant system almost doubled in size as a consequence.

Only **r* and **y* did not have a palatal counterpart, the latter for obvious reasons. The absence of **r^{y}* is also common cross-linguistically (Hall 2000). Perhaps pre-PT **ċ*, too, was not palatalized, as we have no reliable examples of original sequences **ty* and **d^{h}y* in environments with a palatalizing front vowel. The only potential instance of this is in the infinitive ending TAB *-t^{s}i*, perhaps from **-tyey* to **-ċey* (assibilation) to **-ċ$^{(y)}$əy* (palatalization?) and to PT **-t^{s}əy*. This does not show any reflex of a **ċ^{y}* as distinct from the development of **ċ* to TAB *t^{s}*. The lack of a distinct palatalized counterpart would make sense if **ċ* was an inherently palatal affricate like *[tɕ] at the time. However, as the infinitive suffix **-d^{h}yey* would be the only available example of a palatalized assibilated dental in any case, caution is advisable. See Ringe (1996: 79) for a discussion on the reconstruction of the infinitive suffix.

An overview of the pre-Proto-Tocharian system is given in Table 2.12 below. Palatalization of **k^{w}* is not distinct from a palatalized **k*, so any potential **k^{wy}* that may have existed at some point had merged with **k^{y}*, or **k^{w}* had already lost its labialization in palatalizing environments early on (cf. 2.3.3.2); **k^{wy}* is not included in Table 2.12. In combination with some of the changes discussed in subsection 2.3.3, this further reduced the number of labiovelars in pre-Proto-Tocharian.

Table 2.12: The pre-Proto-Tocharian consonant inventory after phonemic palatalization.

pre-PT	labial	dental/alveolar	palatal	velar	labiovelar
stop	*p p^{y}*	*t t^{y}*		*k k^{y}*	*k^{w}*
		“ḋ ḋ^{y}”			
affricate		*“ċ”*			
sibilant		*s s^{y}*			
nasal	*m m^{y}*	*n n^{y}*			
liquid		*l l^{y}*			
trill		*r*			
glide	*w w^{y}*		*y*		

From the system obtained after palatalization, the changes to Proto-Tocharian, and also onward to Tocharian A and B, are relatively minor. In general, we see that many palatalized consonants developed into true palatal consonants, although the palatalized sibilant **s^{y}* actually became retracted to a retroflex *[ʂ]. At some point, word-initial **“ḋ”*

(see the discussion in 2.3.4.4) merged with *"*ċ*" to yield PT *t^s, while the palatalized counterpart *"*ḋ^y*" became */ɕ/ and merged with palatalized *k^y.

Table 2.13: The developments of the pre-Proto-Tocharian consonant inventory to Proto-Tocharian.

pre-PT	labial	dental	retroflex	palatal(ized)	velar	labiovelar
stop	*p p^y*	*t*		*t^y*	*k*	*k^w*
affricate		*t^s* < *ċ, ḋ*				
sibilant		*s*	*ʂ* < *s^y*	*ɕ* < *k^y, ḋ^y*		
nasal	*m m^y*	*n*		*n^y*		
lateral		*l*		*l^y*		
trill		*r*				
glide	*w w^y*			*y*		

Finally, with some minor adjustments, the Tocharian A and B systems can be obtained. The palatalized labial consonants were lost as distinctive segments, with *w^y developing into TA *w* and TB *y*. The result of *k^y may have been an affricate still at the Proto-Tocharian stage (Adams 2000: 19–20; Peyrot 2008: 78), but in the attested languages it had become a non-affricate palatal sibilant TAB *ś*. The shared consonant system of both Tocharian A and B is given again in the table below, this time represented in the (transcribed) spelling of both languages.

Table 2.14: The synchronic consonant inventories of Tocharian A and B.

TAB	labial	dental/alveolar	retroflex	palatal	velar	labiovelar
stop	*p*	*t*			*k*	*kw*~*k$_u$*
affricate		*ts*		*c* [cɕ]		
sibilant		*s*	*ṣ* [ʂ]	*ś* [ɕ]		
nasal	*m*	*n*		*ñ* [ɲ]	(*ṅ* [ŋ])	
lateral		*l*		*ly* [ʎ]		
trill		*r*				
glide	*w*			*y*		

Until now the discussion of palatalization has centred around changes in the consonants. In the next subsections, the connected vowel developments will be considered in more detail, with a special focus on the developments of PIE **i*, **e* and **u*.

2.4.2 Developments of PIE **i* and **e*

To understand palatalization in Tocharian, it is important to understand the developments of PIE **i* and especially **e*. These two short front vowels are both reflected in Tocharian as PT **ə*, generally with palatalization on the preceding consonant. In word-initial position, they developed an onglide, becoming PT **yə-*.

27. PIE **gʷenh₂-* > PT **kʸəna* > TB *śana*, TA *śäṃ* 'wife'
34. PIE **h₁itr-* → ?**h₁itōr* > PT **yətar* > TB der. *ytārye*, TA *ytār* 'road, path'
35. PIE **h₁ekʷo-* > PT **yəkwe* > TB *yakwe*, TA *yuk* 'horse'
65. PIE **kʷetʷores* > PT **kʸətwerə* > TB *śtwer*, TA *śtwar* 'four (4)
73. PIE **limn̥-* > PT **lʸəmə* > TB *lyam*, TA *lyäm* 'lake'
81. PIE **-nti* > PT **-nʸtʸə* > TA *-ñc* prs.act.3pl. verbal ending
114. PIE **ueǵʰno-* > PT **wʸəkne* > TB *yakne*, TA *wkäṃ* 'way, manner'

Since the reflexes of PIE **i* and **e* are mostly the same, it is difficult to tell when exactly the two merged. Their behaviour is not entirely identical, however, so that some change must have occurred before they both became PT palatalizing **ə*. The difference is that PIE **i* appears to have had a lower rate of success in palatalizing preceding consonants. This has led to the view that **i* was not as capable of causing palatalization as **e*; e.g., Pinault (2008: 423) writes that it did not palatalize labials, velars, labiovelars and **s*.

It did, however, clearly cause palatalization in some instances (cf. PIE **limn̥-* and **-nti* in the list above), and it would be expected to palatalize more than **e*. In the linguistic typology of palatalization, the high front vowel ±[i] normally has more of a palatalizing effect than a lower front vowel ±[e], with an implicational hierarchy about the vowels that will cause palatalization. According to Bateman, "if lower front vowels trigger palatalization, so do higher front vowels" (Bateman 2011: 597) and "[i] is the best palatalization trigger in in any language, regardless of what the other triggers are" (Bateman 2007: 64). It may therefore be more natural to assume that **i* had been changed into a different vowel in those cases where it failed to palatalize. This is generally the assumption of sequences involving **i* and a labial glide or a labiovelar (see Ringe 1996: 66; Hackstein 2017: 1326).

14. PIE **dṷito-* > "**wuto*" > PT **wəte* > TB *wate*, TA *wät* 'second'
67. PIE **kʷi-so* > "**kuso*" > PT **kʷəse* > TB *kᵤse*, TA *kus* 'who'
107. PIE **u̯iso-* > "**wuso*" > PT **wəse* > TB *wase*, TA *wäs* 'poison'
123. PIE **suidie-* > "**suye-*" > PT **səya* > TB *səya-* 'sweat'

After the changes affecting **kʷi* and *wi*, **e* and **i* could theoretically be merged as palatalizing **i*, although this is not strictly necessary (cf. 4.3.1). Assumed long **ī*

developed differently and was able to palatalize a preceding **w*, as seen in PIE **dṷidḱm̥t* > **dṷih₁ḱm̥t* > **wīkəmt* > PT **wʸəykən* 'twenty' (13).

Other instances of lacking palatalization before original **i* have an intervening laryngeal.[24] The chief example is TB *skiyo* 'shadow', derived from PIE **skHieh₂* (Ringe 1996: 18–19; Del Tomba 2020: 144–145), but there are a few other potential instances as well (see the discussion by Friis 2024: 372–373).

22. PIE **gnh₁-ih₁-* > PT **kənəy-tər* > TA opt.mid.3sg. *knitär* rather than ***kñitär*
96. PIE **sh₂i-nu-* > PT **səyn-* > TB *səyn-* 'satiate' rather than ***ṣəyn-*
99. PIE **skHieh₂* > PT **skəyo* > TB *skiyo* 'shadow' rather than ***ṣiyo*

Ringe suggests that the laryngeals may have become **ə* rather than the usual **a* before a vocalic **i*, which would explain these examples (Ringe 1996: 19). In this way, a form like **sh₂i-nu* would regularly have yielded **səynə-*, whence TB *səyn-*. Since **ə* does not palatalize, as seen from the reflexes of syllabic resonants (2.3.2), this explanation would account for the lack of palatalization in the examples with apparent non-palatalizing **i* without the need for an unnatural development whereby **e* has more palatalizing potential than the higher vowel **i*.

2.4.3 Developments of PIE **u*

In Proto-Tocharian, PIE **u* yielded PT **ə*, just like PIE **e* and **i* discussed in the previous subsection. However, there are a few marked differences in the development of **u* as opposed to the other two short high vowels. Most important is a consistent lack of palatalization before PT **ə* from original **u*. Instead, it may have caused labialization of a preceding velar to **kʷ*, although there are only a few examples, none of which are certain. Such a development could constitute an additional source of PT **kʷ*, besides those labiovelars that were preserved from Proto-Indo-European (cf. 2.4.3). A labial onglide developed from **u-* in word-initial position, yielding **wə-*, similar to how initial **i-* and **e-* became **yə-*.

- ? PIE **g⁽ʰ⁾ul-* > PT **kʷəla-* > TB *kʷəla-*, TA *kʷälā-* 'fail, weaken, recede'
- PIE **udr-io-* > PT **wəriye* > TB der. *wriyeṣṣe* 'pertaining to water, TA *wri** 'watery'
- PIE **h₂eḱu-* > PT der. **akʷətˢe* > TB *akwatse* /akʷətˢe/ 'sharp'

[24] Pinault (2008: 241) has argued that PIE **si* did not undergo palatalization either, and he is followed in this by Hackstein (2017: 1312). The Proto-Indo-European reconstructions and etymologies of these examples are not compelling due to their imperfect semantic derivations (proposed PIE **luksi-* 'white, shining' > TB *laks* 'fish' and **h₃esilo-* 'ash-tree' > TB *esale* , TA *asäl* 'post').

Labialization of velars due to PIE **u* has been posited in preconsonantal position as well. This would provide a straightforward way to account for TB *lak$_u$tse* 'bright, light' from an earlier **luk-* + *-tio-* (Pinault 2008: 426, see also Kim 1999: 164). The difficulty is that there may have been analogical interference from the full-grade of the root, **leuk-*, which regularly became PT **l^{y}uk-*. Additionally, TB *lak$_u$tse* could also be derived from a secondary *u*-stem quasi-PIE **luk-u-ti̯o*, in which case the labiovelar should result from the syncopated sequence **ku* via **k^{w}ə* as in TB *laṅk$_u$tse* from quasi-PIE **h$_1$l̥ng^{wh}u-ti̯o*. Furthermore, no labiovelar appeared in PIE **leug-* → **lug-ro-*, which is posited as the ancestor of TB *lakle* 'sorrow' with assimilation of the **r* to **l* (Kim 1999: 164; perhaps it could rather be a different formation in **-lo* from the same root, i.e., **lug-lo-*). There is no labiovelar reflex, i.e., no ***lak$_u$le* (or ***lkwale*).

It is clear that intervocalic velars were not labialized by a preceding **u*, chiefly on account of TB *tkācer*, TA *ckācar* 'daughter' from PIE **d^{h}ugh$_2$tēr* via pre-PT **tukat^{y}er*. If labialization had occurred, we should find TB ***tkwācer* instead. The labiovelar in TB *sakw*, TA *suk* 'happiness' from Skt. *sukha-* 'happiness', the only example of **u* causing labialization of a following intervocalic velar, may have arisen from loan adaptation strategies rather than from genuine sound law. I think that the evidence of 'daughter' outweighs it (*pace* Kim 1999: 162 incl. fn. 53). Pinault (2008: 426) takes *lak$_u$tse* and TA *lukśanu* as from an unattested root noun **lək^w* (TB **lakw*, TA **luk*), but since the same type of phonological change would be involved, just in word-final position, this does not make much of a practical difference.[25]

21. PIE **d^{h}ugh$_2$tēr* > PT **təkat^{y}er* > TB *tkācer*, TA *ckācar* 'daughter'
71. PIE **leuk-* 'shine' → ?**luk-ti̯o* > PT **lək^{w}tse* > TB *lak$_u$tse* 'bright, light'
74. PIE ?**luglo-* > PT **ləkle* > TB *lakle* 'sorrow'

Since PIE **u* merged with **i* and **e*, and with the **ə* that arose from the vocalization of syllabic resonants, it is not immediately obvious up until which point all of these vowels continued to exist as separate entities. However, PIE **u* did take part in various umlaut phenomena that happened after palatalization (see subsections 2.6.2 and 2.6.3), which means that the existence of PIE **u* as such is assured until a relatively late stage. It therefore needs to be kept separate from **i*, **e*, and **ə* in our interpretation of the phonological system as palatalization took place (cf. subsection 2.5.1).

[25] The development of PIE **uksō-* 'ox' is often cited in this context, since the initial *o-* in TB *okso* has been ascribed to earlier sequence **wək^{w}so*, where the **u* labialized the following velar (Ringe 1996: 127; Kim 1999: 164–165 incl. fn. 54). In my opinion, it is more likely that the change to *o-* was brought about by *o*-umlaut, as discussed in subsection 2.5.2.

2.4.4 Palatalization as it affected the vowel system

As the general Tocharian palatalization created new oppositions between neutral and palatalized consonant phonemes, some oppositions once inherent to the vowel system were lost. To reach the system directly preceding palatalization, it is preferable to reconstruct bottom-up. This is the same approach I took in Warries (2022). Although there are some minor differences in the current presentation, the conclusions remain the same.

2.4.4.1 Syllable types and palatalization

We will start by taking a look at the types of syllables that were found in Proto-Tocharian. Each Proto-Tocharian reconstruction is followed by its Proto-Indo-European source in parentheses. Palatalization is marked in a uniform way with a superscript ^{y}.

- */C^{y}i/ : PT **w^{y}ikən* 'twenty' (13. **dm̥ih₁ḱm̥t*), PT **w^{y}ike* 'place' (116. **u̯eiḱos*)
- */C^{y}e/ : PT **w^{y}ente* 'wind' (53. **h₂u̯eh₁nto-*), PT **m^{y}en^{y}e* 'moon' (77. **meh₁nē-*)
- */Ce/ : PT **keme* 'tooth' (24. **ǵomb^{h}os*), PT **y^{y}ək^{w}e* 'horse' (35. **h₁eḱu̯os*)
- */Ca/ : PT **pat^{y}er* 'father' (91. **ph₂tēr*)
- */Co/ : PT **prot^{y}er* 'bother' (2. **b^{h}reh₂tēr*), PT **wəlo* 'king' (118. **uḷHōnts*)
- */C^{y}u/ : PT **l^{y}uke-* 'light up' (72. **leuko-*)
- */Cu/ : PT **ku* 'dog' (61. **ḱu̯ō*), PT **wu* '2' (15. **du̯oh₁*)
- */C^{y}ə/ : PT **t^{y}ək^{w}e* 'river' (108. **tek^{w}os*), PT **l^{y}əmə* 'lake' (73. **limn̥*)
- */Cə/ : PT **kənte* '100' (60. **ḱm̥tom*), PT **təpre* 'high' (20. **d^{h}ub^{h}ros*)

The syllable types */Ca/ and */Co/ did not originally have palatalized counterparts */C^{y}a/ and */C^{y}o/, as these were only created later due to umlaut (see 2.5.3 and 2.5.5). This means that */Ca/ and */Co/ can be projected to a stage before palatalization simply as such. It should be noted, however, that the situation surrounding these two Proto-Tocharian vowels is not at all straightforward. They will be dealt with in more detail in the next subsection below (2.4.5). PT */C^{y}i/ did not have a non-palatal counterpart and is therefore easy to reconstruct to a stage before palatalization took place. The difference between palatalized */C^{y}e/ and non-palatalized */Ce/ can be reconstructed as an earlier contrast between */Ce/ and */Cë/ respectively. Based on just internal reconstruction */C^{y}u/ and */Cu/ could be rendered as */Cü/ and */Cu/, but we will return to these vowels soon. The schwa */ə/ can also be reconstructed with a front (palatalizing) and back (non-palatalizing) variant, thus "front */Cə/" and "back */Cə/" respectively.

This is generally in line with the known different Proto-Indo-European sources of these syllable types. For example, the distinction between */Ce/ and */Cë/ reflects PIE **Cē* and **Co*, with front and back vowels respectively. The precursors to "front */ə/" and

"back */ə/" are more varied. The palatalizing "front */ə/" derives from PIE **e* and, partially, **i* (cf. subsection 2.5.2), while the non-palatalizing "back */ə/" emerged from original syllabic resonants **R̥* (cf. subsection 2.4.2) as well as from PIE **u* (cf. subsection 2.5.3). These latter two can sometimes be distinguished in Tocharian, since the sequences of the type PIE **Ku-* became PT **kʷə-* with labialization, while for instance **km̥-* developed into PT **kəm-*. This indicates that there used to be at least a three-way contrast, between a palatalizing short vowel*/i/ (from PIE **i* and **e*), a labializing short vowel */u/ (from PIE **u*) and a neutral short vowel */ə/ (originating mainly from the syllabic resonants).

On the basis of the umlaut phenomena involving PIE **u* (2.5.1, 2.5.2), we can conclude that at least */u/ did not merge with */i/ and */ə/ when palatalization first took place. This means that palatalization involved just the merger of the vowels */i/ and */ə/, with the former causing palatalization. */u/ remained as a separate vowel. This in turn indicates that */ə/ can be understood as a back counterpart to */i/, most naturally to be interpreted as ±*[ɨ].[26] The development outlined in Table 2.15 reflects this more detailed understanding of the mergers between the three high short vowels into PT **ə* and their relationships with palatalization.

Table 2.15: Updated intermediary overview of the origins of Proto-Tocharian schwa.

PT	<	<	<	PIE
Cʸə	Cʸə	Cʸi	Ci	Ci, Ce
Cə	Cə	Ci	"Ci"	CR̥
Cə (kʷə)	Cə (kʷə)	Cu	Cu	Cu

These short *i*-vowels in Table 2.15 are not to be confused with the vowel of the syllable type */Cʸi/ in Proto-Tocharian (examples like **wʸike* 'place', cf. above), which mainly derives from the diphthong **ei*. For that sound, a reconstruction of pre-PT */ī/ or */iy/ is possible, differentiating it from the short **i* until that was reduced to **ə*. Similar to PIE **ei* yielding PT palatalizing **i*, PT */Cʸu/, which we provisionally reconstructed as */Cü/ above, goes back to PIE **eu*. A reconstruction of the type */iw/ may thus be adequate before palatalization took place. PT */Cʸu/ existed next to PT */Cu/, which may be rewritten as either */Cū/ or */Cuw/ at the pre-Proto-Tocharian stage before palatalization, parallel to the other high vowel PT */Cʸi/ as pre-PT */Cī/ or */Ciy/.

[26] A phonetically high realization of the **ə* phoneme as ±*[ɨ] instead of a "normal" undefined vowel ±*[ə] could be due to a push from the mid unrounded vowel **ë* that resulted from PIE **o*.

2.4.4.2 The changing system

Based on these considerations, the vowel system right before palatalization occurred must have contained at least the following vowel phonemes: */i e a o u ə ë/. PIE **ei* and **eu* can at this stage be reconstructed as */iy/ and */iw/; the distinction between PIE **i* and **e* is not maintained in Tocharian, so that they may have merged as **i* before palatalization took place; some considerations on the timing of this merger were given above in subsection 2.4.2. With the phonemicization of palatalization, the contrasts between */Ci/ and */Cə/ and between */Ce/ and */Cë/ seem to have been given up as their distinctions were transferred to the preceding consonant. The process is illustrated in the table below with various stages represented.

Table 2.16: An overview of the vowel developments from Proto-Indo-European to Proto-Tocharian with various intermediary stages.

PT	palatalized	pre-PT	post-PIE	PIE
C^{y}i	C^{y}iy	Ciy/Cī	Cei, Cī	Cei, Cih$_1$?
C^{y}e	C^{y}e	Ce	Cē	Cē, Ceh$_1$
Ce	Ce	Cë	Co	Co, Ch$_3$e?
Ca	Ca	Ca	Ca	C$\underset{\circ}{\text{H}}$, Ch$_2$e
Co, Ca?	Co	Co	Cō	Cō, CoH
Co (?)	Cå	Cå	Cā	Ceh$_2$
C^{y}u	C^{y}iw/C^{y}uw	Ciw	Ceu	Ceu
Cu	Cuw	Cuw/Cū	Cū	Cō#, Cuh$_1$?
C^{y}ə	C^{y}i	Ci	Ci, Ce	Ci, Ce
Cə	Ci	"Cɨ"	"Cɨ"	C$\underset{\circ}{\text{R}}$
Cə (k^{w}ə)	Cu	Cu	Cu	Cu

Further changes in the vowel system, leading to the creation of PT */C^{y}o/ and */C^{y}a/, were brought about by a number of umlaut phenomena. These will be discussed in the following section (2.5). Some developments involving PIE **ō* and **eh$_2$/ā* will receive more attention first, as that is a very complex aspect of Tocharian historical phonology relevant for our understanding of the relative chronology of vowel changes.

2.4.5 Developments of PIE **ō* and **eh$_2$/ā*

From a careful study of the table at the end of the previous subsection it will be clear that the development of PIE **ō* is split. The PIE syllable type **Cō* appears in the rows for PT */Co/, */Ca/ and */Cu/. The development of this vowel is a great problem in the relative chronology of Tocharian sound changes and it affects our interpretation of the loss of vowel length. The relation between the developments of **ō* and PIE **eh$_2$/ā* in the vowel

system especially difficult, since *$eh_2/\bar{a}$ yields PT *o as well. The present subsection will further address this situation.

2.4.5.1 Three reflexes of PIE *$\bar{o}$

The vowel phoneme PIE *$\bar{o}$ is considered to be the antecedent of three separate vowels PT *u, *o and *a. At least the reflex PT */Cu/ can be dealt with quite easily, since it concerns a development that was restricted to word-final position. It is to be understood as a relatively early development, preceding the loss of several final consonants and clusters, and it is usually reconstructed as a change from PIE *-$\bar{o}$# to post-PIE *-$\bar{u}$#. This can be seen in the examples listed below.[27]

15. PIE *$d\underset{\cdot}{u}oh_1$ 'two (2)' > *$dw\bar{o}$ > *$dw\bar{u}$ > PT *wu > TAB wu 'two (2)'
61. PIE *$\acute{k}\underset{\cdot}{u}\bar{o}$ 'dog' > *$kw\bar{o}$ > *$kw\bar{u}$ > PT *$k^{(w)}u$ 'dog' > TAB ku 'dog'
50. PIE *$\underset{\cdot}{h_2}ntb^hoh_1$ 'both' > *$antb^h\bar{o}$ > *$antb^h\bar{u}$ > PT *$ampu$- > TA $\bar{a}mpuk$ 'both'
84. PIE *$o\acute{k}t\bar{o}$ 'eight (8)' > *$okt\bar{o}$ > *$okt\bar{u}$ + *k*∂*mt 'ten' > PT *$oktu$- > TA der. $oktuk$ 'eighty'

In some words, PIE *$\bar{o}$ has a Proto-Tocharian word-final reflex *-o as well, without raising to *-u. This occurs in (*$m\bar{o}$)n-stems and nt-stems, indicating that *-$\bar{o}$ was protected by word-final consonants at the time that *-$\bar{o}$# became *-$\bar{u}$#. When the final consonants had disappeared due to apocope, the new final *-$\bar{o}$# did not change much further, only losing its length to become PT *-o (Ringe 1996: 11). The main examples are the following:

59. PIE *$\acute{k}leum\bar{o}n$ > PT *kl^yomo > TB $klyomo$, TA $klyom$ 'noble'
124. PIE *$uks\bar{o}n$- > PT *$okso$ > TB $okso$, TA pl. $opsi$ 'ox'
118. PIE *$\underset{\cdot}{u}\underset{\cdot}{l}Hm\bar{o}nts$ > PT *$w\partial lo$ > TB $walo$, TA $w\ddot{a}l$ 'king'

These words clearly show a development from PIE *$\bar{o}$ to PT *o. A number of other examples are considered to suggest a development from PIE *$\bar{o}$ to PT *a, however. The reconstructions of most of these are problematic in some way. The precise origin of final -ar is difficult and could be analogical in PT *$y\partial tar$ 'road' (Kortlandt 1988b: 84–85). The derivation of PT *$w^y\partial sar$ 'wheat' from Proto-Indo-European is uncertain.

[27] It is also often considered to have affected the PIE thematic ending of the first person singular *-oH > *-$\bar{o}$, which may be the origin of the Proto-Tocharian first-person ending *-u/-w (Ringe 1996: 89). The issue with this reconstruction is that it turns up in verbal classes where we would expect the athematic ending *-m instead, so that this etymology is disputed (but see Malzahn 2010: 39–41). However, alternative explanations for PT *-u/-w usually involve an irregular, unparalleled change from *-m, so that they in turn fail to be particularly convincing.

37. PIE *h_1i-tr- → ?*$h_1$itōr > PT *yətar* 'road'
120. ? PIE *ues-r → ?*u̯esōro- > PT *w^yəsare* > TB *ysāre*, TA *wsār* 'wheat'

An additional potential example is the preterite 3sg. of TB *yok-* 'drink', namely TB *yāś*, which may go back via *$yōg^{wh}$- to a perfect *h_1e-h_1og^{wh}-*e* (K.T. Schmidt 1997: 261; Peyrot 2013: 449). Clearly, this example would be of a different order, as the *$ō$ here should derive from a contraction without any exact parallels. Some very doubtful examples of putative PIE *$ō$ to PT **a* include the following:

23. PT **aknat^{s}a* 'foolish' from PIE *$n̥ǵneh_3nti̯h_2$ (the first and last nasals were lost due to dissimilations) may rather go back to *$n̥ǵnh_3nti̯h_2$ with a zero-grade and a development to PT **a* from a vocalized laryngeal (Friis 2021);
23. The corresponding verbal stem PT **kna-* 'know' from *$ǵneh_3$- 'know' could likewise reflect a zero-grade with a vocalized laryngeal, i.e., *$ǵnh_3$-;
30. PT **kramər* 'heaviness, weight' from PIE "*g^wroh_2-$mr̥$" could also reflect a zero-grade *g^wrh_2-$mr̥$ with vocalization of the laryngeal to **a*. This formation found in Tocharian is not paralleled in any other branch, so that its precise structure is difficult to determine, and a development via *$ō$ from *oh_2 is by no means assured.
54. PT **anse* 'shoulder' from PIE "*$ōmsos$", with a long vowel based on the Greek reflex ὦμος 'id.'. The original form may have been PIE **$h_2$emsos* instead, with PT **a* from PIE **h_2e* in Tocharian (Beekes 2010: 1680; Kloekhorst 2008: 178), or perhaps **$h_3$emsos* (so reconstructed by Kroonen 2013: 25) if **h_3e* could become PT **a* like **h_2e* (cf. 2.3.1);
94. PT **puwar* 'fire' from PIE "*$peh_2uōr$" via **$ph_2uōr$* may reflects PIE **puh_2r* instead (Winter 1965a: 192–193);
· PT **tano* 'grain' from putative PIE "*d^hoHneh_2" should rather be regarded as a borrowing from Iranian (Peyrot 2018b: 257–259), and is thus uninformative.

Due to the great uncertainty surrounding these examples supposedly showing the change from PIE *$ō$ to PT **a*, I remain (see Warries 2022) sceptical about this development. There is a clear continuation of PIE *$ō$ as PT **o* in examples like PT **walo* and PT **kl^{y}omo*. The change from PIE *$ō$ to PT **a* is also surprising on account of the opposite development of earlier *$ā$ (PIE *eh_2) to PT **o*. If both developments turn out to be correct, it seems like the two vowels need to have crossed paths or swapped places in some way. This is assumed by Ringe in his relative chronology. While I do not find the evidence for PIE *$ō$ to PT **a* compelling, it is commonly assumed, and in the end the interpretation rests on a number of reconstructions on which opinions differ. Therefore, in the next subsection, I will show that the assumption of such a change does not have

to affect our relative chronology and the overall interpretation of the phonological system in this dissertation, even if one does believe in it, or if better evidence is found. This would entail that after original word-final *$\bar{o}$ had become *$\bar{u}$, new word-final *$\bar{o}$ again developed separate from non-final *$\bar{o}$, with the latter eventually becoming *a.

2.4.5.2 PIE *$\bar{o}$ to PT *a and the other vowels

To understand the putative development from PIE *$\bar{o}$ to PT *a within the pre-Proto-Tocharian vowel system, Ringe tentatively reconstructs an intermediate pre-PT long *$\ddot{\bar{e}}$. This would parallel the development of short PIE *o to pre-PT *$\ddot{e}$, and affect the relative chronology of long vowel changes. The hypothetical pre-PT long *$\ddot{\bar{e}}$ could then merge with older *a (from laryngeals) after the long *$\bar{a}$ had become a rounded vowel like *$\bar{\mathring{a}}$ on its way to becoming PT *o. Ringe dates this in the relative chronology before the loss of vowel length (Ringe 1996: 93). This idea is summarized in Table 2.17.

Table 2.17: Potential development of PIE *$\bar{o}$ and adjacent vowels according to Ringe (1996), with a merger of *$\bar{o}$ and *a upon the loss of vowel length.

PT	<	<	PIE
e	ë	ë	o
a	ā	$\ddot{\bar{e}}$	ō
a	a	a	a
å	$\bar{\mathring{a}}$	$\bar{\mathring{a}}$	ā

While such a scenario is theoretically possible, I think that it is difficult in practice, since long *$\ddot{\bar{e}}$ then needs to merge with short *a without interfering with short *$\ddot{e}$. The latter, in turn, would have to merge with original long *$\bar{e}$. These mergers are most understandable if vowel length did not play a role anymore at the time. On the other hand, vowel length would still be required to separate *$\ddot{e}$/o from *$\ddot{\bar{e}}$/$\bar{o}$ in Ringe's model. We may therefore see how a similar sequence of changes can work perfectly well without a retention of vowel length.

When short PIE *o changed to *$\ddot{e}$, long PIE *$\bar{o}$ can take its place as the new short *o. Similarly, if short PIE *e had merged with PIE *i, we can assume that long PIE *$\bar{e}$ became a new pre-PT *e. At this stage, vowel length has ceased to be distinctive without any problems. After palatalization had occurred, early pre-PT *$\ddot{e}$ and *e merged as *e, with the latter having caused palatalization of the preceding consonant (i.e., */Cë/ and */Ce/ became */Ce/ and */C^{y}e/). This then would create the possibility for non-final *o to become a new *$\ddot{e}$, and from that position it can have been lowered and merged with *a. If we reconstruct an intermediary stage *$\mathring{a}$ (±*[ɒ] or ±*[ɔ] as opposed to ±*[o]) for

original long PIE *$eh_2/\bar{a}$, that vowel is then continuously differentiated from PIE *$\bar{o}$ in its development via short *o and *$\ddot{e}$ to eventual *e. This is summarized in Table 2.18.

Table 2.18: Potential developments of *$\bar{o}$ and adjacent vowels with a merger of *$\bar{o}$ and *a after loss of vowel length.

PT	<	<	<	PIE
e	e	e	e	ē
e	e	e	ë	o
a	ë	o	o	ō
a	a	a	a	a
å	å	å	å	ā

This is the same type of movement around the vowel space as Ringe proposes with vowel length intact, from PIE *$\bar{o}$ to intermediate *$\bar{\ddot{e}}$, and from there down to *a, but it turns out that vowel length is not necessary. However, vowel length cannot be lost before *$\ddot{e}$ merges with *$\bar{e}$ to avoid merging a of *$\ddot{e}$ and *$\bar{\ddot{e}}$, while the merger of *$\ddot{e}$ and *$\bar{e}$ depends on a neutralization of vowel length.

It would thus in any case be preferable to keep long PIE *$\bar{o}$ as *o initially, until short PIE *o has developed through *$\ddot{e}$ and with palatalization to eventual *e, merging with *e from *$\bar{e}$. To avoid a merger of long PIE *$\bar{o}$ and *$\bar{a}$ as *o, the *$\bar{a}$ can be considered as still lower *$\mathring{a}$ (*[ɒ] or *[ɔ]). This vowel only later merged as PT *o with the PIE long *$\bar{o}$ that remained in (secondarily) word-final position (type TB *walo* 'king'). Some additional discussion of the PT *o will be given in parts of the next subsection, 2.5.

2.4.6 Summary

Palatalization was one of the most far-reaching changes that took place in pre-Proto-Tocharian, with significant ramifications for the phonological system, as discussed in this subsection. Most consonant phonemes were split into a palatalized and a non-palatalized version. The schematic summary of the changes that took place is reproduced here in Table 2.19.

Table 2.19: An overview of the vowel developments with palatalization from Proto-Indo-European to Proto-Tocharian.

PT	palatalized	pre-PT	post-PIE	PIE
C^{y}i	C^{y}iy	Ciy/Cī	Cei, Cī	Cei, Cih
C^{y}e	C^{y}e	Ce	Cē	Cē, Ceh$_1$
Ce	Ce	Cë	Co	Co
Ca	Ca	Ca	Ca	CH, Ch$_2$e

PT	palatalized	pre-PT	post-PIE	PIE
Ca, Co#	Co	Co	Cō	Cō, CoH
Co	Cå	Cå	Cā	Ceh$_2$
C^{y}u	C^{y}iw/C^{y}uw	Ciw	Ceu	Ceu
Cu	Cuw	Cuw/Cū	Cū	Cō#, Cuh$_1$?
C^{y}ə	C^{y}i	Ci	Ci, Ce	Ci, Ce
Cə	Ci	“Cɨ”	“Cɨ”	C$\underset{\circ}{\text{R}}$
Cə	Cu	Cu	Cu	Cu

The developments of the vowel system can be summarized as first a reduction in the number of long vowels, hand in hand with an increase in the number of vowel qualities. Since palatalization involved a merger of long PIE *$\bar{e}$ and short PIE *o, the length distinction of the former must have been nullified at the time of palatalization. Marginal *$\bar{\imath}$ (from *ih_1) and *$\bar{u}$ (from word-final *-$\bar{o}$) did not merge with their short counterparts. Since short *i, *e and *u remained distinct from one another until after the occurrence of certain types of umlaut, the original of length distinction must have been preserved among the high vowels in some form. Since *$\bar{\imath}$ and *$\bar{u}$ did coalesce with the diphthongs *ei and *eu, they could be considered as *iy and *uw instead.

Table 2.20: The vowel inventories of two stages of pre-Proto-Tocharian, before the loss of phonemic vowel length on the left and after the loss of phonemic vowel length on the right.

pre-PT	front	central	back	>	front	central	back
high	*i ī*	“*ɨ*”	*u ū*		*i iy*	“*ɨ*”	*u uw*
mid	*e ē*		*o ō*		*e*	*ë*	*o*
low		*a ā*				*a*	*å*

With palatalization, short *i (having subsumed PIE *e and most PIE *i) and *$ə$ (derived from vocalized syllabic resonants and some PIE *i after labiovelars and labials) merged as pre-PT *i, with only the former causing palatalization. Similarly, *e from PIE *$\bar{e}$ and *$ë$ from PIE *o merged as pre-PT *e, again with only the former causing palatalization.

Table 2.21: The vowel inventories of two stages of pre-Proto-Tocharian, before phonemic palatalization on the left and after phonemic palatalization on the right.

pre-PT	front	central	back	>	front	central	back
high	*i iy*	“*ɨ*”	*u uw*		*i iy*		*u uw*
mid	*e*	*ë*	*o*		*e*		*o*
low		*a*	*å*			*a*	*å*

The remaining changes needed to derive the Proto-Tocharian vowel system are relatively simple. Short **i* and **u* finally merged as PT **ə* so that **iy* and **uw* can now be rewritten as **i* and **u* (alternatively, **əy* and **əw*). These developments are illustrated in Table 2.22.

Table 2.22: An overview of the vowel inventories of a late stage of pre-Proto-Tocharian on the left, with Proto-Tocharian on the right, after the final changes had taken place.

pre-PT	front	central	back	> PT	front	central	back
high	*i iy*		*u uw*		*i*	*ə*	*u*
mid	*e*		*o*		*e*		*o*
low		*a*	*å*			*a*	(*å?*)

It is sometimes assumed that the merger of PIE **ē* and **o* did not actually take place until after Proto-Tocharian (Ringe 1990: 222–226; Adams 2013; Kümmel 2009). At various points in the next section (2.5), it will become clear why it is not necessary to keep these vowels separate until palatalization. Similarly, (pre-)PT **o* from post-PIE **ō* and (pre-)PT **o* from PIE **eh₂* (=**ā*) are often kept separate as well as PT **o* and **å*, a view which I also held in an earlier article (Warries 2022). I now think it possible that these two vowels were merged in pre-Proto-Tocharian already as **o* (cf. the table above). This will also be discussed further in the next section, specifically in 2.5.6.4.

2.5 Late vowel developments

Some types of syllables that existed in Proto-Tocharian were left out of consideration in the previous section about palatalization because they are the result of a different phenomenon: umlaut. This most importantly concerns the syllables of the type TB /Cʸo/ and /Cʸa/, both of which are derived from earlier */Cʸe/ due to effects of the vowel in the following syllable. It is important to discuss these changes, as they inform a number of decisions made about the interpretation of the relative chronology in the previous section, and there is no consensus.

2.5.1 *u*-umlaut of pre-PT **e*

It has been known for a long time that an original **u* caused umlaut of an **e* in the preceding syllable. Some scholars use a special symbol for the product of this umlaut, namely "*o̰*" with a dot underneath. This reflects the view that "the new *o*-vowel did not merge with other *o*'s in PT" (Ringe: 1996: 98). I write the potentially different *o*-vowels as "*o*" and "*å*" rather than "*o̰*" and "*o*", and the issue will be discussed throughout this section.

Furthermore, while Ringe only discusses *u*-umlaut of the **ë* that arose from PIE **o*, also (post-)PIE **ē* underwent the same development (cf. Hilmarsson 1986: 23–28; Del

Tomba 2020: 88, 91). We can thus simply say that it was the **e* that arose from the merger of **ë* (PIE **o*) and **e* (PIE **ē*) after palatalization that was targeted by *u*-umlaut. Representative examples are listed below.

10. PIE **dor-u-*, **dre-u-* 'wood' >> **eru* > PT **orə* > TA *or*, TB *or*
82. PIE **nokʷtu-* > **nektu-* > **nokt-* > TA obl. *nokte* 'last evening', der. *noktiṃ* 'in the evening'
84. PIE **oktō* 'eight (8)' > **ektū* > PT **oktə* (analogical **-ə* after **sʸəptə* 'seven (7)') > TB *okt*; TA *okät* (earlier *oktu-* visible in TA *oktuk* 'eighty' from PT **oktu-kə*)
100. PIE **sneh₁-ru* 'sinew' > **snēru* > **sʸnʸeru*[28] > PT **ṣñorə* > TB *ṣñor*
102. PIE **solh₂u-* > **selu+me*[29] > PT **solme* > TB *solme* 'complete(ly), altogether'

Hilmarsson (1986: 20-21) dated *u*-umlaut after the Proto-Tocharian period, but this cannot be correct, since most **u* that caused this sound change had disappeared by that point in the merger with original **i* and **e* as **ə* (cf. also 2.7). The *u*-umlaut of **e* indicates both that PIE **o* and **ē* were already phonologically merged when the umlaut took place, after palatalization, and conversely, that PIE **u* and **i*/*e* had not yet merged at that point in time.

2.5.2 *o*-umlaut of pre-PT **u*

Original PIE **u* and **eu* seems to have been turned into PT **o* when followed by an **o* in the next syllable (Hilmarsson 1986: 38–42). There are only two examples.

59. PIE **ḱleumōn-* > ?**klʸuwmo* > PT **klʸomo* > TB *klyomo*, TA *klyom* 'noble'
124. PIE **uksōn-* > *ukso* > PT **okso* > TB *okso*, TA pl. *opsi* 'ox'[30]

This development of PIE **u* is best seen in PIE **uksōn* to PT **okso* 'ox' (TB *okso*, TA pl. *opsi*). PIE **ḱleumōn* to PT **klʸomo* 'noble' (TB *klyomo*, TA *klyom*) shows a similar development to that in **ukso*, although the vowels in initial syllables of these words are

[28] With metathesis from **-ur* to **-ru*, see Del Tomba (2020: 87–94).

[29] TA *salu* 'complete(ly), altogether' is likely from *solh₂wV* with consonantal **u* as **w*, which did not cause umlaut.

[30] The vowel in **okso* is sometimes instead taken to be due to rounding between to labials with a reconstruction of pre-PT **wəkʷso* (Ringe 1996: 127; Kim 1999: 164–165 incl. fn. 54), but on account of the *o* in TA pl. *opsi*, this particular sound change cannot have been the reason for the TB *o* here (Pinault 2008: 432–433). This is because all the other examples of this sound change, namely TB *kokale* : TA *kukäl* 'chariot' from **kʷəkʷle* and TB *pokkāka* : TA *pᵤkāk* 'call!' from **pəkʷaka* show that no development to TA *o* took place in this context (cf. 2.5.6.3). Umlaut is thus the best way to account for the vocalism in TB *okso* and TA *opsi*.

not identical. The fact that we do not see this change in PIE *$\underset{\circ}{\underset{\frown}{u}}\underset{\circ}{l}h_2\bar{o}nts$, which became PT *$wəlo$ 'king' (TB *walo*, TA *wäl*), indicates that the umlaut took place before initial *u- developed into *$wə$- in pre-PT *$ukso$. Otherwise, we would have had a development from PIE *$\underset{\circ}{\underset{\frown}{u}}\underset{\circ}{l}h_2\bar{o}nts$ to PT *$wolo$ as well, to yield TB **olo, TA **wal. This development in its entirety can thus be dated to a time before short *u merged with *i and *$ə$ as PT *$ə$, a change in which the rise of initial *$wə$- from *u- was probably implicated. If the first *o in PT *kl^yomo is the result of the same development, it must have taken place after palatalization, just like the other umlaut developments.

2.5.3 *o*-umlaut of pre-PT *e

The vowel *o from earlier PIE *eh_2 also appears to have caused a preceding *e to turn into *o (Hilmarsson 1986: 29ff; Ringe 1996: 163; Pinault 2008: 432). Reflexes of this umlaut are only clearly visible in Tocharian B, where there are a few examples.

89. PIE *$proskeh_2$ > *$preskå$ > PT *$prosko$ (*$pråskå$) > TB *prosko~proskiye*, ? TA *praski* 'fear'
95. PIE *seh_2i- → prt.ptc. pre-PT *se-$soyəwə$ (*se-$såyəwə$) > TB *sosoyu*, ? TA *sasyu* 'satisfied'
97. PIE *sem- → som-eh_2m > *$semåm$ > PT *$somo$ (*$såmå$) > TB *somo*, TA *ṣom* 'one (1), some' (f.obl.sg.)

Pinault (2008: 436–438) writes that "il n'existe pas de correspondence sûre avec le tokh. A qui garandisse un Umlaut de ce type au stade tokh. commun" (437). This is because the clearest examples in Tocharian B do not have straightforward counterparts in Tocharian A. Each example listed above has its own problems.

TA *praski* 'fear' is superficially similar to the secondary Tocharian B nominative *proskiye*, but since this is a later creation the two cannot go back to Proto-Tocharian as such. TA *praski*, with its alternant plural *praskintu* shows different morphological behaviour from TB *prosko~proskiye*, leading Peyrot (2008: 103) to conclude that the two are most likely independent nominal formations. It is thus uncertain if TA *praski* shows that the *o*-umlaut visible in TB *prosko~proskiye* was restricted to Tocharian B only.

It seems like the counterpart to the preterite participle TB *sosoyu* to *soy*- 'be sated' is TA *sasyu*, with *a*-vocalism rather than *o*-vocalism (Pinault 2008: 436; Ringe 1996: 163; Malzahn 2010: 955). However, Peyrot has pointed out that TA *sasyu* is synchronically the participle to a different verb TA *säy*(*n*)- 'satiate; be satiated', which is the cognate of TB *səyn*- 'be satiated'. The cognate to the verb TB *soy*- 'be sated', from which TB *sosoyu* is formed, is not attested in Tocharian A. The preterite participle TA *sasyu* is parallel to the preterite participle TA *raryu* from the verb *räy*- 'give up', cognate to TB *rəyn*- 'id.' (Peyrot 2013: 833 fn. 949). This means that TB *sosoyu* and TA *sasyu* could reflect two different

formations, so that TA *sasyu* cannot give us reliable information on the status of *o*-umlaut caused by the reflex of PIE *$*eh_2$*/*ā* in Proto-Tocharian.

The feminine of the numeral 'one (1), some', which shows umlaut caused by the reflex of *$*eh_2$* in TB *somo*, pl. *somona*, is analogically modified in Tocharian A *ṣom*, pl. *ṣoman*, by at least a change of the initial sibilant from PT **s-* to TA *ṣ-* on the basis of the masculine TA *ṣom*, pl. *ṣome* (TB *ṣeme*, *ṣemi*). It is therefore uncertain whether the analogical change affected just the initial segment of an umlauted pre-TA **som*, pl. **soman*, or if the analogical modification affected the first two segments of an earlier pre-TA **sam*, pl. **saman* without umlaut (see also Pinault 2008: 437–438). Affection caused by the labial *m* may also have caused a non-umlauted PT **e* to become TA *o*, as in the development from *$*n^yemə$* to TA *ñom* 'name', beside TB *ñem* 'id.' (cf. 2.5.6.1).

There is one example that appears to show *o*-umlaut of **e* caused by an earlier PIE **ō*, namely TB *orkamo*, TA *orkäm* 'darkness; dark', via pre-PT **erkmo*.[31]

42. PIE *$*h_1reg^w$- → *$*h_1rg^w$-*mōn* > **erkmo* > PT *orkmo* > TB *orkamo*, TA *orkäm* 'darkness; dark'

This example would also show that the umlaut was shared between Tocharian A and B, since both have the same vowel *o*. However, it could alternatively reflect a secondary derivation from a *u*-stem adjective *$*h_1rg^{(w)}$-*u*-, thus *$*h_1rg^{(w)}$-*u*-*mōn*, where *u*-umlaut would have caused the change from **e* to **o* instead: *$*h_1rg^{(w)}$-*u*-*mōn* > **erkumo* > **orkumo* > **orkəmo*. Such a reconstruction circumvents the need to posit *o*-umlaut of **e* caused by PIE **ō* in the basis of this single example, but we would expect a labializing effect on the velar caused by **u*, thus TB ***orkwamo* vel sim. (cf. TB *laṅk$_u$tse* in 2.3.3.3).

It is conceivable that PIE *$*eh_2$*/*ā* and PIE **ō* were still different when *o*-umlaut took place, and that they caused changes to different vowels, one with the reflexes TB *o* (TB *prosko*, *sosoyu*), TA *a* (*praski*, *sasyu*), and the other with TAB *o* (TB *orkamo* and TA *orkäm*). With so few examples and the complicated interpretation of TA *praski* and *sasyu* it seems impossible to further substantiate this possibility.

In conclusion, the available examples in Tocharian A do not allow us to determine with certainty whether *o*-umlaut caused by the reflex of original *$*eh_2$*/*ā* already took place in Proto-Tocharian, or if it only occurred in Tocharian B. It is furthermore possible that the resulting vowel of this umlaut was not the PT **o* that became TAB *o*, but rather PT

[31] Another similar case is that of TB *kolmo*/*a**, TA *koläm* 'basin', which has long been interpreted as the result of *o*-umlaut from pre-PT **kelmo* (Pinault 2008 432; Del Tomba 2020: 146 fn. 218; Hilmarsson 1986: 29). Huard has recently argued that the vowel should be the result of *u*-umlaut of **e* instead (Huard 2022: 411-416). This word is best excluded from the evidence due to the uncertainties involved.

**å*, which is treated further in the next subsection. Whether original PIE **ō* caused *o*-umlaut in (pre-)Proto-Tocharian, and whether the resulting vowel was different from the *o*-umlaut was different from the vowel obtained from *o*-umlaut caused by original PIE **eh₂/ā* cannot conclusively be determined.

2.5.4 *o*-umlaut of pre-PT **a*

Especially Tocharian B shows additional cases of umlaut where the *o*-vowel that caused the umlaut arose secondarily due to contractions (e.g., Pinault 2008: 433–436). It is thus etymologically different from both PIE **ō* and PIE **eh₂/ā*. The vowel that resulted from this umlaut must also have been different from PT **o* in phonological terms, since it always became TA *a* rather than ***o* (i.e., not TA ***ora-* 'cease' prs.[4], etc., from a PT ***oro-*). The vowel sign **å* may be used to express this. The development here is thus separate from the PT **o* from at least the products of *u*-umlaut of **e* (PIE **o* and **ē*), *o*-umlaut of **(e)u*.[32]

While the precise development of the second-syllable **-o* in the present class [4] is uncertain (Malzahn 2010: 385–389, Peyrot 2013: 592–593), it is clear that the vowel it created by umlaut was different from the vowel belonging with the two other cases of *o*-umlaut discussed above in subsections 2.5.2 and 2.5.3. The Tocharian A forms cannot reflect the original, non-umlauted vowel, since that should have become TA ***ā* rather than *a*. The examples are given in the list below.

49. TB *onolme* 'human' from **ånålme*, earlier **anålme* contracted from **ana-elme*; no TA cognate
- TB *oro-*, TA *ara-* 'cease' prs.[4] from **årå-*, earlier **arå-*
- TB *ortto-*, TA *arta-* 'praise' prs.[4] from **årttå-*, earlier **arttå-*
- TB *oso-*, TA *asa-* 'dry out' prs.[4] from **åså-*, earlier **aså-*
- TB *yoto-*, TA *yata-* 'be capable of' prs.[4] from **yåtå-*, earlier **yatå-*
- TB *plonto-*, TA *planta-* 'be glad' prs.[4] from **plåntå-*, earlier **plantå-*
- TB *kloyo-*, TA *klawa-* 'fall' prs.[4] from **klåwʸå-*, earlier **klawʸå-*

[32] Pinault (2008: 435–436) considers the word for 'elephant', TB *oṅkolmo* and TA *oṅkaläm* to have undergone the type of umlaut under discussion here, from an earlier **ankålme* (with a secondary inflection in *-o* after *okso* 'ox, cow'). This would show a Tocharian A reflex *o* in the first syllable, counter to the verbs of present [4]. Since the etymology of 'elephant' in Tocharian remains a thorny problem, and analogical influences from both the ancestor of TB *onolme* 'man' and of TB *okso* 'ox, cow' are assumed, this word is best not used for determining regular sound change. See Adams (2013: 118–119) for an overview of other etymological suggestions of TB *oṅkolmo*, TA *oṅkaläm*, with references.

The correspondences in these verbs crucially show that a vowel **å* must be reconstructed for Proto-Tocharian. The relation of this vowel to the result from PIE **eh₂/ā* is not entirely clear, see the discussion below in 2.5.6.4.

2.5.5 *a*-umlaut of pre-PT **e*

The final type of umlaut that we will discuss is the *a*-umlaut of pre-PT **e*. Just like the *u*-umlaut of **e* (2.5.1), this *a*-umlaut affected both **e* from PIE **o* and **e* from PIE **ē*, and it added the syllable type */Cʸa/ to the system. Its effects are mainly seen in verbal morphophonology in Tocharian B. Winter first published about *a*-umlaut of pre-PT **e* in a short aside, connecting the ablaut grades of *a*-preterites [class 1] or *a*-subjunctives [class 5] with their respective root counterparts, and offering the plural noun TB *ārwa* to *or* 'wood' from earlier **erwa* as a parallel outside of the verbal system (Winter 1962: 32–33). The *a*-umlaut allows for an interpretation of, e.g., 3sg. *tārkaṃ* /tárkan/ 's/he dismisses, emits' and 3sg. *tekäṃ* /tekən/ 's/he touches' as having the same type of ablaut grade, namely an original **e*. This original **e* was then changed to **a* in the former, due to the presence of a root-final *-a*. In the third person plural forms, such as *tarkaṃ* /tərkan/ 'they dismiss, emit' and *takäṃ* /təkən/ 'they touch', both verbs have the same *ə*-grade, which remained unaffected by umlaut (e.g., Pinault 2008: 594–596; Peyrot 2013: 69).

Table 2.23: The effects of *a*-umlaut in the *a*-subjunctive (*tārkaṃ*) next to an example from the root-subjunctive without *a*-umlaut (*tekäṃ*).

	ablaut	example	reconstruction
no final *-a*	*e* : *ə*	/tékən/ : /təkən/	
final *-a*	*a* : *ə*	/tárkan/ : /tərkan/	< */térkan/ : /tərkan/

This situation is not found in Tocharian A as such. The form corresponding to TB *tārkaṃ* is TA *tarkaṣ*, with initial TA *a* reflecting a PT **e*, rather than ***tārkaṣ* with TA *ā* from PT **a* due to umlaut. In general, the Tocharian A subjunctives [5] with stem-final **-a* do not show the *a*-umlaut seen in their Tocharian B counterparts. According to Cowgill (1967: 176–177), who is followed by Ringe (1996: 160–163), Pinault (2008: 428), Hackstein (2017: 1320), etc., this is because *a*-umlaut did not happen in stressed syllables in Proto-Tocharian. The *a*-umlaut of stressed **e* only took place in Tocharian B, but in Tocharian A *a*-umlaut was restricted to unstressed syllables, from the point of view of Tocharian B stress. Cowgill's parallel outside the verbal system is the feminine form for 'four (4)' in TA *śtwar*, rather than ***śtwār* from **śtwéra*; cf. TB *śtwāra* (Cowgill 1967: 177).

The reduplication syllable in preterite participles and the negative prefix do show *a*-umlaut in Tocharian A as well as B, so that an earlier **nenákuwə* from the root **nak-* 'reproach' has become PT **nanákuwə* and thence TB *nanāku* and TA *nānku*. Similar are

pre-PT **kekámawə* from **kama-* 'bring' to PT **kakámawə* yielding TB *kakāmau* and TA *kākmu*, and pre-PT **enáywatˢe* 'unpleasant' to PT **anáywatˢe* yielding TB *anaiwatse* and TA *ānewāts* (Ringe 1996: 160). Examples such as these are the reason to assume that *a*-umlaut also occurred in Tocharian A, but that its operation was limited to unaccented syllables (Cowgill 1967). This is summarized in Table 2.24.

Table 2.24: An illustration of a-umlaut affecting unstressed syllables.

pre-PT	*a*-umlaut	TB	TA
**nenákuwə*	*nanákuwə*	*nanāku* /nanákəw/	*nānku*
**kekámawə*	*kakámawə*	*kakāmau* /kakámaw/	*kākmu*
**enáywatˢtˢe*	*anáywatˢtˢe*	*anaiwatstse* /anáywatˢtˢe/	*ānewāts*

However, there are also some instances where Tocharian A shows traces of *a*-umlaut in syllables that are accented in Tocharian B. Jasanoff adduces the *a*-subjunctive of the stem **mənta-* 'hurt', which has a non-ablauting subjunctive showing umlaut from **ménta-* to TA *mānta*, as seen in TA *māntlune* (Jasanoff 2013: 116 fn. 37). Similarly, the sbj.3pl. TA *pālantär* from TA *pälā-* 'praise' would have been TA ***palantär* without umlaut from earlier **pélantər*.[33] These examples support the idea that *a*-umlaut of also stressed **e* actually did apply in the ancestor of Tocharian A as well, but was partly undone. Jasanoff argues that any Tocharian A *a*-subjunctives that seem to reflect **e*-grade could have been remade on the basis of the productive ablaut pattern of TA *a* : *ä* (type PT **e* : **ə*), which is seen in many subjunctives without root-final *-a* and could plausibly have replaced the ablaut of TA *ā* : *ä* that resulted from *a*-umlaut (Jasanoff 2013: 107 fn. 7).

Ringe instead suggests that the **ā* in the root in examples such as these arose in the preterite, where the accent was on the stem-final **-a* instead (Ringe 1996: 161). However, of *a*-umlaut did not only occur in verbal forms; its effects are also seen in a few nouns. The change of pre-PT **erwa* (plural of *or* 'wood') to PT **arwa*, whence TB *ārwa* /árwa/, has already been mentioned above, and it constitutes a case of *a*-umlaut of stressed pre-PT **é* in Tocharian B. The Tocharian A counterpart to this particular word is not attested, but Pinault has convincingly argued that the same plural suffix **-a* did cause *a*-umlaut in TA *wmār* 'jewel'. The cognate in Tocharian B is *wamer* 'id.' The discrepancy in the vowel TA *ā* : TB *e* (expected would be TA *a* : TB *e* or TA *ā* : TB *a*) can be understood if we consider the kind of plural that TB *wamer* forms, namely *wmera* /wméra/. If we start from such a form in pre-PT **wəméra* and apply *a*-umlaut, this should become PT **wəmára*. Such a PT **wəmára* is then the regular source of TA *wmār*, which could be reinterpreted as a

[33] TB *pālāntär* from THT 562 a2 is uninformative, and the 1sg. *pālamar* in THT 240 b6 (arch.) is not reliable, but the subjunctive *pāloymar* PK AS 5A a5 points to initial accent in this subjunctive stem.

singular when the ending was lost due to apocope (Pinault 2011: 160–164 and 171–173). The TB plural *wmera* for expected **wmāra* has been analogically changed to better fit the singular TB *wamer* again. The only way to continue the belief that accented **é* specifically was not affected is to say that the Tocharian B accentuation is innovative, and that the accent was somewhere else at the time of *a*-umlaut. Such an approach inevitably leads to a circular argumentation, since the connection between *a*-umlaut and the accent is based on the place of the accent in Tocharian B in the first place.

The lack of *a*-umlaut in TA feminine *śtwar* 'four (4)' rather than **śtwār* can probably be attributed to analogical influence from the masculine *śtwar*. The numeral for 'one (1)' also merged in Tocharian A due to analogical influence from the masculine form *ṣom* (PT **sʸeme*) on the feminine TA *ṣom*, with the initial retroflex sibilant taken over from the masculine. As a process, *a*-umlaut was not active anymore in either Tocharian language, and its results were clearly prone to being analogically remodelled. It can thus be assigned to the pre-Proto-Tocharian period, and, contrary to common belief, applied to both stressed and unstressed pre-PT **e*.

The examples of *a*-umlaut that we have seen so far affected pre-PT non-palatalizing **e*, which derives from PIE **o*. Yet there are also a few examples of palatalizing TAB *a*, and these can be derived from PIE **ē* by way of *a*-umlaut. These examples are restricted to the preterite of a small group of verbs, namely *lyaka-* to *lək-* 'see', *śawa-* to *śəw-* 'eat', *plyawa-* to *pləw-* 'complain', and *lyawa-* to *ləw-* 'rub'. The relevant forms derive via *a*-umlaut from earlier **lʸeka-*, **kʸewa-*, **plʸewa-*, and **lʸewa-* respectively. The origin of the *ē*-grade in these verbal forms has long been a contentious issue in the Tocharian scholarly literature, see for instance Peyrot (2012b: 102–109, with further references). However, the clear effects of *a*-umlaut seen here imply that PIE **o* and **ē* had phonologically merged before *a*-umlaut, with the two originally distinct vowels both being affected in the same way. This aspect of *a*-umlaut agrees with *u*-umlaut (2.5.1). The fact that original PIE **ē* did palatalize the preceding consonant before becoming **a*, in turn, confirms that *a*-umlaut, again like *u*-umlaut, took place after palatalization.

2.5.6 Effects of labial consonants on vowels

Aside from the various types of umlaut discussed in the preceding subsections, some effects of labial consonants on neighbouring vowels needs to be taken into account as well. These effects of labial consonants largely occurred in the separate histories of Tocharian A and B, and led to the creation of additional *o*-vowels. It is mostly Tocharian A that was affected, but some instances of are known from Tocharian B as well. The conditioning environments are different, however, and not always entirely clear. Apparent irregularities have been interpreted as the result of a retained contrast between PIE **ē* and **o* until after Proto-Tocharian, contrary to our relative chronology, so that they need to be discussed.

2.5.6.1 Rounding due to *p/m in Tocharian A

Instead of turning into the expected reflex TA *a*, PT **e* next to a labial *p* or *m* has in some cases changed to TA *o*. A comparison with Tocharian B shows how the Tocharian A vowel has changed, with Tocharian B here preserving the more original situation.

Since this change only affects Tocharian A, it took place after the split of Proto-Tocharian. It is nevertheless important to address, as it has been adduced as evidence that the contrast between pre-PT **e* (PIE **ē*) and **ë* (PIE **o*) was retained for longer than assumed above in our discussion of palatalization above. This is argued by Ringe (1990: 223ff.), based on the fact that both instances of **o* with a certain Indo-European etymology go back to **ē* (viz., *ṣom* 'one (1)' and *ñom* 'name'), while those with **a* go back to PIE **o* (viz., *masäk* 'joint, link', *maku* 'nails', *pats* 'husband', *kam* 'tooth'). For this reason, Ringe (1990) set up a continued contrast between PT **e* from PIE **ē* and PT **ë* from PIE **o*. The other words are borrowings (e.g., *paräṃ* 'axe' from Iranian), or in some other way difficult to interpret.

Here we will look at the material again, to see if Ringe's continued separation of PIE **ē* and **o* as different phonemes in Proto-Tocharian is indeed warranted, in spite of the indications to the contrary (cf. especially 2.5.1 and 2.5.5). We will start in Table 2.25 with a list of the cases where TA *o* corresponds to TB *e*.

Table 2.25: Examples of TA *o* corresponding to TB *e*.

TA		pre-TA	TB
opäntäs	'in between'	**apäṅkta* (?)	*epiṅkte*
opäśśi	'clever, skilled'	**apästya*	*epastye*
opyāc	'memory'	**apäyāc*	*epiyac* (OSIr. **abiyāta-*)
omäl	'hot'	**amällä*	*emalle*
omlyi	'heat'	**amälyā*	*emalya*
ñom	'name'	**ñamä*	*ñem*
cmol	'birth'	**cämal*	*camel*
porat	'axe'	**paratä*	*peret* (OSIr. **paratu-*)
ṣom	'one (1), some'	**ṣama*	*ṣeme*

In the labial environments in the following words in Table 2.26, PT **e* changed to TA *a* per the usual development. It is immediately obvious from the relative length of this list that a development to TA *a* is more common than to TA *o*.

Table 2.26: Examples of TA a corresponding to TB e in a labial environment.

TA		pre-TA	cf. TB
anapär	'before, in front of'	**anapra*	*enepre*
apärkär	'a long time'	**a(m)pärkra*	*emparkre*
apälkāt	'unworried'	**a(m)pälkātte*	*empalkatte*
kapśañi	'body'	**kakśañi*	*kektseñe*
kam	'tooth'	**kama*	*keme*
pañi	'splendour'	**pańiya*	*peñiyo*
pats	'husband'	**patsä*	—
paräṃ	'glory'	**parna*	*perne* (OSIr. *farnah-*)
der. *parno*	'glorious'	**parnawä*	*perne*$_u$
pare	'debt'	**paray*	*peri*
pal	'(right) way, law'	**pala*	*pele*
der. *paloṃ*	'praise'	**palawnā*	*palauna* (analogical *a*)
maku	'nails'	**makwa*	*mekwa*
mañ	'moon'	**mańa*	*meñe*
malañ	'nose'	**malañä*	*meli* (diff. pl. suffix)
malke	'milk'	**malka*	*malkwer* (parallel deriv.)
malto	'first'	**maltawä*	*melte* (TA is derived)
masäk	'joint, link'	**maska*	*meske*
ram	'witness'	**rama*	*reme*
wram	'thing'	**wrama*	*wreme-*
slam	'flame'	**slama*	*sleme*

The words with a reflex TA *o* can be dividied into two groups: one where the vowel precedes the bilabial that is supposed to affect it and one where the vowel follows the bilabial. These two environments may have had different kinds of effects, so we will consider them separately. Further subdivisions in the first group can be made as in Table 2.27. The majority of cases of rounding to TA *o* are in this group, but there are also clear instances where no rounding occurred.

Table 2.27: Reflexes of PT **e* before a labial consonant in Tocharian A, classified into four categories according to the environment and the different reflexes TA *o* and *a*.

Cam < Cem	C^{y}om < C^{y}em	#op/m < #ep/m	#ap < #emp
kam < keme *ram < reme* *wram < wreme* *slam < sleme*	*ñom < n^{y}emə* *ṣom < s^{y}eme*	*opäśśi < epəst^{y}ye* *omäl < eməlle* *omlyi < eməlya* *opyāc < epəyac* *opäntäṣ < ep^{y}ənkte*	*apärkär < empərkre* *apälkāt < empəlkatte*

A subdivision into four categories can be made, depending on the position of PT **e* in the word. Here it becomes clear that PT **e* became TA *o* when followed by **m* if the preceding consonant was palatalized, but not when it was neutral. This corresponds to Ringe's PT **e* from PIE **ē*, which always caused palatalization of the preceding consonant, but formulates the same correspondence in terms of a differently conditioned development. In word-intitial position, PT **e* became TA **a* if the TB cognate shows a consonant cluster *-mp-*, which was probably original (cf. Hilmarsson 1991: 192–199). In other cases, when the TB cognate has a simple *-m-* or *-p-*, the initial PT **e* followed by a labial is reflected as TA **o*. Since the words in each of the four groups belong to a different environment, the double reflex of PT **e* in Tocharian A in front of bilabials can be seen as a conditioned development, so that it does not require the existence of two separate Proto-Tocharian *e*-phonemes like "**e* vs. **ë*" or "**e* vs **æ*".

According to the material with a PT **e* with a following bilabial, there is a difference in the development of those words where the preceding consonant is non-palatalized as opposed to those where it is palatalized. The development of TA *ñom* from **nʸemə* and *ṣom* from **sʸeme* is similar to the rounding rule that produced PT **yok-* 'drink', **yop-* 'enter' (sbj. of *yəp-*) and **yom* 'get' (prt. of *yəm-*) from earlier **yekʷ-*, **yep-* and **yem-*. A change from **yeP* to **yoP* is shared with Tocharian B, while **Cʸem* to **Cʸom* is not, but they can in principle be seen as the same type of change.

In the group of words where the vowel follows the bilabial, we can see that the picture there is more difficult. The majority of examples do not get TA **o*. Of the only two examples where **e* changed to TA *o*, TA *cmol* from PT **cəmel* stands out as the only instance containing a second-syllable sequence of bilabial + **e* + consonant. In the Tocharian B counterpart *camel* /cə́mel/, the *e* is unstressed. The other example, TA *porat* from PT **peretə* is originally a loan from Old Steppe Iranian **faratu-* (Bernard 2023: 43–48), and thus does not directly reflect any inherited Proto-Indo-European vowel. Its development of TA *o* seems to be truly unexpected, next to a number of other original trisyllabic examples that have TA *a* instead.

Table 2.28: Reflexes of PT **e* following a labial consonant in Tocharian A.

paC < peC	maC < maC	p/moC < p/meC
pañi < **penʸəyo*	*maku* < **mekwa*	*cmol* < **cəmel*
paräṃ < **perne*	*mańk* < **menke*	*porat* < **peretə*
parno < **pernewə*	*mañ* < **menʸe*	
pare < **perey*	*malke* < **melkey*	
pal < **pele*	*masäk* < **meske*	
paloṃ < **pelewna*	*malañ* < **meleñə*	
pats < **petˢə*	*malto* < **meltewə*	

Since it is possible to come up with secondary developments to get the unexpected *o*-vowels in Tocharian A in all cases except TA *porat*, which is a borrowing from Old Steppe Iranian, there is no need to set up a continued differentiation of PIE **ē* and **o* as PT **e* and **ë* respectively to account for it. Rather, only a single PT **e* is needed. This is also in line with the evidence from both pre-Proto-Tocharian *u*-umlaut (2.5.1) and *a*-umlaut (2.5.5), which targeted the results of PIE **ē* and **o* in the same way, implying a merger between the two vowels.

In conclusion, the Proto-Tocharian vowel **e* seems to have been regularly changed to TA *a*, with secondary rounding to TA *o* in the environments *C*^{y}*_m*. The only unexpected TA *o* from PT **e* are found in TA *cmol* 'birth' and in TA *porat* 'axe', the latter of which is not an inherited word from Proto-Indo-European, but instead a borrowing from Old Steppe Iranian **faratu-*. These words cannot uphold a separation of PIE **ē* and **o* until after the Proto-Tocharian period in the face of the other evidence in support of their merger in pre-Proto-Tocharian.

2.5.6.2 Rounding due to **k*^{w} in Tocharian A

Some vowels in the vicinity of labiovelars were also rounded in Tocharian A. This very clearly affected the PT vowel **ə*, which surfaces as TA *u* in words such as PT **yək*^{w}*e* 'horse', TA *yuk* (TB *yakwe* /yákʷe/) and PT **səkʷə* 'happiness', TA *suk* (TB *sakw* /sákʷ/). The sequence may also be written as TA *$_{u}$k* in derived and inflected forms, such as in the oblique singular *y$_{u}$kas* from *yuk* and the derived adjective *s$_{u}$kaṣi* from *suk*.

Given its effect on **ə*, it is not surprising that **k*^{w} is often thought to cause rounding of PT **e* in a similar way to **p* and **m* as well. This has been proposed as the reason for the TA *o* in at least the following words:

- TA *oṅk* 'man' beside TB *eṅkwe* 'id.'
- TA *oṅkrac* 'immortal' beside TB *oṅkrocce* 'id.'
- TA *opṣäly* 'season; time for action (of a Buddha)' beside TB *ekṣalye* 'id.'
- TA *orkäm* 'darkness; dark' beside TB *orkamo* 'darkness; dark'
- TA *śorki* 'mortal fear' and *śorkmi* 'strings' beside TB *śerkw* 'cord, string'
- TA *nokte* 'at night', *noktiṃ* 'last night' beside TB *nekciye* 'at night; last night'
- TA *koṃ* 'dog obl.sg.' beside TB *kweṃ* 'id.'

The first two examples, TA *oṅk* 'man' and *oṅkrac* 'immortal' are structurally very similar, but an important difference is seen in the Tocharian B counterparts TB *eṅkwe* and TB *oṅkrocce*. Only the former, TB *eṅkwe*, has the vowel TB *e*, whereas TB *oṅkrocce* has *o*-vowels just like TA *oṅkrac* (in the first syllable). This allows for the possibility that the Proto-Tocharian form was already **oṅkrocce* rather than **eṅk*^{w}*rocce* or the like, and that the first *o* was caused by *o*-umlaut from the second *o*. The processes at work in TA *oṅkrac* :

TB *oṅkrocce* may thus not be truly parallel to those seen in TA *oṅk* : TB *eṅkwe*, and this word is left aside here. The same applies to TA *orkäm* next to TB *orkamo*, which is claimed to result from affection in PT **erkʷmo* by Kim (1999). However, the matching vowel TAB *o* and the lack of a labiovelar in Tocharian B indicate that this TAB *o* arose at a different time and due to a different development (see above).

The TA *o* in *oṅk* 'man' is mostly considered to be caused by the labiovelar **kʷ* (Hilmarsson 1989a: 108–109; Kim 1999: 167–168).[34] With the shift from **e* to **a* and general apocope of word-final vowels in Tocharian A, PT **enkʷe* might be expected to yield ***anku*, especially when compared to the development in a number of words with similar structure, like PT **kəntwo* 'tongue; language' to TA *käntu* (TB *kantwo*), PT **mekʷa* 'nails' to TA *maku* (TB *mekwa*) and PT **sekʷe* 'pus' to TA *saku* (TB *sekwe*; cf. Hilmarsson 1989a). These all realise the original labiovelars or clusters with **-w-* as a final *-u* in Tocharian A, while no *-u* is found in TA *oṅk* 'man'.

The development of PT **enkʷe* 'man' to TA *oṅk* shares this feature with examples like PT **yəkʷe* 'horse' to TA *yuk* (TB *yakwe*) and PT **tənkʷə* 'love' to TA *tuṅk* (TB *taṅkw*), which have no final *-u* either; and just like TA *ṅonk* they have undergone a change in vowel quality instead. TA *tuṅk* even shows affection of the vowel by **kʷ* across an intervening velar nasal *ṅ* [ŋ], whose presence also separates TA *oṅk* from the likes of TA *saku*, *maku* and *käntu*. The development can thus be considered as one from pre-TA *[aŋʷkʷ] to [oŋk], with loss of the labiovelar articulation upon vowel affection (cf. Kim 1999: 167–168). Considering the lack of affection in TA *saku* and *maku* (rather than ***sok* and ***mok*), the presence of certain clusters appears to be a prerequisite for rounding of PT **e* or pre-TA **a* to TA *o* by a following labiovelar.

A similar case is found in TA *śorkmi* 'strings', which derives from a PT stem **śerkʷ-* reflected in TB *śerkw* 'cord, string' (see Malzahn 2014: 87–90 for more details) with affection of PT **e* to TA *o*. The only element that could have caused this vowel change is the labiovelar, which was subsequently lost, like in the case of TA *oṅk* from PT **enkʷe*. It thus turns out that the effect of the labial quality could traverse an intervening *r* as well as [ŋ], but did not apply when **kʷ* immediately followed the vowel PT **e*. Only the "weaker" vowel PT **ə* was more generally affected by a following labiovelar, as in TA *yuk* 'horse' from PT **yəkwe* and TA *tuṅk* 'love' from PT **tənkwə*.

The case of TA *opṣäly* '(fit) season; time for action (of a Buddha)' and TA *nokte* 'at night' together with *noktiṃ* 'last night' is somewhat different again, since the TB cognates

[34] Pinault takes TA *oṅk* as the result of *u*-umlaut from an earlier **enku*. TB *eṅkwe* then constitutes a different, independent development, a secondary derivation from the same stem that was spared from *u*-umlaut (Pinault 2008: 452). It seems preferable to derive the word for 'man' in Tocharian A and B from the same Proto-Tocharian source, however, which for Tocharian B certainly can only be reconstructed as **enkʷe*.

ekṣalye and *nekciye* to not have a labiovelar attested. This may be due to spelling inconsistencies known to plague the writing of k_u (e.g., Fellner 2005) or late changes to this cluster over the course of Tocharian B attestation (see Peyrot 2008 on a number of such developments), but the Tocharian A vowels can be explained differently as well. For TA *opṣäly* and TB *ekṣalye*, Adams reconstructs a PT **ekʷsʸəlye* to achieve the rounding of the first vowel in Tocharian A (Adams 2013: 80-81), but in Tocharian A there was a general change from PT **kS* to TA *pS* that could provide a *p*; cf. TA *opsi*, the plural of PT **okso* 'ox'. The *p* that resulted from this may have affected the preceding PT **e* or pre-TA **a* in the same way as the original *p* in TA *opyāc* 'memory' beside TB *epiyac*, etc., discussed above.

TA *nokte* and *noktiṃ*, meanwhile, have an additional third cognate in TA *nakcu* 'last night, at night'. This is a closer cognate to TB *nekciye* and does not show a TA *o*, i.e., not ***nokcu*. If a labiovelar had affected the vowel in TA *nokte* and *noktiṃ* from **nekʷt-*, the same effect should have been expected in TA *nakcu* if from **nekʷtʸ-*. Since this is not the case, a different explanation may be preferred: the vowel TA *o* in *nokte* and *noktiṃ* can be explained as the result of *u*-umlaut, based on an earlier *u*-stem PIE **nokʷtu-*. The resulting pre-PT **nektu-* became **noktə-* to become the base of TA *noktiṃ*, while the derived adjective **nokʷtewyo* yielded PT **nektʸəwʸe*, whence TB *nekciye* and TA *nakcu* (Pinault 1990: 184–190).

Finally, in TA *koṃ* (TB *kweṃ*), the oblique singular of *ku* 'dog', there is a clear change from PT **e* to TA *o* next to an original labiovelar. It appears that a labiovelar was not preserved in word-initial position next to pre-TA **a* (**kʷan*) and merged with the vowel. No similar process is seen in TA *śtwar* '4' (not ***śtor*) from PT **kʸətwerə*, suggesting that the effect was caused by PT **kʷ* in particular, rather than by any consonantal **w*.[35]

This concludes our overview of rounding phenomena next to labials and labiovelars in Tocharian A. While a few changes remain obscure, it should be clear that no two distinct vowels for PT **e* need to be reconstructed to account for these developments (*pace* Ringe 1990): such a distinction does not properly solve the remaining issues either. This is also clear from the limited labial affection in Tocharian B, treated in the next subsection.

2.5.6.3 Rounding due to labials in Tocharian B

Just like Tocharian A, Tocharian B had its own marginal labial affection as well, specifically between two labials. There are two examples of this, although there are also two instances where it did not take place. Both of the examples belonging to this later

[35] The development of TA *por* from earlier **puwar* (TB *puwar*) does not quite belong here, since it involves the vowel PT **a* (TA *ā*) rather than **e* (TA *a*).

group are derived from the same verbal root *pəkʷ-*, so that some kind of paradigmatic levelling may have occurred. Perhaps only unstressed **ə* was affected.

64. PIE **kʷekʷlos* > "**kukʷlos*" > PT **kʷəkʷle* > TB *kokale*, TA *kukäl* 'wagon'
- PT **pə-kʷaka* > TB *pokkāka*, TA *pᵤkāk* imp. of *kaka-* 'call'
- cf. PT **pəkʷəmane* > TB *pkwamane*, not ***pokamane* or ***pkomane*, prt.ptc. of *pəkʷ-* 'cook'
- cf. PT **empəkwətte* > TB *empakwätte*, not ***ompokätte* vel sim., privative of *pəkʷ* 'cook'

In the available Tocharian A cognates, the vowel is reflected as *u* or as part of a labiovelar *kᵤ*, not as TA *o*. The word TB *troṅk*, TA *truṅk* 'cavity, hollow, inner part of the torso; cave' shows this same vowel correspondence between Tocharian B *o* and B *u*, but it does not quite share the same phonological environment, and its etymology is not certain (see Adams 2013: 341).

Some word-initial PT **e* were also rounded to TB *o*, similar to the change observed in Tocharian A (see above). A number of examples have a TB *o* in the second syllable as well, which would regularly have caused **e* to change to TB *o* due to umlaut, but four examples remain where no such explanation is applicable, namely *oṅkipṣe* 'shameless', *ompakwättäññe* 'unreliability', *ompalsko* 'meditation' *oṣṣale* 'north'.

- TB *oṅkipṣe* 'shameless' derived from TB *kwipe* 'shame'
- TB *omotruññaiṣṣe* 'southern', cf. TA *mäcriṃ* 'southern'
- TB *ompakwättäññe* derived from *empakwätte* 'unreliable'
- TB *ompalsko* 'meditation' derived from TB *palsko* 'thought'
- TB *ompostäṃ* 'afterwards' derived from TB *postäṃ* 'after'
- TB *omprotärtstse** 'related as brothers' derived from TB *procer*
- TB *oṣṣale* 'north' derived from TB *ṣale* 'mountain', cf. TA *ṣul* 'mountain' from PT **sʸwʸəle* (**wʸ* regularly changed to **y* in TB, which is lost in the initial cluster in **ṣyale* > *ṣale*, but is responsible for the gemination in *oṣṣale*)

It is very difficult to differentiate between TB *eṅkwe* 'man' from PT **enkʷe* and TB *oṅkipṣe* 'shameless' from earlier **enkʷipsʸe*, and between TB *empakwätte** /empə́kʷətte/ 'unreliable' and *ompakwättäññe* /ompə́kʷəttəńńe/ 'unreliability' from PT **empəkʷətte* by a regular sound change. Especially the latter cannot be differentiated on the basis of features like stress placement or syllable structure, but we may assume that these differences were caused by analogy and do not all reflect regular sound change. All instances do contain an intervening element between the affected vowel and the labial. The change in TB *oṣṣale* 'north' needs to have happened before PT **wʸ* had change into

TB *y*, after which no labial would have been present anymore to affect the preceding vowel. The list of examples below shows that, more commonly, TB *e* is found in this kind of environment.

- TB *eṅkwe* 'man', cf. TA *oṅk*
- TB *epastye* 'skillful', cf. TA *opäśśi*
- TB *epiṅkte* 'between', cf. TA *opäntaṣ*
- TB *epiyac* 'memory', cf. TA *opyāc*
- TB *eprer* 'sky', where the *e*- actually alternates with an *i*- as *iprer*
- TB *emalle* 'hot, warm', cf. TA *omäl* 'hot'
- TB *eprete* 'resolute, steadfast', cf. TA *pratiṃ* (without the prefix)
- TB *eplyuwai* 'swimming' derived from *pləw*- 'swim'
- TB *empreṃ* 'truth'
- TB *empakwätte* 'unreliable' derived from *pək*^{w}-
- TB *emparkre* 'wide(ly)', derived from TB *parkare* 'wide'
- TB *empalkatte* 'unworried', cf. TA *apälkāt*
- TB *empele* 'terrible, horrible', derived from *pele* 'way, law' → TA *empele*

The next subsection will continue with an opposite development of unrounding that seems to also have taken place in Tocharian A under some very specific circumstances.

2.5.6.4 Potential unrounding in Tocharian A

One remaining thorny problem is the Tocharian A correspondences to TB *o*. In the non-first syllable, this is regularly TA *a*, while in the first syllable it is generally TA *o*.

1 TB obl. *pokai* vs. TA obl. *poke* 'arm'
59. TB *klyomo* vs. TA *klyom* 'noble'
84. TB *okt* vs. TA *okät* 'eight (8)'
85. TB m.pl. *poñc* vs. TA m.pl. *poñś* 'all' (and other forms of the paradigm)
109. TB (arch.) *toṃ* vs. TA *tos-äṃ* pronoun stem f.pl.
124. TB *okso* vs. TA pl. *opsi*, sg. *okäs** 'ox'

- TB obl. *kroścaṃ* vs. TA obl. *krośśäṃ* 'cold'
- TB *yoñiye* vs. TA *yoñi* 'path'
- TB *kroriya* 'horn' vs. TA *kror* 'crescent of the moon'
- TB *moko* vs. TA *mok* 'old'
- TB *oṅkolmo* vs. TA *oṅkaläm* 'elephant'
- TB *oṅkrocce* vs. TA *oṅkrac* 'immortal'
- TB *kos* vs. TA *kos* 'how much'

- TB *yok* vs. TA *yok* 'colour; hair'

There are just two examples with TA *a* in the first syllable instead of the *o* found in a direct Tocharian B cognate. In both cases, the vowel is a precious reflex of PIE *$*eh_2$*/$\bar{a}$. The discrepancy between the words in the two lists below is the main reason to distinguish the result of *$*eh_2$*/$\bar{a}$ from the product of *u*-umlaut as separate Proto-Tocharian vowels *$*å$* and *$*o$* or alternatively *$*o$* and *$*ǫ$* (cf. Pinault 2008: 420; Ringe 1996: 98).

2. TB *procer* vs. TA *pracar* 'brother' from PIE *$*b^hreh_2tēr$*
115. TB *ost* vs. TA *waṣt* 'house' from PIE *$*u̯eh_2stu$-*

This same correspondence is found in a number of *o*-presents [class 4], but the origins of those vowels are probably of a different nature, and they are discussed in subsection 2.5.4. It is problematic, however, that TA *pracar* 'brother' and TA *waṣt* are the only two examples of a development from PIE *$*eh_2$*/$\bar{a}$ into TA *a*, while the other examples show TA *o*. From the examples listed above, this concerns the following:

1 TB obl. *pokai* vs. TA obl. *poke* 'arm' from PIE *$*b^heh_2ǵ^hu$-*
85. TB m.pl. *poñc* vs. TA m.pl. *poñś* 'all' from PIE *$*peh_2$-nt-*
109. TB (arch.) *toṃ* vs. TA *tos-äṃ* pronoun stem f.pl. from PIE *$*teh_2$-ns*

TA *poke* 'arm' (TB *pokai*) from PIE *$*b^heh_2ǵu$-* could conceivably have its *o* in Tocharian A from a kind of *u*-umlaut (cf. the *u*-umlaut of pre-PT *$*e$* in 2.5.1), but there is no parallel for *u*-umlaut of *$*eh_2$*/$\bar{a}$. The word that is structurally most similar is PIE *$*u̯eh_2stu$-*, which has a different vowel correspondence. In TA *poke* 'arm' and *poñś* 'all', the initial labial consonant *p-* may have had a rounding effect, but the vowel of TA *tos-äṃ* cannot be easily explained in the same way. It is difficult to assume some kind of umlaut due to a suffixed element in this word because the correspondence of TB *-ṃ* (/n/) to TA *-s* is normally found only in word-final position. This development thus contradicts the evidence of TB *procer*, TA *pracar* 'brother' and TB *ost*, TA *waṣt* 'house'. It therefore indicates that a reconstruction of two different *o*-phonemes PT *$*o$* and *$*å$* from distinct Proto-Indo-European sources might not be sufficient to differentiate these varying correspondences.

Burlak and Itkin suggested effects of the initial labial on the vowel in TA *waṣt*, with a sound change from PT *$*#wo$-* to TA *#wa-* and TB *#o-* (Burlak & Itkin 2003: 31). This might work, as a dissimilation is phonetically plausible and there are no counterexamples. The explanation given by Burlak and Itkin for the TA *a* in *pracar* 'brother' is unacceptable, however: they claim that this goes back to PT *$*pərotʸer$*, so that the PT *$*o$* stood in the second syllable, where it was regularly changed to TA *a* as in TB *oṅkolmo* vs. TA *oṅkaläm*

'elephant' (Burlak & Itkin 2003: 33). This explanation relies on the assumption that all pre-Proto-Tocharian clusters of the type **CR* were broken up with epenthesis to **CəR*, for which there is simply no evidence. Furthermore, the lack of a TA *a* in TA *klyom* 'noble' (not ***klyam* from ***kəlyomo*) and *kror* 'crescent of the moon' (not ***krar* from ***kərorə*) directly contradicts such an explanation for TA *pracar*. The TAB *o* in PT **klyomo* and **krorə* is not derived from late sporadic simplifications of a diphthong **eu* to **o*, as Burlak and Itkin claim (2003: 26–27), but from umlaut (see 2.5.1 and 2.5.2 with references).

If we assume with Burlak and Itkin a dissimilation from PT **wo* to TA *wa* in TA *waṣt* 'house', perhaps a similar dissimilation can have affected PT **pro* in TA *pracar* 'brother' from PT **protʸer*. This would amount to a potential rule whereby **o* was unrounded in Tocharian A in certain labial environments. However, **po* and **mo* would clearly have remained unaffected; cf. TA obl. *poke* 'arm' (TB *pokai*) and *mok* 'old' (TB *moko*). If dissimilation occurred in **pro* one might also expect it in **po*, perhaps more strongly because the labial is adjacent to the vowel. However, this may not necessarily be the case, since we can see in the development of TA *oṅk* 'man', *śorki* 'mortal fear' and *śorkmi* 'cords' as opposed to *maku* 'nails' and *saku* 'pus', that **kʷ* had actually had a different effect on an adjacent vowel (in this case rounding) when there was an intervening resonant present. The situation for **pro* as opposed to **po* could thus have been similar.

There is an additional, though more uncertain example which may further support this sound change: TA *kroś* 'cold' is once attested as *kᵤraś* instead, which displays both an initial TA labiovelar *kᵤ* and a change of the vowel to TA *a*. Itkin (2011: 251) has characterized TA *kᵤraś* as "tout simplement due à une faute de scribe". The variant *kroś* in the nominative singular is attested twice for sure, two more times potentially (damaged context), and *kᵤraś* is found only once.[36] It would be a rather strange scribal error, however, so that one may suppose that *kᵤraś* constitutes a real form, and reflects a real phonological process.

If a sound law changing PT **wo-* and **pro-* (and perhaps ?**kʷro-*) to TA *wa-* and *pra-* (and perhaps *kʷra-*) is assumed, this would eliminate the only remaining reason to reconstruct the Proto-Tocharian reflex of PIE **eh₂* as different from the PT **o* that arose due to *u*-umlaut of **e* and different from the second-syllable **o* in PT **wəlo*, **klʸomo*, etc. from PIE **ō*. They could furthermore be written exactly as the *o*-vowel resulting from *u*-umlaut and *o*-umlaut, all simply PT **o*. The Proto-Tocharian vowel **å* still existed, but it was the result of contraction and umlaut caused by the contracted vowel (see subsection 2.5.4 and also 5.4.1). Its correspondences are uniformly TB *o* : TA *a*. Because the uncertainties involved cannot easily be resolved, it might be prudent to provide reconstructions with PT **å* as an alternative, thus PT **wostə* (**wåstə*) 'house'.

[36] The meaning in all cases seems clear.

2.5.7 Summary

Two types of umlaut, namely *u*-umlaut of **e* and *a*-umlaut of **e*, applied in the same way in Tocharian A and B, and can thus be dated to pre-Proto-Tocharian as a set of sound laws (see 2.5.1 and 2.5.5). There is no difference in the application of these various types of umlaut to the Proto-Indo-European vowels **ē* and **o*, which can thus have merged as (pre-)PT **e* without any issue. The labial affection that turns some PT **e* into TA *o* rather than TA *a* appears to have operated in certain restricted environments and was not demonstrably confined to a PT **ë* or **e* as reflexes of PIE **o* or **ē* (see 2.5.6.1 and 2.5.6.2).

The Proto-Indo-European vowel **ō* seems to have caused *o*-umlaut of pre-PT **u*, although there are only two potential examples (see subsection 2.5.2). A similar *o*-umlaut of pre-PT **e* caused by a following PIE **eh₂/ā* clearly affected Tocharian B, but the evidence for this umlaut in Tocharian A is uncertain (see subsection 2.5.3). It is difficult to determine the reflex of the vowels PIE **eh₂/ā* in Proto-Tocharian was merged with the PT **o* from *u*-umlaut of **e* and *o*-umlaut of **u*, or if it became identical with the contracted vowel PT **å*. Three examples suggest that PIE **eh₂/ā* was identical with **o* and two suggest a merger with **å*. The TA *a* rather than *o* in TA *pracar* 'brother' (TB *procer*) and TA *waṣt* 'house' (TB *ost*) could be the result of a secondary, dissimilatory unrounding phenomenon that affected the sequences **pro-* and **wo-*, but it does not seem possible to come to a firm conclusion (see 2.5.6.4).

2.6 Information from language contact

Parts of the relative chronology can be put in a broader perspective on the basis of Tocharian words that were borrowed from early Iranian languages. The phonological characteristics of these loanwords can help determine the state of early Tocharian phonology at the time of contact with Iranian. Some Tocharian words derive from another, unknown prehistoric language, thought to be related to the language of the BMAC, and a little additional information may be gleaned from this as well.

2.6.1 Early Iranian languages

At a certain point, (pre-)Proto-Tocharian came into contact with an Old Iranian language, based on Old Iranian loanwords that have entered into the Tocharian lexicon. Examples include TB *perne*, TA *paräṃ* 'glory' from Old Iranian **farnah-*. The earliest Old Iranian donor language has been labelled "Old Steppe Iranian" by Chams Bernard (2023), and the contact seems to have taken place a short while before the split of Proto-Tocharian into pre-Tocharian A and pre-Tocharian B. Bernard has established regular sound correspondences between the Old Iranian donor language and the Tocharian borrowings, where, for example, each Old Steppe Iranian */a/* corresponds to a Proto-Tocharian **e*, while Old Steppe Iranian */ā/* corresponds to a Proto-Tocharian **a*. On the

basis of the regularity of these vowel correspondences, Bernard argues that the contact with Old Steppe Iranian took place over only a short, concentrated period of time, in which neither it nor (pre-)Proto-Tocharian underwent significant phonological change (Bernard 2023: 242–246).

The Old Steppe Iranian layer of borrowings in Tocharian can give us some insight into which Tocharian sound laws had already run their course. Based on the borrowing TB *yentuke* 'India(n)' from putative Old Steppe Iranian **i̯anduka-* or **i̯enduka-* (Bernard 2023: 60-62), we can tell that *u*-umlaut of **e* applied before contact with Old Steppe Iranian. Since *u*-umlaut changed earlier **e-u* into **o-u*, this word should have become ***yontuke* if it had been borrowed before *u*-umlaut. This indicates that the phonological stage of the pre-Proto-Tocharian language as it was when it came into contact with Old Steppe Iranian was already quite advanced, since *u*-umlaut of **e* was a relatively late development that took place after palatalization. Additionally, TB *eñcuwo*, TA *añcu** 'iron', borrowed from an early form of Khotanese (Peyrot, Dragoni & Bernard 2022) shows no *u*-umlaut either. The apparent *u*-umlaut in TB *mot* 'alcoholic beverage', borrowed from a descendant of Proto-Iranian **madu-* likely took place in the Old Steppe Iranian source language, and not in Tocharian itself (Bernard 2023: 103–104).

A similar argument can be made for *a*-umlaut of **e* (PIE **o*, **ē*), on the basis of the archaic attestations of Tocharian B *ainake* 'evil' as *eynāke*, borrowed from Old Steppe Iranian **a̯ināka-* (Bernard 2023: 27–29). If *a*-umlaut had taken place after the adoption of this word, no *e* in archaic TB *eynāke* would have been expected, as it should have changed into *a* as well: archaic TB ±***āynāke*. Both *u*-umlaut and *a*-umlaut occurred after palatalization, so that the contact itself can also be dated after palatalization. Indeed, the way in which certain Old Steppe Iranian sounds were adopted corroborates this; e.g., TB *waipecce* 'property, possession' from OSIr. **hwai-paϑya-* and TB *śāte* 'rich' from OSIr. **čyāta-* 'happy' (Bernard 2023: 73–75).

The situation is more nuanced in the case of the *o*-umlaut of pre-PT **u* (type **ukso* to TB *okso* 'ox'), because two borrowings from early Khotanese-Tumshuqese seem to show this sound change as well. Two further borrowings did not undergo this change, however, so that no clear conclusions can be drawn. Perhaps TB *cowo** 'stealing' and TB *koto** 'hole, pit; excrement (?)' were borrowed at a relatively earlier stage, but the exact timeline of developments remains elusive.

- TB *cowo** 'stealing' from pre-Khot. acc.sg. **dyūwu*, derived from PIr. **dab-ya* 'stealing', to be compared with LKhot. *dyūka-* 'robber' (Dragoni 2023: 118–120).
- TB *koto** 'hole, pit; excrement (?)' from PTKhot. or pre-Khot. acc.sg. **gūϑu*, whence Khot. *gūha-* 'excrement, faeces' (Dragoni 2023: 103–105).

- TB *wicuko* (not **wicoko*) 'cheek, (jaw)bone' form pre-Khot. **wi-jwäka-*, a derivative of K °*jv-* 'chew' based on the Tocharian borrowing (Dragoni 2023: 169–170).
- TB *sumo* (not **somo*) 'libation (?)' from LKhot. *ysūma-* 'broth' (Dragoni 2023: 196–197).

The situation for *o*-umlaut of pre-PT **a* is similar, since there are Khotanese-Tumshuqese borrowings that have undergone this sound change in Tocharian B. This may point to a relative late date for this sound change, perhaps restricted to pre-Tocharian B, after Proto-Tocharian had ceased to exist as a unitary language.

- TB *kompo** '?' perhaps via **kampo* from OKhot. acc.sg. *ggaṃphu* 'plain', or its PTKhot., pre-Khot. or OK ancestor (Dragoni 2023: 106)
- TB *kontso** '?' perhaps via **kantso* from OK *ggaṃjsā-* 'flaw', or its PTK or pre-K ancestors. (Dragoni 2023: 105)
- TB *koro* 'donkey' perhaps via **karo* from Khot. acc.sg. *kharu* 'donkey', alternatively from PTK **gōru* 'wild ass', in which case no umlaut is needed (Dragoni 2023: 106–107)

An aspect of the Tocharian phonology that remains difficult to pin down in the timeline is the general voicelessness in the stop system. Voiceless stops are used in borrowings from Old Steppe Iranian and Khotanese-Tumshuqese when the source has a voiced stop, but whether this is caused by an incorporation of the borrowings into the native sound system, or due to a later loss of the voicing opposition after the adoption of these words is not strictly visible. In both scenarios, the attested situation would show the stops as voiceless.

However, an argument in favour of general devoicing prior to contact with at least Indo-Iranian can be made as follows: Starting from contact with Old Steppe Iranian, Tocharian would have continuously been in contact with some form of Iranian until its eventual attestation. Since all the Iranian languages with which Tocharian was in contact, like Khotanese and Sogdian, also had a voicing opposition, there would have been no pressure for Tocharian to undergo a general devoicing of its stops. Rather, a retention of a voicing distinction, if still present at the time, should have been facilitated by the abundance of other languages with this distinction in the vicinity. This suggests that the devoicing of the Tocharian stops took place at least before contact with Old Steppe Iranian.

It is also clear that the sequence PIE **t̯*/*d*ʰ*t̯* had already become pre-PT **t*ˢ, since Old Steppe Iranian **t̯* or **d̯* became **cc* instead (e.g., OSIr. **gr̥d̯a-* → TB *kercci* 'palace'; Bernard 2023: 157–159). All available information thus points to an advanced stage of

development in the Proto-Tocharian phonological system at the time of contact with Old Iranian, meaning that the specific Tocharian sound changes had developed at an even earlier stage already.

2.6.2 "BMAC language"

Potentially somewhat farther back in time, Tocharian was in contact with a language related to the unknown language of the BMAC culture. This hypothesis is based on a set of similar loanwords that are found both in Tocharian and in Indo-Iranian languages that would have been in more direct contact with the BMAC Culture known from archaeology (e.g., TB *kercapo* ~ Vedic *gardabhá-* 'donkey', Bernard 2023: 199–232 with extensive discussions and references)
Bernard has argued that phonological differences in the borrowings that are not easily ascribed to either Tocharian or Indo-Iranian suggest that the language encountered by the early Tocharians was not quite the same as the one encountered by the Indo-Iranians farther to the south-west, but rather a related variety (Bernard 2023: 232–241).

Although little is known directly about the BMAC language (family), the shape of the borrowings in Tocharian can give us some crucial information about the state of the Tocharian phonological system at the time of borrowing. As with the Old Steppe Iranian borrowings, it looks like Tocharian phonology had developed to a relatively advanced state already, with palatalized consonants being present (e.g., in TB *kercapo* 'donkey', TB *kroṅkśe* 'bee', TB *śerwe* ~ TA *śaru* 'hunter'; Bernard 2023: 215–228). However, no sequences liable to umlaut are found in the words that have so far been identified as borrowings from a variety of the BMAC language, so that it not possible to determine whether or not those developments had already taken place. The nature of the stop system of the BMAC language is also difficult to accurately determine, so it cannot be conclusively argued that Tocharian should have lost its stop distinction before contact with this language.

2.6.3 Summary

The general picture as can be gleaned from traces of interactions between Tocharian and Iranian languages like Old Steppe Iranian and early Khotanese-Tumshuqese in the later stages of Proto-Tocharian is that most major sound changes had occurred at an even earlier stage. Palatalization had certainly already taken place, and at least two types of umlaut seem to have occurred before contact with at least Old Steppe Iranian as well, namely *u*-umlaut of **e* (on account of TB *yentuke* 'India(n)') and *a*-umlaut of **e* (on account of archaic TB *eynāke* 'evil'). It cannot be ascertained whether *o*-umlaut of **u* had occurred, or if **u* and **i* had merged as **ə*. If the early Khotanese-Tumshuqese borrowings TB *cowo** 'stealing' and TB *koto** 'hole, pit; excrement' indeed show *o*-umlaut

of **u*, this sound change may have occurred relatively later. The Tocharian B *o*-umlaut of PT **a* is reflected in a number of Khotanese-Tumshuqese borrowings found in this language, and it may therefore have been a late sound change as well, although the available data is sparce and partially contradictory.

Palatalization is confirmed by borrowings from an unattested language related to the "BMAC-language" that donated loanwords to various early Indo-Iranian languages as well. Due to the limited information available about this language, not much else can be said on the basis of the current evidence.

2.7 Overview of Tocharian relative chronology

The developments that occurred in pre-Proto-Tocharian can be roughly split in two groups: those that took place before palatalization, and those that took place after palatalization. The relatively more significant changes belong to the first group, and the second group mostly consists of umlaut phenomena and a few additional vowel changes, as well as further shifts in the palatalized pronunciation of certain consonants, e.g., **s^{y}* > TAB *ṣ*, **k^{y}* > TAB *ś*.

We will begin our overview once again with the Proto-Indo-European phonological system. The phonemes that were affected by changes at an early stage were the laryngeals and the various velar series. A few changes are datable relative to laryngeal developments, which as a whole belong before palatalization.

Table 2.29: The Proto-Indo-European consonant inventory.

PIE	labial	dental	palatal	palatovelar	velar	labiovelar	laryngeal
stop	*p b^{h}*	*t d d^{h}*		*ḱ ǵ ǵ^{h}*	*k g g^{h}*	*k^{w} g^{w} g^{wh}*	
unknown							*h$_1$ h$_2$ h$_3$*
sibilant		*s*					
nasal	*m*	*n*					
liquid		*l r*					
glide			*i̯*			*u̯*	

Table 2.30: The Proto-Indo-European vowel inventory.

PIE	front	back
high	*i*	*u*
mid	*e ē*	*o ō*
low	(*a*)	

Changes involving laryngeals seem to have happened particularly early on. This caused an early increase in the number of vowels in the system due to developments such as *h_2e to *a, *eh_2 to *$\bar{a}$, *ih_1 to *$\bar{\imath}$, $\bar{o}$# > $\bar{u}$#, and the frequency of certain vowel phonemes increased as well by developments like those from *eh_1 to *$\bar{e}$. The vowels that arose in this way behave exactly like their more original counterparts in further developments, so that the colouring effects of the laryngeal can be considered particularly early. After laryngeal colouring, the three laryngeals can no longer be distinguished in the phonology, so that they might be merged as *H and eventually lost (2.3.1). The vocalization of syllabic resonants from *$\underset{\circ}{R}$ to *"$\dot{\imath}R$" added one further vowel to the system (2.3.2), but it is difficult to date exactly. It had certainly taken place before the disappearance of laryngeals and before phonemic palatalization. An additional source of the new vowel *∂ may have been the laryngeal in sequences *Hi (2.4.2). Other changes either did not interact with the laryngeal developments, or occurred demonstrably later.

Table 2.31: An overview of early vowel changes from Proto-Indo-European to pre-Proto-Tocharian.

PIE	> post-PIE	PIE	> post-PIE
i	i	ih_1?	$\bar{\imath}$
e	e	$\bar{e}$, eh_1	$\bar{e}$
u	u	-$\bar{o}$#	-$\bar{u}$#
o, h_3e?	o	$\bar{o}$, oH	$\bar{o}$
$\underset{\circ}{H}$, h_2e	a	eh_2	$\bar{a}$
$\underset{\circ}{R}$	"$\dot{\imath}R$"		

Table 2.32: The Proto-Indo-European vowel inventory (left) and the new early pre-Proto-Tocharian vowel system after the first changes (right).

PIE	front	back		front	central	back
high	i	u		i $\bar{\imath}$	"$\dot{\imath}$"	u $\bar{u}$
mid	e $\bar{e}$	o $\bar{o}$	>	e $\bar{e}$		o $\bar{o}$
low	(a)				a $\bar{a}$	

The depalatalization of palatovelars probably took place very early, but it cannot be dated relative to any change other than the general Tocharian palatalization, which it preceded (2.3.3.1). Delabialization of labiovelars next to PIE *e and *$\bar{e}$ must have happened before *$\bar{e}$ merged with *$\ddot{e}$ from PIE *o as (pre-)PT *e, which occurred after with palatalization (2.4). PIE *d changed in some uncertain way (see 2.3.4.4) at least in clusters before resonants before assibilation of PIE *$t\underset{\wedge}{\imath}$ and *$d^h\underset{\wedge}{\imath}$ took place, and in its turn, assibilation also happened before palatalization (2.3.4.2 and 2.3.4.3). The putative Tocharian version of Grassmann's Law would probably have taken place even earlier

(2.3.4.5). The details of the developments of PIE **d* in Tocharian remain unclear, however, and I have decided to provisionally use the symbol "*ḋ*" to mark an altered state of PIE **d* that is of unknown quality (see 2.3.4.4 and 2.3.4.6).

These changes happened before obstruents were generally devoiced and de-aspirated. The simplification of the cluster PIE **mbʰ* to **m* belongs here too (see 2.3.4.1). Since no distinctions in voicing or aspiration are necessary to understand the outcomes and effects of palatalization in general, it can be assumed that the loss of these features occurred before palatalization, but this cannot be confirmed. A summary is provided in Table 2.33 with an overview of the new consonant inventory in Table 2.34.

Table 2.33: An overview of early changes to the Proto-Indo-European consonants and their relative chronology.

Sound changes		Chronology	Treated in
A1 ? *dʰVDʰ*	> *dVDʰ*	if this did happen, before A2., D.	2.3.4.5
A2 *d*	> "*ḋ*"	before A3, B, D.	2.3.4.3, 2.3.4.4
A3 *dṷ, dị, dr*	> *w, y, r*	likely before B, D.	2.3.4.3
B *tị, dʰị*	> "*ċ*" ("*ż*ʰ")	before D.	2.3.4.2
C ? *mbʰ*	> *m*	before D.	2.3.4.1
D1 *p, bʰ*	> *p*		2.3.4 *passim*
D2 *t, dʰ*	> *t*		2.3.4 *passim*
D3 *ḱ, ǵ, ǵʰ*	> *k*		2.3.4 *passim*
D4 *k, g, gʰ*	> *k*		2.3.4 *passim*
D5 *kʷ, gʷ, gʷʰ*	> *kʷ*		2.3.4 *passim*

Table 2.34: The pre-Proto-Tocharian consonant inventory after the first consonant changes had taken place.

pre-PT	labial	dental	palatal	velar	labiovelar
stop	*p*	*t* "*ḋ*"		*k*	*kʷ*
affricate		"*ċ*"	~ "*ċ*"		
sibilant		*s*			
nasal	*m*	*n*			
liquid		*l*			
trill		*r*			
glide	*w*		*y*		(*w*)

One further development that seemingly took place after the loss of laryngeals but before palatalization was a development from PIE **wi* and **kʷi* to **wu* and **k(ʷ)u*. This, in turn,

preceded a possible development of PIE **e* to **i*, as it eventually merged with original **i* that had not undergone the aforementioned centralization. The full merger of these two vowels may have happened around the time of palatalization, or followed soon thereafter (2.4.2).

Some additional vowel changes that took place before palatalization involve the loss of distinctive vowel length. This process involved an unrounding of PIE **o* (potentially including **h_3e*) to a pre-PT **ë*, and a shift from PIE **eh_2*/**ā* to **å*. With the shortening of PIE **ē* (including **eh_1*) to **e*, perhaps original short **e* (including **h_1e*) had already merged with **i*. The very marginal early **ī* ends up as equivalent to PIE **ei* and may be represented as **iy*; in parallel, marginal **ū* may be represented as **uw* and PIE **eu* as **iw*.

Table 2.35: An overview of further early changes to the vowel pre-Proto-Tocharian vowel system and their relative chronology.

Sound changes		Chronology	Treated in
A1 *k^we, $k^w\bar{e}$*	> *ke, kē*	uncertain change	2.3.3.2
A2 *{w,k^w}i*	> *{w,k^w}u*	before C1.	2.4.2
B1 *o*	> *ë*	before B2.	2.4.4
B2 *ō*	> *o*		2.4.5.1
C1 *i, e*	> *i*	before D.	2.4.2
C2 *ei, ew*	> *iy, iw*		2.4.4
C3 *ī, ū*	> *iy, uw*	after A2.	2.4.4
C4 *Hi*	> *əy*		2.4.2
D *ē*	> *e*		2.4.4
E *ā*	> *å*		2.4.4

Table 2.36: Two stages of the pre-Proto-Tocharian vowel inventory before vowel length was given up (left) and after vowel length was given up (right).

post-PIE	front	central	back		front	central	back
high	*i ī*	*ə*	*u ū*		*i iy*	*“i”*	*u uw*
mid	*e ē*		*o ō*	>	*e*	*ë*	*o*
low		*a ā*				*a*	*å*

Palatalization involved a merger of the vowels **i* and **“i”*, and of **e* and **ë*. In the resulting system, these can be represented as new **i* and **e* respectively. Almost every consonant gained a palatalized counterpart. Only palatalized counterparts to **r* and **“ć”* cannot be established. Palatalized original labiovelars are identical to palatalized **k* (2.3.3.2), and these eventually merge with the palatalized **“ḋy”* (counterpart to **“ḋ”*) as TAB *ś*.

Developments that took place before palatalization are the shift from **o* to **ë*, the loss of vowel length, backing of certain PIE **i* before (labio)velars, assibilation and the depalatalization of palatovelars. Any change that can be dated before these developments also took place before palatalization, namely laryngeal colouring and the vocalization of syllabic resonants, the possible early delabialization of labiovelars next to PIE **e*, the change (whatever the exact nature, see 2.3.4.4) of PIE **d* to **"ḋ"* and the putative Tocharian Grassmann's Law. The phonological system resulting from the Tocharian palatalization is given in the tables below.

Table 2.37: An overview of sound changes that took place as phonological palatalization came into effect.

Sound changes		Chronology	Treated in
A CV_{front}	C^yV	before B.	2.4 esp. 2.4.1
B1 *i*, *"i"*	> *i*		2.4.2
B2 *e*, *ë*	> *e*		2.4.4

Table 2.38: The pre-Proto-Tocharian consonant inventory after phonemic palatalization.

pre-PT	labial	dental	palatal	velar	labiovelar
stop	*p* p^y	*t*	t^y	*k* k^y	k^w
		"ḋ"	*"ḋy"*		
affricate		*"c"*			
sibilant		*s*	s^y		
nasal	*m* m^y	*n*	n^y		
liquid		*l*	l^y		
trill		*r*			
glide	*w* w^y		*y*		

Table 2.39: The pre-Proto-Tocharian vowel inventory before (left) and after (right) phonemic palatalization.

pre-PT	front	central	back	+palatalization	front	central	back
high	*i iy*	*"i"*	*u uw*		*i iy*		*u uw*
mid	*e*	*ë*	*o*	>	*e*		*o*
low		*a*	*å*			*a*	*å*

After palatalization had taken place, we see *u*-umlaut of pre-PT **e* (from PIE **ē* and **o*) and *o*-umlaut of pre-PT **u* (from PIE **u*), which were both followed by the centralization of **u* to **ə* and its concomitant merger with earlier **i* (from PIE **i*, **e*, and secondarily

from vocalized PIE syllabic resonants and earlier centralized **i*). The *a*-umlaut which affected pre-PT **e* (from PIE **ē* and **o*) also took place after palatalization had run its course. The potential changes from pre-PT **o* (PIE non-final **ō*) to PT **a* and from pre-PT **å* (PIE **eh₂/ā*) to PT **o*, are uncertain, but they would also have taken place after palatalization and in that order (2.4.5.2).

Table 2.40: An overview of late vowel changes in pre-Proto-Tocharian, like umlaut and contraction.

Sound changes		Chronology	Treated in
A1 *e-u*	> *o-u*	before B.	2.5.1
A2 *u-o*	> *o-o*	before B.	2.5.2
A3 *e-a*	> *a-a*		2.5.5
B *i, u*	> *ə*		2.4.4
C1 *a.e*	> ? *å*		2.5.4, 5.4.1
C2 *a-å*	> *å-å*		2.5.4

Table 2.41: The late pre-Proto-Tocharian vowel inventory (left) developing into the Proto-Tocharian vowel inventory (right).

pre-PT	front	central	back	to PT	front	central	back
high	*i iy*		*u uw*		*i* (=*əy*)	*ə*	*u* (=*əw*)
mid	*e*		*o*	>	*e*		*o*
low		*a*	*å*			*a*	(*å*)

The character of a few consonants was further changed as well. The palatalized **sʸ* became a retroflex TAB *ṣ* [ʂ] while palatalized **kʸ* and **"ġʸ"* became a palatal sibilant *ś* [ɕ]. In the case of **kʸ*, this development may have taken place through an intermediary stage as a palatal affricate *[cɕ] or *[tɕ], so that the full palatalization of **tʸ* to TAB *c* [tɕ] probably occurred after **kʸ* had become de-affricated.

Several final consonants and consonant clusters had been lost before this point, but the exact dating in the relative chronology is difficult. This apocope did not affect the nature of the consonant inventory at any point and has therefore been left out of consideration in this overview.

Table 2.42: An overview of the remaining changes that took place in the consonants to yield the Proto-Tocharian system.

Sound changes		Chronology
A1 *sʸ*	> *ṣ* [ʂ]	before A2.
A2 *kʸ*, "*ḋʸ*"	> *ś* [ɕ]	before A3.
A3 *tʸ*	> *c* [tɕ]	
B "*ḋ*", "*ċ*"	> *tˢ*	

Table 2.43: The Proto-Tocharian consonant inventory.

PT	labial	dental	retroflex	palatal	velar	labiovelar
stop	*p pʸ*	*t tʸ*			*k*	*kʷ*
affricate		*ts* < "*ḋ*", "*ċ*"				
sibilant		*s*	*ṣ* < *sʸ*	*ś* < *kʸ*, "*ḋʸ*"		
nasal	*m mʸ*	*n*		*nʸ*		
lateral		*l*		*lʸ*		
trill		*r*				
glide	*w wʸ*			*y*		

The great majority of these changes took place before pre-Proto-Tocharian came into contact with Iranian languages in the relatively late period preceding the split of Proto-Tocharian, and perhaps even before it came into contact with a language related to the BMAC language (2.6). This means that the developments that turned Proto-Indo-European phonology into what we recognize as Tocharian phonology happened in the deeper past.