

## Pāli Variants: A Typology (Part I)

*Bryan G. Levman*

**ABSTRACT**—There are thousands of variants in the Pāli canon. This paper examines the reasons and processes by which they arise with many examples. There are eight major factors involved: 1) The nature of the source transmission, i.e. the different dialects and/or *koiné* in which the Buddha taught prevalent in north India (§2.1). 2) Natural, diachronic language change over time which tended to simplify by, for example, voicing or eliminating unvoiced stops, replacing aspirated stops with aspirates only, etc. (§2.2). 3) Sanskritizations which acted to “restore” Pāli to its putative “original” OI form (§2.3). 4) Linguistic diffusion from neighbouring IA languages and dialects where one dialect might interact with and alter another (§2.4.1); linguistic diffusion from autochthonous languages (§2.4.2); linguistic diffusion due to bilingual speakers of Dravidian and Middle Indic (MI), whose native language was non Indo-Aryan (IA), and who had to adapt a foreign phonemic system into the MI transmission (§2.4.3); linguistic diffusion due to foreign word borrowing from non-IA languages into MI, confusion due to the transcription of these “foreign” words, and lack of knowledge of the Pāli language (§2.4.4). 5) Oral transmission ambiguity and errors due to memory, recitation and auditory issues (oral/aural);

a pot pourri category subsuming all seven of the above processes (§2.1–§2.4.4) and additional variants which are purely sonic in nature (§2.5). 6) The introduction of explanatory glosses into the main transmission, a practice which was going on probably from the time of the Buddha himself (§2.6). 7) Orthographic variation in spelling and copyist errors, sometimes due to the influence of the copyists' native language, whose phonetic/phonemic system was foreign to MI (§2.7). 8) Harmonization and standardization of the canon by the later grammarians (§2.8). All of these factors introduce changes into the *Buddhavaṇṇa* which are preserved in the canon (or can be reconstructed from it). A series of examples from the *Theragāthā*, one of the oldest of Buddhist works, is given to illustrate these processes.

**KEYWORDS:** Pāli language, Dravidian language, Munda language, Sanskritization, *koiné*, oral transmission, diachronic change, dialect change, Theragāthā

## 1. Introduction

There are thousands of variants in the Pāli canon. This paper is about the types of variants that are found and the reasons and processes by which they arise. Variants arise in the transmission of the *Buddhavaṇṇa* from one generation to another in both the oral and written tradition, according to the Buddha's teaching of *anicca* ("impermanence") which affects all phenomena. As is well known, according to tradition, the teachings of the Buddha were memorized by Ānanda as he taught, and later were transmitted by Ānanda to the core of five hundred monks at the First Council after the Buddha's *parinibbāna*. Upāli provided the same service for the Vinaya.<sup>1</sup> As we learn from Ven. Purāṇa in the Vinaya (see below §1.2), each *bhikkhu/bhikkhunī* was in effect his/her own tradent of the teachings, memorizing and passing on the *Buddhadhamma* to the public and to other religious as he/she remembered it. There was also a

---

<sup>1</sup> To avoid confusion, all Pāli words discussed here are italicized (except place names, and words borrowed into the English language and found in the Oxford English Dictionary), including proper names, when their etymology is being discussed, since most of the proper names are equivalent to or derived from ordinary Pāli words which would normally be italicized.

specialized group of *bhāṇakas* (“reciters”) appointed by each Sangha (as the original Sangha diversified into many immediately after the Buddha’s passing), who specialized in one or more *Nikāyas* or the Vinaya and transmitted the teachings to their successors for approximately three hundred years, until they were written down in Sri Lanka in the first century BCE. Along the way many changes were introduced, especially in the oral phase, most probably involuntarily, as the tradents and *bhāṇakas* tried to preserve the Buddha’s teachings to the best of their abilities.

The PTS edition of the canon is usually based on just two manuscript traditions, the Sri Lankan and the Burmese, and usually includes the *mūla* text and the commentary where available; occasionally the Thai/Khom lineage has been consulted if manuscripts, or printed editions, were available; the PTS DN, for example, sometimes references a Thai manuscript in “Kambojian” script marked “K” in DN 2: viii.<sup>2</sup> Even when only two traditions are employed, variants are considerable. In the new Dhammachai Tipiṭaka project, which has so far produced a pilot edition of the *Silakkhandhavagga* of the *Dīghanikāya* (2013), there are thousands of variants compiled from Burmese, Sinhalese, Khom (central Thai) and Tham (northern Thai) manuscripts (Levman 2016b).

### 1.1 Pāli recensions

There are four primary recensions of the Pāli canon: Sinhalese, Burmese, Central Thai (Khom/Cambodian) and northern Thai (Tham); there is also a Laotian canon in Pāli. The PTS or European edition (Ee) a composite, diplomatic version, may be considered another recension. The Sinhalese is considered to be the oldest, (but not on that account necessarily the most accurate, because of the very complex interaction with other parts of south-east Asia, especially Burma), while the others are thought to be younger.<sup>3</sup> In addition, we have fragments of the canon composed in other MI languages such as Gāndhārī and Buddhist Hybrid Sanskrit (Edgerton 1953/98), which may be viewed as a Prakritized Sanskrit or a Sanskritized Prakrit, depending on the extent of change (e.g. the Patna *Dharmapada*, the *Mahāvastu*, etc.).

<sup>2</sup> For a detailed discussion on text critical practices in PTS editions, see, for instance, Chris Clark’s PhD thesis (2015: §3.6, pp. 65–80).

<sup>3</sup> However, note that in the Thai tradition, coeval with Mahinda bringing the canon and commentary to Sri Lanka, another group brought the *Buddhadhamma* to Suvannabhūmi (Thailand) in the third century BCE (Kusalasaya 1965/2005: 4).

From Sri Lanka, monks exported the teachings of the Buddha to south-east Asia under a “highly complex system of intertwining historical, geographic, political, and cultural circumstances” (Keown 2003/2004: 275), which naturally involved various to and fro movements between south-east Asian polities. Although the canon was allegedly written down in Sri Lanka in the first century BCE, its oral transmission continued for centuries afterwards and variants continued to accrue, through inexact memorization and recitation, faulty or idiosyncratic recollection or deliberate attempts to improve the text, through dictation process errors (such as faulty pronunciation by bilingual speakers or faulty hearing by bilingual scribes), manuscript copying faults and many more issues discussed below (§2.1–§2.4.4).

An attempt was made to fix the commentary to the canon (and to that extent, the canon itself) in the fourth/fifth century by Buddhaghosa, but changes continued to accrue and were harmonized in part by Aggavaṃsa in his grammar of the twelfth century CE in Burma. The oldest Pāli manuscript we possess (part of the Vinaya) is dated to the ninth century CE (von Hinüber 1991) and continues to show numerous variants from the “standard” Sinhalese, Burmese or Thai recensions. Even after Aggavaṃsa, thousands of variants are found in the manuscripts utilised by the Dhammachai Tipiṭaka Series, which generally are to be dated between the seventeenth and nineteenth centuries. No attempt is made here to unravel the canon’s complex history of exportation and transmission across south-east Asia. There is, however, a rationale underlying this paper that a good deal of the variants are due to the oral transmission process and the various phonological and phonetic variations which are inherent in it. As noted, the manuscripts are not old, but the age of the manuscript has nothing to do with the age of the content; the inference here is that many if not most of the significant phonological and phonetic variations discussed below stem from the oral practices (memorization, recitation and dictation) of the *bhāṇakas*, and that many of them go back to the earliest Buddhist traditions, although it is impossible to date them, except (sometimes) relatively. Presumably, if manuscripts had existed at the time these variants occurred, they would have constrained and prevented many of the sound-type *anicca* witnessed here; I am not suggesting that all phonological/phonetic variants occurred in the absence of manuscripts (which, especially in CE times would not be true), but that the oral/aural tradition of memorization, recollection, recitation and interpretation (and later, dictation) exercised sway even over manuscripts; before the

manuscript tradition emerged, it was the only source of transmission. There are of course variants introduced due to the writing process itself (§2.7), but these are of secondary importance to the oral transmission. See “A Note on Theragāthā Recensions” below, under Works Cited.

## 1.2 The variant process

Although variation is a natural process and occurs at every stage in the chronology of the canon, this paper is principally concerned with variation in the oral transmission process (which includes of course the later aural scribal rendition of the oral transmission). It is generally believed that the Buddha taught in many different dialects (Norman 2002: 111 and also perhaps in non-IA languages as well; Levman 2023a), and that these teachings were memorized by Ānanda and others tradents and transmitted orally through generations of *bhāṇakas* whose job was to memorize the *Buddhavacana* and preserve its integrity. The Buddha himself may have used two different forms of one word to communicate to different dialect speaking audiences, or more likely, the *bhāṇakas* interpreted the transmission they received in terms of the phonemic structure of their own dialect or language, which resulted in variation. The various *reasons* for variation are discussed in detail below; the *fact* of variation is indisputable and no doubt in part paralleled the branching of the teaching into various sects, which happened from the earliest times, certainly no later than the Second Council. Considering Ven. Purāṇa’s statement, who declines to join the First Council recitation because “he would bear in mind the Buddha’s teachings as he heard it,” not as the Council prescribed,<sup>4</sup> variation probably occurred right after the Buddha’s *parinibbāna*. There is of course the possibility of two separate transmissions from the Buddha in similar contexts,<sup>5</sup> but when the words are phonologically related, a much more parsimonious inference is that different *bhāṇakas* and tradents

<sup>4</sup> “Your reverences, well chanted by the senior bhikkhus are *Dhamma* and discipline, but in that way that I heard it in the Lord’s presence, that I received it in his presence, in that same way will I bear it in mind” (Horner 1938–66/2001–07: 2396–97). Se Vin 7, 389<sup>10–12</sup>; PTS Vin 2, 290<sup>6–8</sup>: *susaṅgīt’ āvuso therehi bhikkhūhi dhammo ca vinayo ca, apica yath’ eva mayā bhagavato sammukhā sutam sammukhā paṭiggahitam tath’ eva aham dhāressāmi ti*.

<sup>5</sup> A suggestion made to me by Stefan Karpik (pers. communication). Possible examples of which are discussed in Levman 2025b: §4.3.18, the *Pāsādikasutta*, where the verb *sarissāma* occurs in Se (“we will/can go”) while *passāma* (“we see”) occurs in the other three recensions and both are possible. They are not phonologically related. See also examples §4.3.21c and §4.3.22 in the same study.

heard certain words differently for the reasons discussed below, and since both forms were suitable to the context, they were preserved from the earliest times. This paper gives several examples of phonologically similar variants from the *Theragāthā*; Levman 2025a (xvi–xvii) gives several other examples from the *Dīgha Nikāya*<sup>6</sup> where one may conclude that the fact of the variants’ preservation points to different, independent *bhāṇaka* traditions, which were retained despite the fact that they contradicted some manuscripts.

The Buddha himself recognized that variant teachings would beset the Sangha after his demise and according to tradition, he provided a means of establishing their authenticity in the *Catu-mahāpadesa-kathā* (“Discourse on the four great precedents”) of the *Mahāparinibbānasutta*, shortly before his *parinibbāna*. He instructed that the “words and letters/phrases” (*pada-byañjanāni*) of such teachings were to be compared to the *sutta* and *Vinaya* and “if they did not fit,” (*na c’ eva sutte otaranti, na ca vinaye sandissanti*) they were to be rejected. But the case where both words suited the context was not discussed. And of course there is the question of *which* suttas and *Vinaya* were to be consulted? It is highly unlikely that such a canon existed before the Buddha had passed away, where there was a “standard” to consult; the whole *mahāpadesa* story is probably a later interpolation. Nevertheless it shows that the Sangha recognized the fact of variants from a very early stage and provides an authority for preserving both variants where they “fit” the context.

### 1.3 Objections

One might object that this is all surmise, and it is impossible to determine whether the variants discussed here can be traced back to the early oral

---

<sup>6</sup> For example *Se sutvā* (“heaving heard”; transmitted as *suttā*, *sutvā* being a Sanskritization) vs. *Be/Ce/PTS suddhā* (“pure”) in the *Mahāgovindasutta*; *Se suddha-āvāsa* (“pure abode”) vs. the other recensions *satta-āvāsa* (“abode of beings”) in the *Mahāpadānasutta*. These words are also confused in the *Brahmajālasutta* commentary (Sv 1 87<sup>11–12</sup>), which notes that *kosiya-sutta* (“silk thread”) interchanges with var. *suddha-koseyya* (“pure silk”) in the *Vinaya*. *Se kammaniya* (“skilful”), *Ce, PTS (kamaṇīya, “beautiful”), Be khamāṇīya* (“bearable”) in the *Mahāśudassanasutta*. *Se* and *PTS amūlha-paṇha* (“successful in questioning”) vs. *Be/Ce amūlha-paṇṇa* (“unconfused wisdom”) in the *Sakkapaṇhasutta*; or *Se bandhanti pārā pārāṃ*, *Be bandhanti apārā pārāṃ*, *Ce bandhanti orā paraṃ*, *PTS bandhanti aparāparaṃ* in the *Mahāparinibbānasutta* (re: building a raft to go from the near to the far shore). The close phonological similarity of all these forms “proves” that they had to come through the oral tradition, and probably are pre-ms, as if there were a ms to refer to, presumably they would have been harmonized to one standard.

transmission of the canon. There are thousands of manuscripts available in each tradition, none earlier than the 9<sup>th</sup> century CE (von Hinüber 1991), and given the complexity of the manuscript situation and the approximately 2500 years that have elapsed since the death of the Teacher, it is impossible to establish a critical edition of what the earliest or “original” transmission might be. This objection has some validity and Buddhist scholars have attempted to ascertain the authenticity of early Buddhist transmissions (of which the *Theragāthā* is one) for over a century;<sup>7</sup> the general consensus among scholars who specialize in the transmission of early Buddhist teachings is that the transmission is, by and large, authentic and goes directly back to the Founder (Rhys Davids, 1877: 15–17; Wynne 2005: 35–66; Sujato and Brahmali 2014: §3.7 and many references therein). The conclusion is based on many factors including the internal consistency of the Pāli canon itself and the lack of any significant additions to it in the post-Asokan (mid-third century BCE) and Sri Lankan periods (first century BCE), with the exception of minor emendations and harmonizations (Norman 1983: 5; Norman 2002: 140; Wynne 2005: 65–66). Anālayo (2012: 246) notes that the canon was “fairly closed” by the first century BCE and argues, along with Rhys Davids (1911: 174), Geiger (1916/2004: 7) and Pande (1974: 16) that the absence of the mention of King Asoka in the canon points to its completion prior to his reign, that is, the mid-third century BCE (p. 243). Von Hinüber (2006: 202) makes a similar observation with regard to the lack of mention of Pāṭaliputra in the *Mahāparinibbānasutta* as the capital of the Maurya empire, suggesting that the text is likely pre-Mauryan. Epigraphical confirmation that some form of canon existed in Asokan and probably pre-Asokan times is provided by the Asoka Bhabra Edict, which mentions several canonical works by name and by almost coeval epigraphical evidence at the Sanchi and Barhut *stūpas* where the terms *dhmma-kathika* (“preacher of the Dhamma”), *peṭakin* (“one who knows the *piṭaka*”), *suttantika/suttantakinī* (“a man/woman who knows a *suttanta* by heart”) and *pañca-ṇeṇāyika* (“one who knows the five *nikāyas* by heart”) are inscribed (Bühler 1894: 92; Rhys Davids 1911: 167–68). A comparison of the Pāli with the oldest surviving Buddhist manuscripts in Gāndhārī, dating from as early as perhaps the first century, provides confirmation of this hypothesis of the antiquity of the Pāli canon: in

<sup>7</sup> “Early Buddhist Texts” are usually taken to include the first four *Nikāyas*, the *Sutta Nipāta*, *Dhammapada*, *Thera-* and *Therīgāthās*, the *Udāna* and *Itivuttaka*, the *Pāṭimokkha* and other parts of the *Vinaya* (e.g. *Khandhaka*); for a useful definition see Sujato and Brahmali 2014: §0.1.

substance and often word-for-word, they agree with the Pāli. Within the Pāli tradition itself over 98% of the material is the same amongst the four main recensions.<sup>8</sup> This is not to maintain that there was an “original” Buddhist transmission, as a more likely scenario is that there were multiple *bhāṇaka* traditions from the earliest times, even if the variation was quite small (but nevertheless significant).

#### 1.4 Reasons for retention

Where there is retained variation (aside from minor orthographic differences), their presence in the canon either argues for a retention from the very earliest transmission layers, where the variant was not understood or could not be rationalised, or represented variation in the early teachings and therefore was preserved by the different *bhāṇaka* traditions as a potentially valid reading. Eschewing “scribal errors” as a sole reason, Norman (2002: 113–15) cites this possibility—tradents not understanding the derivation or meaning of a word—as one of the principal causes of variation retention. This is especially applicable to technical terms, where, in some cases, there are dozens of variant forms that have been preserved (see, for example, a partial list of variants on the word *pācittiya* in Levman 2023a: 90, n. 59).

Variant retention was not idiosyncratic to the Buddhist *bhāṇaka* tradition. Bloomfield and Edgerton (1932/1979: §1) report thousands of phonetic variants in the Vedic oral tradition, “accompanied by or resulting in lexical and morphological change at the same time.” They note that rime and phonetic confusion constitute the “prime motive” in the variation, but that lexical change was a “real fact of the tradition of a given school” and that these secondary readings have “their own right to exist ... the genuine readings of their respective schools” (§3). Most if not all of the variants found in the Buddhist transmissions are also found in the Vedic tradition. Bloomfield and Edgerton summarize the findings of their study *Vedic Variants* as follows:

There is, however, one kind of interchange which runs as a

---

<sup>8</sup> See Levman 2025a (introduction, page xii ) which compares the Thai recension to Be, Ce and PTS and concludes that there is a 98.5% concordance between the four. Of the 1.5% which consists of variant readings, only 0.005% represents “significant” variants, which means 0.5% of the variants readings.



thread through Vedic tradition, and which is so important that it should receive special treatment and emphasis. We refer to shifts which suggest possible dialectic influence from popular speech, by their resemblance to the phonetics of the later Middle-Indic dialects ... The large mass of variants of this kind, clearly pointing to extensive influence of Middle-Indic phonetics in the earliest periods of the languages seems to us one of the most important results of this volume of the *Vedic Variants* ... We find, all in all, the most definite proof of phonetic changes not only in the direction of Prakritism, but also (no less interesting) in the reverse direction, ‘hyper-Sanskritism,’ which latter indicates a rather definite consciousness, on the part of the handlers of the texts, of the antithesis between the phonetics of the high speech and of the popular dialects (§20).

Such variants as consonant lenition, interchange of aspirate and non-aspirate stops, or replacement of the latter with aspirate only, interchange of labials *m* and *v*, reduction of sibilants, assimilation of conjuncts, etc. are provided in detail; often the variant is purely phonetic but many have lexical (semantic) import as well. Each case has to be individually considered; in the Buddhist tradition the situation is equally complex. Sometimes the variant was preserved because it was not understood or the tradent considered it a potentially valid dialect form. In other cases the variant was understood, but it was not clear which variant was “correct,” so both were preserved, either by the individual tradent or *bhāṇaka*, or by different *bhāṇaka* traditions. Or the variant was simply a mistake, a “slip of memory” by a senior monk, which was not corrected and transmitted to the next generation (Anālayo 2022: 36). Often the variant seems to be purely phonetic; but often there can be a change in meaning. In the *Mahāgovindasutta* Se DN 2, 278<sup>2</sup>; PTS DN 2, 244<sup>11</sup>), the Thai recension has *sutvā* (“having heard”) vs. Be/Ce/PTS *suddhā* (“purified”), which could only have occurred and been preserved in an oral environment. Both were understood, but the Thai tradition preserved *sutvā* as *Buddhavacana*, while Be/Ce preserved *suddhā*. Since we know that *sutvā* was in fact a later back-formation/Sanskritization for MI *suttā* (von Hinüber 1983/2005: 7–8/190–91), it makes the argument for confusion during the earlier oral tradition even more compelling; that is, the Thai *sutvā* was probably transmitted as *suttā* and the difference between the two *suttā/suddhā* would not be apparent to a Dravidian bilingual who had no phonemic voiced stops or aspirates in his native

language. This is also useful for dating; as von Hinüber notes (*ibid.*: 8/191), because of the presence of both forms in the older part of the *Milindapañha*, the *Gāndhārī Dharmapada* and the *Buddhavaṃsa*, “it may be assumed that the restoration of the absolutive ending in *-tvā* had indeed begun before the Theravāda canon was written down in the first century B.C. [...] the roots of the shaping of this particular linguistic form of Pāli must reach back into the last centuries B.C.” OI conjunct absolutes were already assimilated by Asoka’s time (e.g. *-tvā* > *-tu* in *Shāhbāzgarhī*, *Kālsī*, *Dhauḷi* and *Jauḡaḡa*; Hultzscht 1924/1969: lxxxiv, xcvi, cxi), and may well go back to the time of the Buddha. As Norman points out (2002: 148) “these features may have formed part of some of the very earliest texts. They may indeed go back to the original language or languages of Buddhism ....”<sup>9</sup> In the examples below the reader will find many instances of variant readings preserved where both make sense in the context—probably one of the main reasons for their preservation—while in other cases they were preserved because, although they were not understood, they were deemed potentially valid readings.

A reasonable approach is to examine the manuscript variants (evidenced in the following examples) together with other supporting factors, ascertain where the evidence points, and evaluate the cogency of the conclusions; that is, to adhere to the methodology of standard historical, comparative linguistics (see Norman 1997/2006: Chapter 4 and throughout; Levman 2016a: 5–8). Rejecting this hypothesis based on the lack of a critical edition and the lateness of the mss on which the printed recensions are based would imply dismissing all of the work of historical phonology done by scholars like Norman and von Hinüber in the last few decades—indeed, all

---

<sup>9</sup> How far awry this can go in the oral transmission is illustrated by one (not uncharacteristic) example of oral confusion from the *Nāmasutta* (PTS SN 1, 39<sup>3-6</sup>) of the *Samyutta Nikāya*, *Sagāthāvagga*, one of the very oldest of the early Buddhist transmissions (von Hinüber 1996/2008: §74, “some parts of the *Sagāthāvagga* seem to be very old, actually very near to Vedic texts”). Here we have six phonologically/phonetically related variants for the word *addhabhavi* (PTS, Be), *anvabhavi* (Se, Ce), *andha-* and *andabhavi* (Se vars.), *aṭṭhabhavi* (Sī v.l.) and *adanvabhavi* (Cambodian v.l.). Bodhi (2000: 130) translates “What has weighed down everything?” taking *addha-bhavi* as an aorist of *abhi + bhū* “to master, command of”, while Se has *anva-bhavi*, the aorist of *anu + bhū* (“to enclose, embrace”) which requires no change of *abhyabhavi* > *abbhabhavi* > *addhabhavi* (cp Oberlies 2019, § 76). Other variants are equally cogent, like *addha-bhavi* (“being in the past”) which is how the *ṭikā* takes it (Be Spk-ṭ: 1, 132: *kāmaṃ pāḷiyaṃ atīta-kāla-niddeso kato, taṃ pana lakkhaṇa-mattaṃ*, “surely in the *mūla* this refers to a past time, but is merely a characteristic”) or *andha-bhavi* (“being blind”), or *aṭṭha-bhavi* (“carrying the meaning”) or *anda-bhavi* (“being a fetter” < OI and “to bind,” *ānda* “one who makes fetters”) or *adanva-bhavi* (?). Except for this last (which may be a typo) all of these are possible in the context, which is presumably why they were retained and transmitted.

work on the historical linguistics of the canon, which seems hardly plausible.

These “supporting factors” are also for the most part pre-canonical (Dravidian language influence, bilingualism, the influence of other dialects and/or *koiné*, etc.), which strengthen the argument presented here as to the antiquity of the variants. The variations preserved in the transmission, or “anomalous forms” as Norman calls them (2002: 141), or “oddities” as Wynne calls them, “are likely to be significant—not produced by the random variation of an oral tradition, but by causes that in theory can be discovered” (Wynne 2004: 124). As Norman intuitively, “We can also assume that traces of this linguistic diversity were retained when the Buddha’s sermons, which had been preached in different areas, were first collected together” (2002: 145). This might refer to the retention of two different dialect forms of a word in two different passages of the canon, or it might refer to the same word in the same teaching, either spoken by the Buddha in a different form at a different time, or differentially remembered by the Buddha’s bhikkhu-tradents and/or *bhāṇakas*.

Another important point is the presence of variants in the commentaries, many examples of which will be shown below from the different commentarial recensions. As Norman and Endo have noted, parts of the commentary may go back to the time of the historical Buddha where the teacher and/or his immediate disciples tried to clarify certain obscure points in the doctrine (Norman 1997/2006: 149; Endo 2013: 5). If these were simply “scribal errors” then the commentators would have corrected them. But not only are variants preserved in different manuscripts, but the commentary itself embeds several alternate readings within its own transmission, with the expression *ti pi pāṭho*, i.e. “so is another reading”; a quick search shows that there are over six hundred of these in the commentaries, again attesting to the antiquity of the variants. Some of these may date from the earliest transmission, others to the time of Mahinda’s transport of the canon and commentary to Sri Lanka in the third century BCE (Endo 2013: 20 calls this SRIOC, “Sinhala Rendition of the Indian Original Commentaries”), others to later versions of the commentary formulated in Sri Lanka.

Although some parts of the commentaries may go back to the *mūla* transmission, they are by and large later than the *mūla*. A peculiar idiosyncrasy of the Se variant transmission is the large number of cases where the *mūla* text is different from the other recensions, but the Se commentary preserves the variant readings of Be and Ce in the headword. The Thai editors were

obviously aware of the variant forms (as they were preserved in their own commentary), but retained the Se reading as a valid Thai tradition, differing from the others, despite the commentary. The *mūla* text here clearly pre-dates the commentary.<sup>10</sup>

Charles Hallisey (2024: 92) notes that “one of the services of the commentaries in the Pāli canonical texts is to establish the text, and careful attentiveness to textual differences and variants eventually became a standard component in the normative practices of Theravādin Buddhist commentators.” The expression *ti pi pāṭho* was only used for variants which were considered “significant,” that is, “worthy of consideration, reflection and preservation”; other variants which were not significant were deemed “careless” (*pamāda-pāṭha*) in the commentaries (p. 92). Each (of the former, significant) variants has its own narrative, enhancing and potentially revising our knowledge of the history of the Pāli language and its transmission, and suggesting heretofore unimagined possibilities (99–100). Hallisey does not attempt to deal with the very complex question of the chronology and interweaving of oral and textual transmission, other than noting that the Pāli was “unitary without being uniform,” a general condition which would simultaneously create a comfort with received textual variants ...” (p. 96). The word “textual” here is not accurate, as Hallisey acknowledges (referring to Gāndhārī manuscripts) that “both oral transmission and written transmission co-existed and interacted with each other in multiple ways and awareness of it enhances what we know about other oral aspects of Pāli textual culture” (p. 101).

Richard Salomon (2024: 20) has also reflected on the importance of variant readings in the canon, noting that “the variants were already present from the beginning of the textual tradition ...” and that the “variability *precedes* [*italics in original*] the fixity and stability of the received, canonized texts of the Bible or of the Buddhist sūtras ... discoveries of manuscripts have made it clear that they involved the suppression of innumerable variants, any of which probably had an equal claim to originality.”

A second objection might relate to the manuscripts scribal tradition, where the canon was copied between various Indic (Brāhmī, Sinhalese, Karoṣṭhī—

---

<sup>10</sup> For example, Levman 2025b: §4.2 has 9 examples of Se *mūla* differing from the Se comm. (in the DN) and Levman 2025c, has 15 examples of the same phenomenon in the MN. These instances are very useful for establishing time-lines, showing *mūla* forms as preserving forms different from and pre-dating the commentary.

actually a Semitic script in which Gāndhārī was written) and other south-east Asian scripts (Burmese, Lanna, Khom, Thai). Certainly mistakes in the rendering of one script in another did introduce variation into the canon (discussed below §2.7); especially problematic in some south-Asian scripts where certain letters were very similar to each other (e.g. Sinhalese *ha* and *bha*; Norman 1987: 30). These changes are all post-oral, however, largely orthographic and are by and large not significant in terms of the change of the transmission's meaning. They should be the subject of a separate study of *Theragāthā* manuscripts, which looks at the similarity between letters and the effect this may have had on variants in the written transmission tradition. As noted above, this study argues that the significant phonological variants mentioned can be traced back to the oral tradition, which is the main subject demonstrated below.

### 1.5 Caution

Historical linguistics is the study of the history and evolution of language by comparing cognate forms and, where possible, tracing back their history to a shared, common ancestor. It is considered a descriptive science, because of its rigorous methodology and the potential for falsifying its result; but like its sister, the descriptive science of palaeontology which proposes cladistic trees of species based on shared characteristics, its results cannot be considered “proof,” but only best inference. So, while we can establish that phonologically cognate forms evolved from a common ancestor and in some cases propose what that ancestor might be, we cannot consider the result to be definitive; it is a hypothetical reconstruction, which is why the ancestral form is prefixed by an asterisk (\*-), as are all the proto-Indo-European roots reconstructed by comparing cognate words in the various daughter languages. Timelines can only be considered relative, not absolute.

The sine qua non of language is sound, and sound like all other elements changes over time, due to the factors discussed below (§2.1–2.5). There are thousands of variants which can be attributed to the manuscript copying traditions; the kind of changes that occur here and their causes have been well established in biblical text-critical studies (Tov 1992/2001). This paper is principally concerned with changes that may be attributed to the oral tradition by examining phonologically cognate forms in parallel passages of the different recensions. They are not copyist or scribe “mistakes”; presumably, if they were, they would have been corrected long ago. They are, for the most part, valid readings preserved by the different *bhāṇaka* traditions, either because they were *Buddhavacana*, or not understood and preserved for that

reason. Although one cannot “prove” that they pre-date the written manuscripts, it is a logical and cogent inference based on the data at hand. By and large they are not minor orthographic peculiarities and idiosyncrasies but significant variants which often affect the meaning of the passages and/or etymology of the words.

To return to our main subject: why so many variants? Here follows a short summary of the principal causes.

## 2. Principal causes of word variation in Pāli

### 2.1 The Prakrits and *koiné*

The linguistic situation at the time of the Buddha was very complex. In addition to the Vedic language (which was exclusively reserved for the religious purposes of Brahmanism), there were several dialects in use in the north-west, west, central and eastern regions of north India. The Asokan edicts, inscribed 150–200 years after the Buddha’s ministry, provide detailed evidence of the linguistic structure in the different regions. Scholars believe the Buddha taught in an eastern Prakrit but his teachings were also restated into other Prakrits, either by himself or his successors, resulting in transmission errors. In addition there is a lot of linguistic evidence which provides a convincing argument for the existence of a cross-India *koiné*, lingua franca or common language that was used to simplify inter-dialect and inter-language communication, in trade and government (Geiger 1916/2004: 3–4; Smith 1952: 178; Norman 1989: 35; Levman 2016a, 2019a); this too led to potential confusion when the simplified word was interpreted by the receivers. Based on the linguistic evidence, it is also reasonable to assume (see below §2.4.2–2.4.3) that the Buddha spoke in one or more indigenous languages, or—since it is not possible to definitively identify the historical Buddha’s teachings—that his early tradents did so.<sup>11</sup> See discussion on the language that the Buddha and early tradents spoke in Levman 2023a: 60–63. As Norman has opined (2002: 111), it is a reasonable assumption that many of the variants are preservations of dialect anomalies due to the Buddha’s sermons being preached in different dialect areas.

---

<sup>11</sup> I omit from discussion here the whole question, recently brought up again by Drewes (2017), of the existence of a historical Buddha. This has already been answered by von Hinüber (2019), Wynne (2019) and Levman (2019b); Drewes’ recent response (2022) does not advance his hypothesis. His suggestion, for example, that “it would be very difficult to get to the idea that early Buddhism focused on either the quest for religious experience or the practice of meditation” (*ibid.*, 20) is ludicrous to anyone who has actually read the Pāli canon, where meditation is an integral part of the three-fold training (*sīla, samādhi, paññā*).

## 2.2 Diachronic language change

In India we can trace a history of change from Old Indic (OI) to Middle Indic (MI) and New Indic. Most of the changes from OI > MI were already present in the Vedas in inchoate form (Bloomfield and Edgerton 1932/79: §20), and they continued and accelerated in the centuries before and after the Buddha, often under the influence of indigenous language speakers. These took such forms as: the weakening of conjunct consonants > geminates; the weakening and loss of intervocalic stops; the levelling of all sibilants; the loss of aspirated stops and their change to aspirates only; the interchange of labials and glides, etc., to name a few of the major changes. These changes are very old and most are preserved or foreshadowed in the Asokan dialects, dating the anomalous features in Pāli to pre-Asokan times (Levman 2010; Norman 2002: 111).<sup>12</sup> Again, this introduced uncertainties into the transmission. Did a geminate (e.g. *-tt-*, as in *satta*) indicate the conjunct *-pt-* (*sapta* “seven” or *śapta* “cursed”, since *s-* and *ś-* > *s-*), *-kt-* (*sakta* “clinging” or *śakta* “competent”), *-tv-* (*satva*, “warrior”) or *-tt-* (*satta*, “seated”)? The meaning was often clear from the context, but not always.

## 2.3 Sanskritization

The priestly language (OI) and a form of the vernacular Prakrits have co-existed since Vedic times (Wackernagel 1896: xvii-xviii), along with the tendency towards Sanskritization, that is, the action of purifying the ritual language of Prakritic influence (Edgerton 1930: 27; Bloomfield and Edgerton 1932/1979: §20, this article §1.4). The ultimate hegemony of the Sanskrit language as a pan-Indic phenomenon began in earnest after the time of the Buddha, as early as the Asokan edicts when it became increasingly important as a unifying cultural and political force in India; from the 3rd century BCE on, Pāli has a lot of words that have been Sanskritized, their Prakrit form “restored” to their earlier OI form (Norman 1983: 4–5; 1997/2012: 95–112; von Hinüber 1982: 138, 1996: 190; Salomon 1998: 84; Levman 2020; Levman 2021a: 266–67; 290–91; for a summary, see Levman 2023b: 6–8). Sometimes the transmitter who interpreted the Prakrit word did not do so correctly. Pāli, for example, retains most intervocalic stops and aspirated stops whereas other Prakrits do not; some of these are interpretative restorations, and these interpretations

<sup>12</sup> For example, the various ambiguities we see between intervocalic voiced and unvoiced stops in the canon (Levman 2014: 475–94), or the voicing of *-tṭh-* > *-ḍḍh-* etc., all of which we find in the Asokan edicts (Norman 2002: 142).

are subject to variations and error. Pāli has many anomalous forms where a simplified Prakrit form is preserved alongside a Sanskritized form: e.g. *pahā* alongside *pabhā*, “shining” (Levman 2021a: 285); *khāyita* alongside *khādita*, “eaten” (PED);<sup>13</sup> *goyāna* alongside *godāna* (proper name); *avāyesi* alongside *avādesi* (Levman 2023b: 3, note 2); *āchāya* (“gift”) alongside *āchāda* (BHS §2.32); words ending in *iya*, alongside the same word in *-ika*, etc., Lüders 1954: §133–38); and many forms where a simplified Prakrit form has been interpreted in variant ways: e.g. *vijita/vidita* pointing to the existence of an earlier form *viyita* (Levman 2023b: 3, 11); *virato/virajo* < *viyato*, (Norman 1980: §3.2). These date from the earliest layer of the transmission, at least from as early as Asokan times (Norman 1997/2006: 125–26) or earlier, as Bloomfield and Edgerton have provided evidence that Sanskritizing of Prakrit was taking place as early as Vedic times (Bloomfield and Edgerton 1932/1979: §20).

### 2.4.1 Linguistic diffusion (synchronic change) from neighbouring

#### IA languages

Different Prakrit speakers pronounced words in different ways. Some tending to voice intervocalic consonantal stops and others tended to devoice them. Some

<sup>13</sup> While it has been suggested that *khādita* may be a retention of an Old Indic form or dialect variation (rather than a Sanskritization), it is clear from the appearances in the canon that *khāyita* pre-dates *khādita* and that the latter is therefore a restoration. DPD reports 19 instances of *khāyita* in the first four books of the Sutta Piṭaka and three in the early *Cūḷa-* and *Mahāvagga* of the Vinaya, while for *khādita* there is only one occurrence in the AN (none in the other three *Nikāyas*) and one in the *Mahāvagga*. Most of the occurrences of *khādita* occur in the commentary (68 in total) as opposed to 20 appearances in the commentary for *khāyita*. Total statistics per DPD are (for Be):

Appearances	Mūla	Aṭṭhakathā	Ṭīkā
<i>khāyita</i>	57	20	40
<i>khādita</i>	11	68	30

See DPD for breakdown by *Nikāya*. Mallik (1970) reports a number of cases which he calls “Sanskritisms” and treats as retentions, but by the same process many of these can be shown to be restorations, where the Sanskritized form post-dates the Prakrit (e.g. *citta*, *citra*, “bright-coloured”, *bhadda*, *bhadra* “good, excellent”; *uddaya*, *udraya* “result, profit, outcome”; *udraya* perhaps a wrong reconstruction of *udaya* “rising, profit, outcome, result” per Cone). One form that is ambiguous is *uttasta/uttrasta*, “frightened” as both conjunct (*-tr-*) and geminate (*-tt-*) forms of this verb (*uttasati*, “to be frightened, terrified”) occur in early works.



omitted the stops altogether or replaced them with a -y- glide,<sup>14</sup> a hiatus bridging sound connecting two vowels. Retroflex stops, an innovation in IA from the Dravidian languages, were often confused with dental stops (Geiger §41.3, §42).<sup>15</sup> Pāli is a mixed language containing dialect forms from different parts of India; for example the OI word *kṣaṇa* (“instant”), which occurs in Pāli in its eastern form (*khaṇa*) and its western form (*chaṇa*), or the word *arya* (“noble”) which occurs as both eastern *ariya* with an epenthetic -i- added and western *ayya* with a geminate; or the western word *taṇhā* (“craving”) which, along with eastern *tasiṇā*, derived from OI *tṛṣṇā*. We could also consider a basic word like *loka*, “world”. Pāli treats the word the same as Sanskrit. In the Asokan edicts it is written as *loga* in the northern Jaugada edict; Ardhamāgadhī transmits it as *loṇa* or *loa*; Gāndhārī has several possibilities including *loga*, *loku* and *loo*. The evidence suggests that the form transmitted to us in Pāli, which one might reasonably expect to be similar to one of the Prakrit forms (Norman 1983: 4–5) was subject to linguistic diffusion and later back-formed to *loka*, that is, it was Sanskritized (Levman 2021a: 276–88).

#### 2.4.2 Linguistic diffusion from autochthonous languages, bilingualism and word borrowing

When the Indo-Aryans migrated into northern India they encountered a settled population of Dravidian and Munda speaking peoples (Emeneau 1954: 282). A strong argument can be advanced that the Buddha’s Sakya clan were historically a Dravidian speaking people in the process of being assimilated into

<sup>14</sup> The y-glide was used as a substitute for a weakened or vanished intervocalic stop in the Prakrits, notated with an overdot -ỵ- in Ardhamāgadhī and just a -y- in other Prakrits. Pischel (§187) calls it a weakly articulated ya (*laghu-prayatnatara ya-kāra* “Lightly articulated -ya-syllable”). Von Hinüber (2001: §171) calls it a “hiatus-eraser” (Hiattilger in German). It is also found in Vedic writings; Bloomfield & Edgerton note the “similar use of the sound [-y-] in Prakrit and Pāli” (1932/1979: §338), but this is more for external sandhi usage, than representing the loss of a consonant. Vedic has examples of stop disappearance, but not, that I am aware of, where they are replaced with a y-glide; see von Hinüber op. cit. §170, Vedic *maireya* “intoxicating drink” <\* *madireya* (Vedic *madirā*); Vedic *prayuga* “forepart of the shafts of a chariot” > *prauga* (Wackernagel 1896: §37b)

<sup>15</sup> Hallisey quotes an illustrative example from the *Sutta Nipāta*, where, in the first verse of the *Uragavagga*, PTS has *visata* (< vi + sṛ, p.p. *visṛta*, “spread”), while all the other recensions (Be, Ce, and Se) have *visaṭa* with retroflex -ṭ-, although Norman (1990: 34; CP 4: 84) writes that Ce has a dental -t- (but not present in the BJT version of the text). This may in fact be a difference in MI dialect pronunciation, as Norman suggests, where the vocalic -ṛ- has changed the dental -t- > -ṭ- in some, but not all cases. The OI rules of retroflexion do not require such a change.

an IA culture, judging by the toponyms of the towns they lived in, the names of many of the monks which are of Dravidian origin (Levman 2021b: 181) and the linguistic and cultural borrowings from Dravidian found in the Pali scriptures (Levman 2021a: Chapters 1–4; Thomas 1931: 23 thought their native language was Munda).<sup>16</sup> This extends far beyond simple word borrowing and includes phonological borrowing, morphological and syntactical influence. We find a significant imprint of Dravidian language features on Pāli from the earliest parts of the language, that is, the early Buddhism of the early *Nikāyas*, the *Sutta Nipāta*, the *Dhammapada* and *Thera/Therīgāthā* (for summary see Levman 2021b: 170–72).<sup>17</sup> The eastern clans were looked down upon by the western Indo-Aryans; their speech was considered *mṛdhra-vācaḥ* (“obstructed”; Deshpande 1979: 254; Levman 2013: 154–55), which is not surprising considering the very different phonetic/phonological system in the indigenous Dravidian and Munda languages. Linguistic diffusion from indigenous south-Asian languages is often not distinguishable from intra-dialect IA language changes (as discussed above §2.4.1). Either one could be the cause of, for example, the weakening or strengthening of an intervocalic stop, or, inter-dialect variations could be catalyzed and accelerated by the constraints of bilingual speakers’ native phonological systems. Included in this category as well, is tradents’

<sup>16</sup> Dravidian Buddhism (which I define as “Buddhism among Dravidian speaking peoples, in Dravidian speaking areas and influenced by Dravidian culture”; Levman 2023a: 59) is a largely unstudied chapter in the history of Buddhism in India. It is a matter of Buddhist history that the *Damīlas* (or at least some of them) were at one time part of the Buddhist Sangha. Asoka brought his *dhmma* message to the south, beyond the frontiers of the *Coḷas* and the *Paṇḍyas* (including present day Tamil Nadu) all the way to Taprobana (i.e. Sri Lanka, Rock Edict 13, Bloch 1950: 130), through the missionary work of his son Mahinda. The *Nāgārjunikoṇḍa* inscription (3<sup>rd</sup>–4<sup>th</sup> century CE) congratulates the Buddhist monks for having converted the *Damīlas* to Buddhism (Lamotte 1958/88: 342). Buddhaghosa himself, who went to Sri Lanka and translated all the lost Pali commentaries from Sinhalese back into Pāli, was a southerner who lived for sometime in Kañcī in Tamil Nadu (von Hinüber 1996/2008: §207; Mp 98<sup>4</sup>), and undoubtedly spoke a form of Old Tamil. *Dhammapāla*, author of the *Paramattha-dīpanī* (commentaries on the *Thera-* and *Therīgāthā*) was a native of Padarattittha, near Kañcī (Lamotte, *ibid.*, p. 350). Buddhadatta, commentator on the *Vinaya* and *Buddhavaṃsa*, was born in Uṛaiyūr and wrote many of his works in the Bhūtamangalagāma monastery in the *Coḷa* country. For a comprehensive treatment of the “Dravidian connection”, see Levman 2023a.

<sup>17</sup> On page 172 of Levman 2021b I say “there is no mention in the commentaries of the Buddha actually teaching in a non-IA language and that teaching being translated into a *koiné* or Pāli.” While there is no explicit statement to that effect, Étienne Lamotte describes an incident in the *Sarvāsativāda Vibhāṣā* where the Buddha speaks in Dravidian, and the commentary to the *Parisāsutta*, strongly suggests that the Buddha was speaking non-IA languages. See below, §2.4.3.

lack of knowledge of the Pāli language which often leads to a confusion of Pāli verbal forms and roots.

### 2.4.3 Linguistic diffusion: bilingualism

It is quite possible that the Buddha himself was bi- or multi-lingual and certainly the early tradents were, teaching in one or more MI dialects (including the aforementioned *koiné*) as well as one or more autochthonous languages. Lamotte (1958/88: 550) remarks that “there is no doubt that the Buddha possessed the gift of tongues and that occasionally, to make himself better understood or to respond to the preferences of his listeners used non-Āryan languages.” He describes a famous incident in the *Vibhāṣā* (3<sup>rd</sup>–4<sup>th</sup> century CE) where the Buddha converts four kings, first by speaking in Sanskrit to the first two, then, when the second two do not understand, speaking in Dravidian and then in Mleccha.<sup>18</sup> Although this is the only incident I am aware of where the Buddha is said to speak Dravidian,

<sup>18</sup> The *Vibhāṣā* is a Chinese translation of a lost OI work. A similar text occurs in the Gilgit Manuscripts (Dutt, 1947: 256–60), from the Mūlasarvāstivādin Vinaya, the relevant section of which reads:

“He addressed the great king Dhṛtarāṣṭra as follows: ‘O Great King, the body is old, feeling has been tranquilized, perception has ceased, mental intentions are calmed, consciousness has disappeared. This is the end of suffering.’ When this Dharma was being taught, the stainless Dhamma eye arose for the great king Dhṛtarāṣṭra and some of his hundred thousand *gandharbhas* in regards to the Dhamma ...” (... *iti hi mahārāja jīrṇaḥ kāyo vedanā śīti-bhūtā saṃjñā nirūddhā saṃskārā vyupaśāntā vijñānam-astam-gatam / eṣa eva-anto duḥkhasyeti* ..., pp. 258<sup>8-9</sup>).

“... Then the Bhagava addressed the great king Virūdhaka as follows: ‘In the seen let there be merely the seen, in what is heard, in what is thought, in what is cognized, let there be merely what is cognized.’ When this Dharma was being taught the stainless Dharma eye arose for the great king Virūdhaka and some of his hundred thousand *kumbaṇḍhas* ...” (... *iti hi mahārāja dṛṣṭe dṛṣṭa-mātram bhavatu śrute cintate vijñāte vijñāta-mātram* ... p. 258<sup>13-15</sup>).

“... Then the Bhagava addressed the great king Virūpākṣa as follows: *ene mene daṣphe daṇḍaṣphe* that is the end of suffering ... When this Dharma was being taught the stainless Dharma eye arose for the great king Virūpākṣa and some of his hundred thousand *kumbaṇḍhas* ...” (... *iti hi mahārāja ene mene daṣphe daṇḍaṣphe eṣa eva-anto duḥkhasyeti* ... p. 259<sup>1-2</sup> with footnote: “Tibetan *e ne me ne dab phye daḍap phe* (ཨ་ནེ་མེ་ནེ་དབ་ཕྱེ་པེ).” “... Then the Bhagava addressed the great king Vaiśravaṇa, ‘Here, for you Great King, *māṣā tuṣā saṃśāmā sarvatra virāṭhi*, that is the end of suffering.’ When this Dharma was being taught the stainless Dharma eye arose for the great king Virūdhaka and some of his hundred thousand *nāgas* ...” (... *atra te mahārāja māṣā tuṣā saṃśāmā sarvatra virāṭhi eṣa eva-anto duḥkhasyeti* ... p. 259<sup>6-7</sup>).

According to Lamotte (1958/88: 551) *ene mene...* represents a Dravidian language and *māṣā tuṣā...* a Mleccha language, but this is not stated in the Sanskrit text. The words *ene mene...* do not appear to be Dravidian, nor is the *māṣā tuṣā* language identifiable.

there are references in the commentary to various monks who spoke non-Aryan languages (including Dravidian speakers; the Kirātas, perhaps of the Tibeto-Burman group; the Savaras, Munda speakers; and the Yavanas or Greek speakers) whose mispronunciation of the Dhamma would invalidate a *kammavācā* (an official act of the Sangha; Levman 2017: 31–33). There is also one statement in the *Parisāsutta*, where, in reference to the eight assemblies, the Buddha is quoted as saying “I appeared just like them, and my voice became like their voice” (Bodhi 2012: 1212); “whatever their language that became mine” (Hare 1935/2006: 205); which suggests that he indeed spoke the Dhamma in many other (non-IA) languages, and the “otherness” of the languages is so stated in the commentary.<sup>19</sup> Although the commentarial references are late, it is clear that the tradition believed the Buddha was bi- or multi-lingual and it should not be discounted out of hand; given all the linguistic data, it is highly unlikely that the Buddha did not speak a non-IA language like Dravidian.

Murray Emeneau identified bilingualism as the major catalyst of change in what he called the “Indian Linguistic Area” (1956; 1974; 1980). The indigenous peoples had to learn the language of the new IA immigrants, a language which was fast becoming politically, economically and culturally dominant. But the Dravidian and Munda languages also left a strong imprint on IA; a

<sup>19</sup> AN 69 (9), 4, 307<sup>21–23</sup>: *tattha yādisako tesam vaṇṇo hoti tādisako mayham vaṇṇo hoti, yādisako tesam saro hoti tādisako mayham saro hoti*. The comm. (Mp 4, 148) says they are “different languages” and describes them as: “Their sounds are broken/interrupted/not continuous/cut off (*chinna-ssarā*), they are incomplete/defective (*khaṇḍa-ssarā*) or roaring (*gaggara-ssarā*), declining/sloping (*pabbharā*), confused, rumbling, stammering, stuttering (*babbhara* < Skt. *balbalā-kṛ* ‘to stammer or stutter,’ *barbara* = Gr. βάρβαρος ‘stuttering; people of an unknown language’), the sound of the crow (*kāka-ssarā*). The Teacher is like the voice of Brahma (*sattā brahma-ssaro va*, the supreme voice that can create all sounds). This is said in respect of different languages.” *te chinna-ssarā pi honti khaṇḍa-ssarā* (vars. *gaggara-ssarā, babbhara, pabbharā*) *pi kāka-ssarā pi, sattā brahma-ssaro va. idaṃ pana bhāsa-antarām sandhāya kathitaṃ*. For a similar description of the “obstructed speech” of the non-Aryans in the *Śatapatha Brāhmaṇa*, see Levman 2013: 255; see also Levman 2017: 32 for a similar description of the *milakkhānaṃ bhāsā* (“language of the barbarians,” the non-Aryans, at Sv 1, 176<sup>26</sup>). Although the above compound *bhāsā-antarām* may also be translated as “different way of speaking” (as well as “different languages”), I interpret it as referring to both the different languages and the different manner in which they are articulated. The *mūla* and comm. passages are also repeated in the *Mahāparinibbānasutta* DN 2, 109<sup>13–15</sup> and Sv 2, 560<sup>22–24</sup>. The Buddha’s supernatural ability to adopt the appearance, vocal mannerisms and (in my interpretation) languages of his audience as recounted in the Mp and Sv commentaries suggests a relatively late date for their composition (like the *Vibhāṣā*, probably early medieval); nevertheless they do shed some light on the tradition’s belief in the Teacher’s putative bi- or multi-lingualism.

good summary of some of the features shared by Indo-Aryan and Dravidian languages (but not by Iranian, Indo-Aryan's closest Indo-European relative) is found in Masica 1976: 178–86, Sjoberg 1992: 55–69, and Krishnamurti 2003: 38–42. These include, *inter alia*, the use of retroflex consonants, the extensive use of non-finite verbs in strings as a compositional principle, the use of the quotative marker in reporting direct speech, syntactic parallels between the proto-Dravidian *-um* suffix and IA *api*, and the use of “echo words” or “expressives.”

Some of the phonological features of the Dravidian and Munda languages led to confusion in the Dharma transmission, if one of the monks (the transmitter or receiver) was a bilingual, MI-as-a-second-language-speaker. Dravidian lacked voiced stops and aspirated stops, to name two major impediments to accurate Dharma transmission; many variant readings which show both voiced and unvoiced stops, aspirated and unaspirated stops are possibly attributable to this cause. Judging from the bilingual or non-IA names of many of the monks in the earliest transmissions (mentioned above §2.4.2), Dravidian speaking monks were present right from the beginning of the Sangha formation, which again points to the antiquity of many of the phonological anomalies to be discussed below.

#### **2.4.4 Linguistic diffusion: foreign words, word borrowing**

The IA immigrants naturally adopted many of the foreign words of the indigenous inhabitants for biota they were unfamiliar with. They also adopted foreign technical terms for cultural practices like the *kaṭhina* practice (Bechert 1968: 324; Hu-von Hinüber 1994: 4–5) and the *kuṭi* meditation hut which they borrowed from the locals (Levman 2021a: 154–56). Since these words were native to a foreign phonetic system, when imported into MI various attempts were made to render them accurately in the IA system and this naturally resulted in variant interpretations and spellings. There are hundreds, perhaps thousands of foreign words in the canon (no complete inventory has been attempted). A good example of the phonetic problems that result from the importation of a non-IA word is the word *jaḷogi* (an alcoholic drink), which occurs in many different variant forms, many or all of which must have been present from the earliest strata of the canon, as the word featured prominently in the second council debates, one hundred years after the Buddha's passing (Levman 2019a: 91–92).

## 2.5 Oral transmission confusion, ambiguity

For over three hundred years the teachings were transmitted orally from *bhāṇaka* to *bhāṇaka* without the aid of a written guide. It is inevitable that mistakes would happen, especially when facilitated by some of the pathways outlined above, the propensity of dialect speakers to communicate their own dialect idiosyncrasies, the natural evolution of language towards simplification over time, and the constraints of MI speakers who learned the language late in life and for whom certain sounds were not natural. Along with oral transmission confusion should be grouped variation due to lack of knowledge of the Pāli language (as demonstrated in Levman 2021b) and the imperfect memory of the *bhāṇaka*, including for example his addition of material that never existed due to the “vicissitudes of memory” (Anālayo 2009); these largely unintentional changes were caused by a “lack of systematic training of the Buddhist reciters” comparable to Vedic reciters, resulting in memory slips not being corrected and being incorporated into the transmission (Anālayo 2022: 36, 202). Exacerbating this phenomenon was the *bhāṇaka*’s imperfect pronunciation, or the imperfect auditory skills of the scribe, when MI was not their native language. To this last factor may be added the influence of the copyists’ language on manuscript transmission, e.g. the Thai language, which lacks certain elements phonemic in Pāli—voiced aspirates, all retroflexes, palatal and velar voiced stops, and palatal nasals—and which may be a factor in some Thai variants.

While it is impossible to quantify this factor, it no doubt had a role to play in the transmission of variants; the *suddhā/suttā* confusion discussed above (§1.4) may be due to this factor, exacerbated perhaps by phonological confusion caused by the linguistic diffusion issues (§2.4). In effect this category is a “pot pourri” which subsumes all the seven processes discussed in §§2.1–2.4.4 above, where it is impossible to determine which one(s) of these processes are operant for a given variant, or where none of these processes are operant but there is nevertheless oral confusion that is purely sonic in nature (like incorrect word division in the example *karontā-gacchanti* and *karontā-āgacchanti* from the *Agāṇṇasutta*; Levman 2025a: v3: 94, n. 204 or syncope (syllable loss) and confused word division in the *Cakkavattisutta*: *tapasa-brahmacārī*, vs. *tapa-brahmacārī* vs. *tava sabrahmacārī*, *ibid.*, v3: 30, n. 52); oral confusion and ambiguity are a common denominator of them all.<sup>20</sup>

---

<sup>20</sup> The interested reader should look these examples up in the Pāli index, as the pagination is likely to change on final printing.

## 2.6 Commentarial glosses

As Endo and Norman have opined, many parts of the commentary may go right back to the time of the Buddha, and possibly even the Buddha himself, when monks would ask the teacher to explain certain points (Norman 1997/2006: 195; Endo 2013: 5). In his mission to Sri Lanka in the mid-3<sup>rd</sup> century BCE, Mahinda brought a complete commentary to the island and translated it into Sinhalese, according to Buddhaghosa's introductory verses to the *Dīgha Nikāya Commentary*.<sup>21</sup> The commentary often has variants which do not exist in the main text, some of which are the commentators' attempt to make sense of a problematic passage, others which may be synonyms for certain words, spoken by the Buddha and remembered by the monks. Sometimes the commentary is misplaced in the *mūla* text (Norman 1997/2006: 200, 206, 210; Anālayo 2022: 147–51). Often the commentary itself talks of alternate readings (*ti pi pāṭho*, discussed above §1.4).

## 2.7 Written transmission errors

Once the text was written down, orthographic variation (differences in transcription protocol) and copying errors, common in the transmission of manuscripts could, and did take place, that is, errors such as incorrect spelling, wrong compound and sentence division, haplography, misreading, dittography, etc. (Tov 1992/2001: 199–286 for the Hebrew Bible). There were also particular problems with Indic scripts like Brāhmī where geminates were not written down and the marks for long vowels were frequently omitted (Norman 1997/2006: 107), and Karoṣṭhī where neither geminates nor long vowels were shown (von Hinüber 2015). Several of the south-Asian scripts (Burmese, Thai, Sinhalese) had letters very similar to each other which could be easily misunderstood. Nevertheless, most of the phonological (sound transmission) confusions would not have occurred if a firm base text was present as a guide. I attribute most of the variants discussed below to the oral tradition, but I also identify instances where orthographic variation may have played a role.

<sup>21</sup> *Sumaṅgalavilāsinī* 114–17: verse 6: *attha-ppakāsana-atthaṃ, atṭha-kathā ādito vasisatehi. pañcahi yā saṅgītā, anusaṅgītā ca pacchā pi.*

“For the sake of explaining the meaning, the commentary, etc., was sung and afterwards repeated by the five hundred Masters [i.e. arahats],

verse 7: *sihaḷa-dīpaṃ pana ābhatātha, vasiṇā mahāmahindena. thapitā sihaḷa-bhāsāya, dīpa-vāsinam atthāya.*

“and brought to the island of Sihaḷa, by the Great Master Mahinda and translated there into the Sihaḷa language, for the sake of the residents of the island.”

## 2.8 Harmonization and standardization

Many problematic areas of the canon were harmonized and standardized by the medieval Pāli grammarians. For example, the absolutive in *-ttā* (OI *-tvā*) was restored throughout the canon to its OI form, although some vestiges still remain (von Hinüber 1983: 7). Problematic passages, especially those involving foreign words were often preserved since no one knew what they meant. Helmut Smith hypothesized that the Pāli that has come down to us is really the product of the medieval grammarians (“la conviction que notre pali est une fonction de celui du 12<sup>me</sup> siècle,” referring to Aggavaṃsa’s *Saddanīti*; 1928–54/2001: vi), who standardized the language. Although Smith’s view is certainly an exaggeration, there is nevertheless some truth to the fact that Burmese grammarians exerted a not insignificant influence on the canonical tradition, especially in Burma. These changes, however, are limited and do not affect Norman’s observation (1983: 5–6) that only minor changes to the canon took place after it was written down in the first century BCE. In any case, this discussion is concerned mainly with changes that can be attributed to the oral transmission tradition.

There is one other possible source of variation, only touched on in this study (§4.7 below), as it is rarely found in the *Theragāthā* verses explored here: the possibility of two separate transmissions, not phonologically/phonetically related, where the Buddha may have used different words and phrases in similar contexts or they were (mis-)remembered that way by different tradents (n. 5 above).

### 3.1 The *Theragāthā*

In the following I chose nineteen examples from the *Theragāthā*, illustrating the processes listed above in section §2.1–2.8. It is one of the oldest of Buddhist works which Norman dates between the end of the sixth century BCE to the middle of the third century BCE (1969/95: xxix). These are utterances of the monks from the early Sangha, most of whom are said to have been contemporaneous with the Buddha. In many cases these verses antedate their rendering into Pali,<sup>22</sup> which began to take place from sometime around the 3<sup>rd</sup>

---

<sup>22</sup> According to Norman (1990: 34), all Buddhist transmissions we possess are translations of earlier compositions. This applies as well to the *Theragāthā* which were probably originally transmitted in the same dialect/*koiné*/language that the Buddha and his early tradents used. Evidence of this earlier form will be adduced in the examples below.



century BCE, perhaps around the same time as when Mahinda and associates brought the then inchoate canon and the commentary to Sri Lanka (Warder 1967: §13; Norman *ibid.*, p. xxix). The examples could have been taken from any of the early works, such as the *Sutta Nipāta*, *Dhammapada*, early *Nikāyas*, etc., which all exhibit the same variant phenomena. I identify the major categories of variation before each example, grouped according to the above list, and highlight some of the important elements, for ease of organization. The category of harmonization and standardization has been omitted as this is difficult to date, and in any case, I am primarily interested in changes during the early oral transmission stage.

### 3.2 Variant summary

As the reader has no doubt noticed, there are a lot of interconnections and overlapping in the above categories. The majority may be grouped under the general category of “oral transmission confusion and ambiguity” (§2.5) with the other processes (seven categories, §2.1–2.4.4) being in effect the causes of this ambiguity. Diachronic variation (lenition and loss of intervocalic stops; loss of aspirated stops and their replacement by an aspirate only; assimilation and/or resolution of consonant clusters; leveling of sibilants and glides, etc.) is a natural process, but is catalyzed and accelerated by a putative *koiné*, which, for ease of inter-communication, removes distinctive dialect differences which might obstruct understanding. A *koiné* might be conceptualized as a logical development of dialect diffusion, where the unpredictable effect of one dialect on another (e.g. a dialect that regularly voices intervocalic consonants vs. one that devoices them) is neutralized by removing the consonant altogether and replacing it with a -y- glide. This in itself led to further variation as tradents tried to interpret what the glide meant, when “restoring” the word to its earlier form. At the same time many of the IA phonemic differences, which do not exist in a native language like Dravidian, are not even perceived by bilingual, non-IA first language speakers, again acting as an accelerant to change and simplification, like, for example, eliminating aspirated stops (not phonemic in Dravidian) and replacing them with an aspirate only. Word borrowing is a particularly good example of the change that results from the other vector, that is, IA speakers trying to interpret how a native, borrowed word or phrase should be spelled and understood in Pāli with a very different phonemic system.

All of these changes were actualized and amplified by the oral/aural<sup>23</sup> vagaries of live communication, recitation, dictation and audition, and alongside these processes was the inchoate impulse to Sanskritize which began as early as the third century BCE (Levman 2023b: 5–8) with the aim of clarifying, defining and fixing the meaning. The transmission process was very complex. Its sole goal was the preservation of the *Buddhadhamma* in its “original” form; but nothing could escape the basic rule of the universe: *anicca*, change (including even the fabled Veda memorization process, which itself contains thousands of variants, §1.4 above), and so, through the mixture of these various factors, thousands of variants resulted, sometimes with quite different meanings equally contextually plausible, but adding a slightly different or very different nuance to the passage, or variants not—or imperfectly—understood and (presumably) preserved for that reason.

#### 4. *Theragāthā* examples

##### 4.1. *Theragāthā* 118, *Kimila-ttheragāthā* (vars. *Kimbilo*, *Kimilo*, *Kimmito*)

##### **Linguistic diffusion from autochthonous languages and the influence of bilingualism**

Geminate/aspirated geminate ambiguity (-tt-/-tth-)

retroflex/dental ambiguity (-ṭṭh-/-tth-)

##### **Diachronic change**

(-mb-/-mm-)

##### **Oral transmission ambiguity**

(*abhisatto*/*abhisitto*)

*Abhisatto* (vars. *abhisatto*, *abhisatṭho*, *abhisitto*; comm.: *abhisāpakato*, *abhisankappa*, *abhisāmakato*, *abhiḷāsapito*, *abhiḷapaṇkato*) *va nipatati vayo, rūpaṃ aññaṃ iva c’ eva santaṃ; tass’ eva sato avippavasato*

<sup>23</sup> By oral/aural I mean transmission of the *Buddhadhamma* from the Buddha to his disciples, from his disciples to other tradents including *bhāṇakas*, from *bhāṇakas* to other monks and other *bhāṇakas* and from *bhāṇakas* to scribes. Along each step of the transmission process there is an oral communication and aural listening process involved whereby variants can be introduced.

*aññass'eva sarāmi attānan ti.*<sup>24</sup>

“(Old) age falls upon one as though ordered (*abhisattho*); the shape, although the same, is as though different. I remember my own self as though of another, although I am the same, not having been away” (Norman, 15).

or

“As if cursed (*abhisatto*), old age falls ...” or “As if anointed (*abhisitto*), old age falls ...”

The PTS has the reading *abhisattho*, while the Burmese, Sinhalese and Thai have *abhisatto*. There are also other variant readings as noted.

A major phonological difference between Dravidian and IA is the former language's lack of phonemic voiced stops and aspirated stops, a very prominent feature in Pāli. So when a word like *abhisattho* is heard by a Dravidian tradent (someone speaking IA as a second language), he or she would probably hear it without the aspirate, i.e. *abhisatto*. This is likely why both readings are recorded in this poem, which, however, mean different things. *Abhisattho* is the past participle of *abhisamsati* (< OI *abhiśams*, “to blame, accuse” which per the commentary has the meaning here of *abhiśās*, “order, direct, assign, allot”); both meanings would work in the context; *abhisatto* is the past participle of OI *abhiśap*, “to curse” so the intended meaning in the *gāthā* is unclear as both make sense.

The commentary glosses *abhisattho* by *anusittho*, *āṇatto* and *abhisattho*, all meaning “ordered” (Be *abhisatto*, instead of PTS *abhisattho*); *abhisattho* is an alternative way of spelling *abhisattho*; here the variation is due to retroflex/dental ambiguity. As is well known, retroflex letters were borrowed from Dravidian into IA and Dravidian speakers had a threefold discrimination of dental, alveolar (not in OI/MI) and retroflex stops, more subtle than native IA speakers; any slight allophonic variation in pronunciation might easily result in this kind of refinement of a dental to a retroflex consonant.

The poem is uttered by *Kimila* in response to the Buddha's creation of the

<sup>24</sup> Burmese: *abhisattho* (Be); *abhisatto*, *abhisattho*; *abhisāpakato*, *abhisāmakato* (comm.). Sinhalese: *abhisitto*, *abhisattho* (Ce); *abhisankappa*, *abhiśāpato*, *abhiśāpato* (comm.). Thai: *abhisattho* (Se); *abhisatto*. CPD (s.v. *abhisattha*): *abhisatthalāpato*, *abhisatthasapito*.

Burmese: *Kimilo* (Be); *Kimilo*. Sinhalese: *Kimilo*, *Kimilo*, *Kimilo* (Ce). Thai: *Kimilo* (Se). Pāli quotes are from the PTS edition, unless otherwise noted.

form of a woman in front of him, standing in the first wave of youth; gradually he showed her overpowered by old age, disease and distress.<sup>25</sup> He recalls how his own youth has so quickly passed away and realizes the truth of *anicca* and with this as a springboard, soon attains to arhathood.

In a previous life *Kimila* venerated the relics of the Buddha *Kakusandha* with garlands of *saḷala* flowers and by building a *maṇḍapa* (“pavilion”).

*Kimila* is itself spelled in three different ways, with a geminate *-mm-*, *Kimmila*, and with the conjunct *-mb-*, *Kimbila*. See below discussion ad Th 95 re: *pakkhando/pakkhanno* for diachronic change of *-nt-/nd-* > *nn-* (nasal + stop > geminate nasal). He was a Sakyan convert so his name was probably not IA, but Dravidian or Munda, perhaps related to OI *kṛmi* (“worm”) cp DED #1614 *kīri* “intestinal worm” and the Munda language, Korku *kīra* “worm, insect, germ”; in Th-a 1, 244 the Buddha conjures up the picture of a beautiful woman disintegrating into old age and disease which spurs *Kimbila* on to reach arhathood. These forms of *asubha* meditation invariably show worms crawling out of the corpse’s nine holes.<sup>26</sup> The change of *-mb-* > *-mm-* may be IA dialectal in origin or more likely appears to be a normal diachronic conjunct weakening and assimilation over time (cp *Gāndhārī*, Brough 1962: §46 where *mbh-* appears as *-mm-* and *-nd-* as *-n-* or *-nn-*). This also occurs in Munda (Kuiper 1948b: 383, *-nd-* > *-n-*, *-mb-* > *-m-*), and historically in Telugu, a Dravidian language (Emeneau 1970: 109, *\*-mp-* > *-mm-*).

The two meanings—cursed (*abhisatto*) and ordered (*abhisattha*)—are quite different, although they do not change the overall message. The *mūla* text has yet another reading *abhisitto* (“sprinkled” < *abhisīñcati*), which is less apt; this appears to be the case of an oral transmission confusion (*abhisatta* > *abhisitta*).

The commentary then gives five other variants or glosses (*abhisāpakato*, *abhisāṅkappa*, *abhisāmakato*, *abhiḷāsapito* and *abhiḷapaṁkato*) a sure sign of an oral transmission issue, and the CPD lists two others without giving the source (*abhisatthalāpito* and *abhisatthasapito*).

What we apparently have here is a Dravidian speaking monk(s) wrestling with the word *abhisatta/abhisattha*, trying to decide what it meant and which

<sup>25</sup> Th-a 1, 244<sup>28-30</sup>: *Satthā paṭhama-yobbane ṭhitaṁ dassaniyaṁ itthi-rūpaṁ abhinimminivā, purato dassetvā, puna anukkamena yathā jarā-roga-vipattihi abhibhūtā dissati tathā akāsi. Taṁ disvā Kimila-kumāro ativiya saṁvegaṁ pakāsento ...*

<sup>26</sup> As in the Buddha’s graphic presentation of a corpse to Kulla the monk (Th-a 167<sup>23-25</sup> ad Th 393-98): *Atha naṁ Satthā tassa pekkhantass’ eva navahi vaṇa-mukhehi paggharamānā suciṁ kimi-kulākulaṁ ativiya bibhacchaṁ duggandha-jeguccha-paṭikkūlaṁ katvā dassesi.*

root it came from (*abhisattha* < *abhiśaṃs* or *abhiśās* or *abhisatta* < *abhiśapta*) with several other suggestions arising both phonologically very close (*abhisitta*, “sprinkled”; *anusitṭho*, “ordered”) with commentarial glosses (the variants above starting with *abhisāpakato*).<sup>27</sup>

#### 4.2 *Theragāthā* 3, *Kaṅkhārevata-ttheragāthā*

##### Diachronic language change, influenced by bilingual speakers

*Paññaṃ imaṃ passa tathāgatānaṃ: aggi yathā pajjalito nisīthe* (vars.  
*nisīve, nisive, nissive*).<sup>28</sup>  
*āloka-dā cakkhu-dadā bhavanti ye āgatānaṃ vinayanti kaṅkhan ti.*

“See this wisdom of the *Tathāgatas*, who, giving light and vision like a fire blazing in the night, dispel the doubt of those who come” (Norman, 2).

The PTS, Thai and Burmese all have *nisīthe* “at midnight, at night” which the commentary glosses *rattiyam*, idem. The OI word is *niśītha* < *ni* + *śī*, to lie down, but the verb is not attested with this prefix, as *śī* already means “to lie down” so the prefix seems superfluous; a better etymology should be sought elsewhere. The word occurs in several forms in OI, all meaning “night”: *niśītā* (*Taittirīya Saṃhitā*),<sup>29</sup> *niśā* (*Sūtras*), *niś* (*Manu*), as well as *niśītha* (*Mahābhārata* = MBh), per Mayrhofer (1956–76: vol. 2: 168). The Pāli form occurs also as *nisīthe* with three variations *nisīve, nisive, nissive*. There is also a later form *nisā* which occurs in Mil and later Pāli works. In AMg the form is *ṇisā, ṇisi-* (in compounds, e.g. *ṇisi-bhatta*, “night-meal”) and *ṇisiha*, “midnight”; there is also a Prakrit form *ṇisā, ṇisi-* (compounds), *ṇisiha* and *ṇisīḍha*.

<sup>27</sup> *Abhisāpakāto*, “curse made” = “as if cursed”; *abhisankappa* ~ OI *saṃkṛp*, Pāli *sankappa*, “thought, intention” = “as if intended”; *abhisāmakāto*, “to make like (*sāma*)”?; *abhiśāpito* “desired” (caus.) < *abhiśa* “to desire, wish for, covet, crave” = “as if wished for”; *abhiśāpamkato*, “to make talk”? (< *abhiśap*, “to talk or speak about”).

<sup>28</sup> Burmese: *nisīve, nisive, nisīthe* (Be). Sinhalese: *nisive, nisīthe* (Ce); *nisīve, nisice* (comm.). Thai: *nisīthe, nisive* (Se); *nissive* (comm.).

<sup>29</sup> The brackets show the first occurrence of the word, which is included for dating purposes, when appropriate.

Although the etymology is not clear,<sup>30</sup> assuming the earliest form is *nisitā* (< OI *niśitā*), all the variants seem to be normal diachronic changes, probably influenced by bilingual speakers, viz.:

*niśitā* > *niśītha* > *nisītha* > *nisīdha* > *ṇisīha* (AMg), *nisīva* (Pāli)

The interpretation of the dental *-t-* as an aspirate stop *-th-* is probably due to bilingual speakers who had no phonemic aspirate stop; the dental *t<sup>h</sup>-* has a natural allophonic aspirate and was probably interpreted in terms of the IA phonological system. As an intervocalic consonant, Dravidian speakers would automatically strengthen it and pronounce it as *-d<sup>h</sup>-*. The reason for the vowel change of *-itā* > *-ītha* is not clear, but probably due to a change of accent from the ultimate to the penultimate syllable. The AMg/Prakrit *ṇisīha* is the normal change of an aspirated stop into an aspirate (Pischel §188). The change of *-dh-* > *-v-* is, however, a little unusual. Norman (1992/2006: 157) cites several examples of a *v/dh* alteration (*dhīro/vīro* Sn 44, 165, 439, 531, 646; *vaṃkaṃ/dhaṃkaṃ* Sn 270–71; *avibhū/adhibhū* Sn 684; *vimutta/dhimutta*, Sn 1071, 1114; Dhp 193) which he says (1997/2004: 117) arise from “the similarity of the two characters in early Brahmi script”;<sup>31</sup> however, there are too many of these for this to be considered a “scribal error,” and in any case the characters are quite different. The *-dh-* > *-v-* change is more likely a not uncommon dialect change (weakening), as there are a number of examples in the Vedic writings of *d* and *v* alternation (Bloomfield and Edgerton 1932/79: §869; e.g. *diśa/viśa*), and as we have seen, the addition of an aspirate to a stop was a common phenomenon amongst bilingual speakers (and also Prakrit speakers, Geiger §40). The *dh/v* alternation also occurs in BHS (e.g. *avodigbhāga/adhodigbhāga*, Study Group 2006: Chapter 9, 56a1). Alternatively the *-v-* is simply a hiatus glide (usually pronounced as a *-w-* in Indic languages); similar in function to *-y-* (von Hinüber 2001: §171), replacing the dental stop,

<sup>30</sup> Pokorny (1959/1969: 762–63) derives the English word “night” from root *nok<sup>w</sup>t-* f., *nok<sup>w</sup>ti*, *nok<sup>w</sup>tu-f.*, *nok<sup>w</sup>t(e)r* from which OI *nak*, *naktam* are also derived; but this of course is a long way from *nisitā/nisītha/niś/niśā*. M1 (op. cit.) suggests *niśitā* and the other words were substitutes for the “Scheu gemiedene alte ‘Nacht’ wort” (“shyly avoided old ‘night’ word”) through the newly imagined derivation from *śi* (“to lie”). Dravidian has the words *narkam* (“night”) and *nallam* (“darkness” < PD root *nāl/na!* “night”; Kirshnamurti 2003: 528) and the compound *nisa narkam* (Gadba) which means, “night stands (still)” pointing to another possible derivation from the proto-Dravidian verb *nil-/nit-/nitt-* “to stand” (Krishnamurti, *ibid.*). The etymology requires further research.

<sup>31</sup> The same phenomenon happens in Gandhārī, e.g. v. 173 where Gandhārī has *viru* and the Pāli Dhp has *dhīro*.

which is also sometimes the case in Gāndhārī with the sound *-h-* (Brough 1962: §39); so *niśitā* > *nisitā* > *nisīva*, *niśiha*. If this is the case there would then be two phonetic pathways involved: this one and *niśitā* > *niśītha* (as above). The variant with geminate *-ss-* appears to be orthographic.

**4.3 Theragāthā 7, Bhalliya-tthera**, (vars. *Bhallika*, *Bhalluka*, also *Bhaddiya* (group of five monks))<sup>32</sup>

**Variation due to word-borrowing from non-IA languages**

**Dialect diffusion due to bilingualism and different IA dialects**

aspirate ambiguity

retroflex/dental ambiguity

phonetic equivalence of *-ll-* and *-dd-*

**Variation due to Sanskritization**

*Yo pānudi Maccu-rājassa senaṃ, naḷa-setuṃ va* (vars. *naḷaṃ*, *daḷaṃ*, *daḷha-aṭṭhaṃ*, *nala-*) *sudubbalaṃ mah'-ogho*<sup>33</sup>

*vijitāvī apeta-bheravo hi, danto so parinibbuto ṭhita-atto ti.*

“He who has thrust away the army of King Death, as a great flood pushes down a very weak bridge of reeds, is victorious, with fears truly gone, tamed, quenched, with steadfast self” (Norman, 1).

Bhalliya-tthera's name in the *Apadāna* was Vallikāraphaladāyaka; he gave the fruit of a creeper plant (*valli-kāra-phalaṃ*) to the Buddha Sikkhi in a previous life.<sup>34</sup> In the time of Buddha Gotama, Bhallika and his brother Tapussa were the first people the Buddha met after his *bodhi*; they were going along the high road from Ukkalā to the district where the Buddha was staying.

The change of *naḷa* to *daḷa* (*n-* > *d-*) is not common in the Prakrits, but does occur, e.g. *namayanti*, “they lead” vs. *damayanti* “they tame” in variant versions of *Theragāthā* 19/Dhp 80; *Se diyyamāna* (“being offered”) vs. *ka. niyyamāna*

<sup>32</sup> Burmese: *Bhalliya* (Be, PTS). AN 1, 26<sup>1</sup>: *Bhallika*. PTS Jā 1, 80<sup>16</sup>: *Bhalluka*. Thai: *Bhalliya* (Se); *Bhallika*. Sinhalese: *Bhalliya* (Ce).

<sup>33</sup> Burmese: *naḷaṃ*, *naḷa* (Be); *daḷaṃ*, *nala*. Sinhalese: *naḷa* (Ce); *daḷhaṭṭhaṃ*. Thai: *naḷa* (Se).

<sup>34</sup> Walters (2018: 777) suggests that this might be the *vallikā* plant, a climbing edible plant with red berries, *Vitis quadrangularis*.

(“being taken”) at Sv 2, 355<sup>6</sup>. The word *daḷa* has the meaning of “petal, leaf” so would fit the context (“bridge of leaves”), while the aspirated form *daḷha* (“firm, strong, solid, steady”), does not suit, nor does *-aṭṭham* (“goal” or “meaning”). The word probably intended here is *attham* (< OI *astam* “home”), “like a flood destroying a very strong house”; this is variation due to retroflex/dental ambiguity because of bilingualism, as is most likely the confusion of *-l-* and its aspirated form (*lh-*), which a Dravidian tradent could not hear as it is phonemically absent in his/her language. As for the variation of *-l-* and *-ḷ-* (*nala* vs. *naḷa*) this is probably orthographic (Pischel §226, §240) as they are often interchanged in Pāli, without any change of meaning, as they are here (e.g. see *ala/aḷa* “crab claw”; *unnala/unnaḷa* “vain, insolent”; *galati/gaḷati* “to drip, flow”; etc.). Dravidian has a finer resolution of the liquids with three “l” sounds, normal *l*, retroflex *ḷ* and a voiced retroflex fricative or retroflex approximant (*ḷ̥*) variously represented as *l*, *r*, *ṛ*, *zh*, and *z*, and transliterated as *ṛ*, *ḍ*, *ḷ*, *y*, *r* and *l* in Dravidian and IA languages (n. 40); Caldwell 1875: 59; Emeneau 1970: 98–99).

The variants are perhaps influenced by the similar sounding Dravidian words *tal* “palmyra tree, toddy palm” and *aṭṭam* “terraced roof” or “scaffold” (DED #93); so the compound *tal-aṭṭam* (echoing *daḷha-aṭṭham*) would mean “a scaffold made of palm trees” (“... like a great flood destroying a scaffold of palm trees”), but the “bridge of reeds” *naḷa-setuṃ*) metaphor seems more apt.

A more important clue to the indigenous influence is in the name of the monk Bhalliya, with two variants *Bhallika* and *Bhalluka*. The word *bhallūka* means “bear” in OI (attested in the MBh, but in Pāli, it occurs only in Sadd, outside of this variant); *Bhalliya* and *Bhalluka* are just variants, partially Sanskritized with presumably the same meaning “bear.” The word may also be cognate with *Bhaddiya* (rather than *Bhaddiya* < OI *bhadra*, “auspicious”), one of the group of five monks who were the Buddha’s first converts, as *da* and *la* are often interchangeable in the Prakrits (Pischel §244),<sup>35</sup> as they are in some of the Dravidian languages (Caldwell 1875: 52). M1 vol. 2: 485 says the word

<sup>35</sup> This would not normally apply to a geminate, but in the oral tradition they might be heard as single consonants if not pronounced distinctly or pronounced quickly, in continuous flow (*eka-baddha*), that is, with degemination (Levman 2021a: 298). Consider these word sets which demonstrate the validity of this hypotheses: *adda/alla* “wet, moist, slippery”; Pāli *khudda*, *culla*, OI *khulla* “small” OI *kṣulla* idem; Pāli *chidda*, AMg *chilla* “hole”; etc., and so *Bhadda-/Bhalla-*.



*bhallūka* is “nicht sicher gedeutet” (not determined with certainty); Przyluski takes the word as derived from AA (1929: 196), relating *bhalluka* to *mallu* (both meaning “bear”), both because of the graphic similarity between *bha* and *ma* in most Indian alphabets and because in AA various dialectal forms coexist with a *b-* and *m-* initial consonant (see also Levman 2011: 52–53 and refs. therein; also in 2021a: 222–24). Przyluski’s AA connection is supported by the Munda Etymological Dictionary (Stampe): cp Bodo-Gadaba, Bondo *balu*, “bear”; Mundari *baluk*; Korku, Birhor, Ho, Kol, Mahali, Santali, *bana*; Juang *banae*, *banai*, *bhalu*; Sora *bud-ən*; Birhor *buria*, *bānā*, *bir-miṇḍi*, *bir-būrhiā*; Mundari *buṛi*; Kharia *bə'nəi*; Koda, Santali, Mundari *b<sup>h</sup>aluk*; all meaning “bear.” The words *b<sup>h</sup>aluk* and *bhalu* are almost identical to OI *bhallūka*, with the *-l-* gemination and the addition of an IA ending *a* or *-ka*. Such a wide distribution in most of the Munda language families suggests an old age for the word, well before its first appearance in the MBh; Southworth opines mid-second millennium BCE for proto-Munda words (2005: 195). Since the word is foreign to MI and has various spellings in the different Munda languages, its variant orthography in Pali is to be expected.

As a final note, the same *thera* in the *Apadāna* is called Vallikāraphaladāyaka, because he gave the fruit of a *vallikāra* plant to the Buddha Sumana (per the *Apadāna*, Sikkhi per Th-a). The *vallikāra* plant is derived from the Dravidian root *valli*, “creeper”; see DED #5317 and cp Tamil, Malayalam *valli*; Telugu *valli*, *vallika*, and others, all with the meaning “creeper, climbing plant” (Caldwell 1875: 466; Burrow 1946: 15; M1 vol. 3: 167). Bhallika’s counterpart at the time of the Buddha Sikkhi was the son of a caravan leader named Ujita, who gave that Buddha (with his friend Ojita) his first meal. *Ujita* is a Dravidian name (< *ucitam*, pronounced *ujitam*) with the meaning “propriety, suitability, fitness; excellence, good quality, transcendence.” *Vallika* = *Bhallika* = “creeper” is another possible derivation of this person’s name, with Sanskritization of *v-* > *bh-*, the change *bh-* > *v-* being a common OI > MI diachronic change (see below, §4.5).

#### 4.4 *Theragāthā* 19, *Kula-tthera* (vars. *Kuḷa*, *Kula*, *Kūla*, *Kuḍḍa*, *Kuḍḍha*, *Kuḍḍala*, *Kuṇḍa*, *Kuṇḍala*)<sup>36</sup>

##### Variation due to word-borrowing from non-IA languages and their interpretation by bilingual speakers

ambiguity of the letter *-l-*, *-ḷ-*

alternation of NC and CC (N = nasal; C = consonant; *-ṇḍ-* and *-ḍḍ-*)

alternation of *-l-* and *-ḷ-* with *-t/d-* and *-ṭ/ḍ-*

##### Different interpretation of an underlying *koiné*

##### Sanskritization

addition of IA suffixes

*Udakaṃ hi nayanti nettikā, usu-kārā namayanti* (var. *damayanti*)  
*tejanaṃ.*  
*dāruṃ namayanti* (var. *damayanti*) *tacchakā, attānaṃ damayanti*  
*subbatā' ti.*

“Truly canal-makers lead water, arrow-makers bend the bow,  
 carpenters bend wood, men of good vows tame the self” (Norman, 3).

For discussion on *kula* etymology, see Levman 2021a: 160–63.

*Kula* is a non-IA name, derived from Dravidian, which accounts for its different spellings, as *kula* (Burmese, Sinhalese), *kuḷa* (Burmese, Thai, Cambodian) and *kūla* (Sinhalese); in addition there are five other variants *kuṇḍa*, *kuṇḍala* and *kuḍḍa*, *kuḍḍha*, *kuḍḍala* (Sinhalese comm.). In Pāli the word *kula* means “clan, family, household” and goes back at least to the earliest recorded Buddhist transmissions in the *Sutta Nipata* and the *Nikayas*; in OI the first occurrence is in the late RV (book 10; *kula-pā*) with the meaning “chief of a family or race or tribe”; there is an earlier occurrence in RV book 1 which Mayrhofer translates as “hollow” or “cavity” referring to a goblet (M1 238). This probably comes from a different root or is a mis-translation (Levman 2021a: 160, n. 256).

<sup>36</sup> Burmese: *Kula* (Be); *Kuḷa*. Sinhalese: *Kūla*, *Kula*; *Kuḍḍala*, *Kuṇḍa*, *Kuṇḍala* (Ce); *Kuḍḍa*, *Kuḍḍha* (comm.). Thai: *Kuḷa* (Se, also PTS); *Kula*, *Kuṇḍala* (comm.). Burmese: *namayanti*, *namayanti* (ka. var. *damayanti*), *damayanti* (Be). PTS *namayanti*, *namayanti* (var. *damayanti*), *damayanti*; Sinhalese, Thai: *namayanti*, *namayanti*, *damayanti* (Se, Ce). Dhṛp v. 145 shows one var. ms (C) where all three verbs have *damayanti*.

*Kula* in the meaning of “family, herd, multitude” is derived from Dravidian *kuḷu* (*kuṛu*) “assembly, flock, herd, heap” (DED #1821; Burrow 1943: 139; 1946: 23; Kuiper 1948a: 55 agrees) as well as *kulai/kula* “bunch” (DED #1810; Levman 2021a: 160–63).<sup>37</sup> These are both very old Dravidian roots. *Kuḷu* exists in PSD (Proto South Dravidian; Southworth: 50) and PCD (Proto Central Dravidian) and may be dated to the earlier part of the second millennium. *Kulai/kula* exists in PSD1 and PSD2 which may be dated to c. 1750–1500 BCE. Both of these words go back to a PD root *\*kul* (“to accumulate”). A third root traced back to PD *\*kul* is *kūḷi* “pit, hollow, hole, pond, well” (DED #1818), which exists in all Dravidian language groups (PSD, PCD, and PND, Proto North Dravidian) and may be dated to c. 2500–2000 BCE.<sup>38</sup> This latter word is probably the source of OI *kūla* which has the meaning “declivity, slope, bank, pond, pool, heap” and Pāli “river-bank, embankment, edge of a well.” Krishnamurti (2003: 526) lists two PD roots: *\*kuḷ* “pit, hollow” (however, the *Paragaramuthali* takes this root back to *\*kul*), as well as *\*kuḷ*, “lake”; cp Tamil *kuḷam* “tank, reservoir, lake,” < PD *\*kuḷ*, which itself is derived from *\*kul* (*\*kul* “to accumulate” > *\*kuḷ* “lake” or accumulation of water).<sup>39</sup>

Another potential source of variation is the letter *-ḷ-* in Dravidian (a retroflex voiced fricative per Zvelebil 1990: §1.7.1),<sup>40</sup> which is transliterated as (“interchanges with”) *ṇ*, *ḍ*, *ḷ*, *y*, and *r* per Caldwell (1875: 59; also Emeneau 1970: 98–99) and I would add “normal” *-l-* as well. OI speakers had no equivalent to

<sup>37</sup> A Dravidian word for family, Tamil *kūḷi* “company, multitude, family” (DED #1915) is fairly late and restricted to PSD1 (the southern branch which excludes Telugu and associated languages; Southworth: 50). It is derived from the PD root *\*kūḷ* (“crowd together, assemble, muster”) which comes from the PD root *\*kuḷ* (“lake”). *Paragaramuthali*, p. 666 [TAMIL VIRTUAL ACADEMY \(tamilvu.org\)](http://TAMIL_VIRTUAL_ACADEMY(tamilvu.org))). See n. 39 below. The usual Dravidian word for family is *kuṭumba* < *kuṭi* “house” (DED #1655) > OI/MI *kuṭumba* “household, family.”

<sup>38</sup> The dates are rough estimates based on Southworth (195).

<sup>39</sup> The etymologies are taken from the Tamil Etymological Dictionary (*Paragaramuthali*, vol. 2 part 2): *kulai* < *kul* (p. 518); *kūḷi/kūḷa* < *kul* (p. 536); *kuḷu* < *kul* (p. 543). For the root *kul* see p. 506, entry 3 “to collect, to accumulate, to heap, to assemble.” The root *kul* itself is derived from *ul* per the dictionary; see vol. 1 part 3, p. 109, entry 3 for *ul* with the meaning “to originate, to appear, to sprout, to grow big.” *kuṭam* < *kuḷ* (p. 344), and *kuḷam* < *kuḷ* (p. 556). Other words to be discussed below are *kuṭṭam* < *kuḷ* (p. 333), *kuṇṭam* < *kuḷ* (p. 377). All are derived from *kuḷ* per the *Paragaramuthali*, with the meaning “lake” (per Krishnamurti 2003: 526; *kuḷ* is not found in the *Paragaramuthali*, only *kuḷam* on pp. 555–56 with the meaning “tank, pond, reservoir, lake”).

<sup>40</sup> The different ways of writing this letter (*ḷ*) in Roman script can be quite perplexing: *ṛ* (Emeneau and Burrow), *ḷ*, (TL and modern writers), *r* (Caldwell), as well as *zh*, and *z*.

the letter *-l-* and approximated it in different ways. So the Tamil word *kuṛu/kuṛām* (*kuḷu/kuḷām*, “herd, flock, crowd”) might be transcribed as *kuḍu/kuḍām*, similar to *Kuḍḍa* (see below).

There are then two basic PD roots *\*kul* and *\*kuḷ*, both related, that account for the four Dravidian words above and the Pāli names *Kula* (“family”), *Kuḷa* (“lake”) and *Kūla* (“pond, pool, heap”). This explains the different spellings and meanings behind these words, in that they are derived from different, but related Dravidian roots. All are united by the general meaning of “accumulation” or “assemble, gathering together.”

The other names for the *thera* are also related phonetically and semantically through the same roots and the propensity in Dravidian (e.g. *kuḷam* “waterpot” ~ *kuṭam* “tank, reservoir” ~ *kuṭṭam* “pond”; see Emeneau 1970: 98–99 for *-l-*, *-ḷ-*, *-r-*, *-ṭ/-ḍ-* correspondences and others) and the Prakrits for the alternation of *-l-* and *-ḷ-* with *-t/-d-* and *-ṭ/-ḍ-* (Pischel §226, §244), and CC > < NC (as in Tamil *kuṭṭam* ~ *kuṇṭam* “pond, pool” and Pāli *santa* ~ *sanna*, “bond, chain”;<sup>41</sup> *Chanda* ~ *Channa*; *ālambana* ~ *ārammaṇa*, “support”), which is not uncommon in both Pāli and Dravidian (also in Gāndhārī, Brough 1962: §46; Geiger §6.3; Norman 2002: 144; Levman 2022: 21).

The noun *kuṇḍa* has several different meanings in Pāli: “water pot, pitcher; hole; iguana (?); bran (?); crippled” (Cone vol. 1: 706; OI “water pot, hole, pit, well, spring”). Here it probably means “water pot” which is derived from the Dravidian *kuṭam* (Tamil “water pot” < *\*kuḷ*; Kannada *guṇḍi*, “jug”; DED #1651/1669, M1 vol. 1: 226); or “spring” < Tamil *kuṭṭam* “pond”; *kuṇṭam* “pool”; Kannada *kuṇḍam* “pit, pool, pond” (also < *\*kuḷ*, see DED #1669 and many more cognates in PD and PCD).<sup>42</sup>

With the meaning “water-pot” the word appears again in Th 15 in the monk’s name *Kuṇḍa-dhāna thera*, which appears to be a bilingual translation (“water pot-receptacle”), the second word *dhāna* being an IA calque of the Dravidian-derived Pāli word *kuṇḍa*; a *kuṇḍa* is also the bowl of Gāṇḍatīriya-tthera (Th 127), which he describes as *chava-sitto’va me patto* (“my bowl is just a sprinkled corpse”) and the commentary explains as *matānaṃ khīra-*

<sup>41</sup> *Nāgo ’va santāni guṇāni chetvā* (Se); *nāgo ’va sannāni guṇāni chetvā* (Be); “Like an elephant who has broken his ties and bonds” at Se Sv 2, 324<sup>13–14</sup>. PTS has *Nāgo va sandāna-guṇāni bhetvā* at PTS Sv 3, 708<sup>23–24</sup>. See Levman 2014: n. 1100.

<sup>42</sup> Compare also OI *kūta* “vessel, heap, multitude” which combines the meaning of *kula* “multitude” and *kuṇḍa* “vessel.” For the meaning “crippled,” see Dravidian *kuṇṭan* “cripple” DED #1688.

*secana-kuṇḍa-sadiso ti attho*, “like a water pot pouring milk on the dead” Th-a 2, 8<sup>31–32</sup>). The word *kuṇḍa* is derived from the same PD root *\*kuḷ* (“lake”) and also takes on the meanings of words derived from *\*kul* (OI/MI *kūla*: “declivity, pond, pool” ~ Tamil *kuḷi*, “pit, pond, well”). *Kuṇḍa* may also be phonetically related to Pāli *kullaka* (“little basket”) and *kulla* “belonging to the family” < OI *kulya*) because of the equivalence of *-ṇḍ-* and *l(l)-* sounds in OI and MI (e.g. *galla-gaṇḍa* “cheek”; *malla-muṇḍa* “tribal name”; *daṇḍa-dala* “stick”; *caṇḍa* “glowing,” *cullī* “fireplace,” etc. (Woolner 1926–28: 67; Levman 2011: 51, also in Levman 2021a: 220). Notice again that all these words (IA *kuṇḍa* and Dravidian *kuṭam*, *guṇḍi*, *kuṭṭam*, *kuṇḍa*) share a meaning based on the root etymology *\*kuḷ* (“lake”), that is “holding water” and *\*kul* “accumulation.”

The name *Kuḍḍa/Kuḍḍha* goes directly back to Dravidian *kuṭam* (“water-pot, well, spring”) and *kuṭṭam* (“pond”), either through the weakening of the geminate *-tṭ-* > *-ḍḍ-* or because *kuṭam* in Dravidian is pronounced *kuḍam*, allophonically weakened between vowels. It is also related, as stated above, to the word *kuḷām*, (“herd, flock, crowd”), which might be transcribed as *kuḍām*, because of the ambiguity of the Dravidian *-ḷ-* sound. The *-la* suffixes for *Kuṇḍa* (*Kuṇḍala*) and *Kuḍḍa* (*Kuḍḍala*) are adjective derivatives in OI (Whitney §1189)—that is, Sanskritization.

All the Dravidian words are much earlier than the first appearance of *kula* in the RV (perhaps around 1200 BCE), going back to the third millennium or mid-to-late second millennium (Southworth: 51, 60, 195, 252, 330; Krishnamurti 2003: 501). Since *kula* has no convincing IE etymology (for some speculation, see M2 vol. 1: 372–73), a Dravidian borrowing seems cogent and is corroborated by the various words and meanings outlined above. In sum, what we have here is an attempt to interpret a whole series of names (eight in total) based on the PD roots *\*kuḷ*, *\*kuḷ* and *\*kul* (*Kula*, *Kuḷa*, *Kūla*) in a foreign phonological system which, because of the ambiguity of the *-ḷ-* sound, the alternation of *-l-* and *-ḷ-* with *-t/d-* and *-ṭ/ḍ-*, and NC > < CC, resulted in several more forms (*Kuḍḍa/Kuḍḍha*, *Kuṇḍa*, *Kuḍḍala*, *Kuṇḍala*) with similar underlying meanings. Since this is very complicated, the following table may help to clarify:

PĀLI VARIANTS: A TYPOLOGY (PART I)

Dravidian	Root	Pāli	Root	Notes
<i>kuḷu</i> “assembly, flock herd”	PD * <i>kuḷ</i> “to accumulate”	<i>kula</i> “family”	<i>kul</i> “to accumulate” (OI)	Pāli root is a Dhātup <sup>43</sup> root only, back-formed from <i>kula</i>
<i>kulai/kula</i> “bunch”	PD * <i>kuḷ</i> , “to accumulate”	<i>kula</i> “assemblage, multitude, herd”	<i>kul</i> “to accumulate”	
<i>kuḷi</i> “pond, well, pit, hollow”	PD * <i>kuḷ</i> “pit, hollow” < PD * <i>kuḷ</i> “to accumulate”	<i>kūla</i> “heap, mound, pond, pool; shore, bank; declivity”	<i>kūl</i> “cover, hide; keep off obstruct” (OI); Pāli, <i>kuḷati</i> “obstruct, constrain”	Pāli root is a Dhātup root only, back-formed from <i>kūla</i>
<i>kuḷam</i> “lake”	PD * <i>kuḷ</i> “lake” < * <i>kuḷ</i> “to accumulate”	not in Pāli or OI; <i>kuḷa</i> in Prakrits alt. form of <i>kula</i>	n/a	Pischel §260. Cp Pāli <i>kuḍuba</i> “measure of volume” (change of -l- > -ṭ/-ḍ-)
<i>kuṭam/kuṭṭam/kuṇṭam</i> “water-pot,” ( <i>kuṭam</i> ) “reservoir, pond, pool”	PD * <i>kuḷ</i> < PD * <i>kuḷ</i> ; -l- > -ṭ-; -ṭṭ- > -ṇṭ- (CC > NC); weakening of -ṭ- > -ḍ-	<i>kuṇḍa</i> (also OI) “spring, well; water-pot, basin, bowl”	< Dravidian <i>kuṭam</i> , Kannada <i>guṇḍi</i> per M1 vol. 1: 226; or < <i>kuṭṭam/kuṇṭam</i>	<i>kuṇḍala</i> , Sanskritization of <i>kuṇḍa</i> with -la suffix (Whitney §1189)
<i>kuṭṭam</i> “pool, pond”	PD < * <i>kuḷ</i> “to accumulate”	<i>kuḍḍa</i> (proper name), <i>kuḍḍha</i>	< Dravidian; aspirate geminate <i>ḍḍh</i> - due to bilingual Dravidian speakers	<i>kuḍḍala</i> , Sanskritization of <i>kuḍḍa</i> with -la suffix (Whitney §1189)
<i>kuṭṭu/kuṭṭi</i> “to pound”	PD < * <i>kuḷ</i> “lake” (per <i>Paragaramuthali</i> )	<i>kuṭṭa/kuḍḍa</i> “wattle and daub wall”	< <i>kottṭeti</i> (OI <i>kuṭṭayati</i> ) “to pound”	Clearly a Dravidian word but etymologically unclear in Dravidian with root * <i>kuḷ</i> “lake”

<sup>43</sup> Dhātup = *Dhātupāṭha* (“recital of grammatical roots”) is the name of an ancient list of roots ascribed to Pāṇini (MW).

There is a Sanskritized version of *pāda* 19-a (*udakaṃ hi nayanti nettikā*, “canal makers lead water”), in the *Udānavarga* 17.10 that reads *udakena nijanti nejakā* (“washer-persons purify with water”), and also a Chinese version (水工調舟船 “sailors control their boats”); for a discussion of these three versions and a reconstruction of the underlying *koiné* which gave rise to the variants, see Levman 2023b: 9.

#### 4.5 *Theragāthā* 22, *Cittakatthera*

##### **Natural, diachronic language change over time**

-bh- > -v-

nasalization, de-nasalization,

vowel lengthening/shortening

##### **Sanskritization**

restoration of -bh- < -v-

##### **Orthographic variation**

presence or absence of nasal and *niggahīta*

##### **Foreign word borrowing**

< Munda word *karam* (*Adina cordifolia*)

*Nīlā sugīvā sikhino, morā Kāraṃviyaṃ* (vars. *Kārambhiyaṃ, Kāraṃviyaṃ, Kāyaṃviya*) *abhinadanti*.

*te sīta-vāta-kīlitā* (vars. *kīlitā, -kalitā, -kaḷibhā, -kadditā*; comm.: *sañjāta-kalitā, kīlitā, kilī, -kalī, kaddita-kalitā*), *suttaṃ jhāyaṃ* (vars. *jhānaṃ, jhāyiṃ*) *nibodhentī ti*.

“Blue, with beautiful necks, the crested peacocks call in *Karaṃvī*; urged on by the cool breeze they awaken the sleeper to meditation” (Norman, 3).

There are four variants in the *mūla* text as above and the commentary has many more. The PTS edition of the commentary reads (with variants in brackets from Sinhalese, Burmese and Thai recensions; capitalization is varied as per the editions).<sup>44</sup>

*Kāraṃviya* (vars. *Kārambhiya*, *Kāraviya*), *kāraṃva-rukkaṃ* (vars. *karambha-rukke*, *kārambha-rukke*, *kāraṃva-rukko*, *kāraṃva-rukkaṃ*). *Kāraṃvī tī vā* (vars. *kāravani vā*, *Kārambhiya*, *Kāravīti vā*) *tassa vanassa nāmaṃ*; *tasmā Kāraṃviya* (var. *kārambhiya*) *kāraṃvi-* (vars. *Kārava-*, *kārambha-*) *nāmake vane ti attho*.

“*Kāraṃviyaṃ*” = the *kāraṃva* tree. Or “*kāraṃvi*” is a name for that forest; therefore “*Kāraṃviyaṃ*” means “in the forest whose name is *kāraṃvi*” (with suitable changes *pari passu*).

This is a hapax legomenon in the canon, only occurring in Th 22, representing the name of a tree or a forest; the Th-a commentary is not sure of which. The word occurs in the Böttlingk and Roth dictionary as *kārambhā*, “name of a plant bearing a fragrant seed (commonly *priyaṅgu*)” but it is not attested in the literature (*Amarakośa* 2.4.2, 26 only). There is also the *karambhā* plant with a short initial -ā- which refers to two plants, *priyaṅgu-vṛkṣa* (*Asparagus racemosus*) and fennel. The word *priyaṅgu* refers to several different plants (q.v.). The only Pāli dictionary which defines the word is the Burmese *Pāli Myanmā Abhidhān* where it is called the *upāsakā* (vol. 5, p. 668), referring to a tree/shrub or forest; the word *upāsakā*, per the Myanmar–English Dictionary (p. 609) refers to the sarsaparilla shrub, *Hemidesmus indicus*. The roots are used as Ayurvedic medicine throughout India, ground into a drink.

The large variety of variants and lack of attestation show that this is a foreign word. For a full list of variants see note 44.

<sup>44</sup> The variations in the recensions are as follows: *Mūla*: Burmese, *Kārambhiya* (Be), *Kāraṃviya*. Sinhalese, *Kāraṃviya* (Ce). Thai, *Kāraviya* (Se), *Kāraṃviya*, *Kārambhiya* (Mc). Commentary: Burmese, *Karambha*. Sinhalese, *Kāraṃviya*, *kāraṃva*, *kāraṃvī*; *kāraṃva*, *kāraṃvī*; *kāraṃva*, *kāravani*, *kārava-nāmake*. Thai, *Kāraviya*, *kāraṃviya*, *kārambhiya*; *kārambha*, *kāraṃva*, *kāraṃvi*, *kāraṃvi-nāmake*. In addition, the CPD lists the following variants not covered above: *Kārambha*, *Kāramba*, *Kārambhī*, *Kārambhiya*, *Kārambiya*, *Kārambhī*. The variants are only given the first time they appear. For vars. *kilitā*, etc., see n. 48.



The alternation of *-bh-* > *-v-* is a dialectal variant per Norman (1989: 373; CP 4: 52), which also occurs in *Vedic Variants*, so it is quite old; however, Bloomfield and Edgerton (1932/79: §220–22) note that the change is “not purely phonetic but always involves tolerable lexical shifts” and that in the change between *abhi* and *vi* “most, probably all” of the former are prior. So it may indeed be a diachronic indication of normal consonantal weakening over time. Von Hinüber (2001: §101) notes the occasional (“vereinzelt”) change in late Pāli (“späten Pali” *abhiyutta* > *aviyutta*) and it is a normal change in Gāndhārī (*-bh-* > *v(h)*) and in the Niya documents, both later than early Pāli, as is the change of Pāli *paṭisambhidā* > BHS *pratisaṃvid*, noted by Norman (*ibid.*) and Lüders (1899: 493; also in 1973: 170). Brough (1962: §44) says that *-bh-* has four possible spellings in Gāndhārī (*-bh-*, *-vh-*, *-v-*, and *-h-*) but notes that it is probable that *-bh-* is “merely a historical spelling” for *-vh-*; in other words it was pronounced as a glide, regardless of how it was written. It is likely that in the *koiné* or common form underlying Pāli it was pronounced as a glide (*-v-*) or aspirate only (*-h-*) as in Gāndhārī, i.e. that *Kāraṃviya* or *Kāraṃva* was the earlier form and the change to *-b-* or *-bh-* was a Sanskritization, restoring the word to its “original” form; this is at least partly confirmed by Sn 443, where variant *vecchāmi* (“I will break”)<sup>45</sup> appears opposite *Mahāvastu bhetsyāmi* (Pāli *bhecchāmi*, “I will break” Mv 2, 240<sup>11</sup>). Here the Mv is definitely later than Pāli and, therefore, a Sanskritization.

The variation between *-b-* and *-bh-* is either because bilingual Dravidian speakers couldn’t hear the aspiration, which is allophonic in their language, or the restoration of *-v-* > *-b(h)-* was just a guess, the two being almost identical. The only other variation of any significance is the *niggahīta -ṃ-* in the second syllable, which is omitted in two cases (*Kāraṃviyaṃ* and *Kāraṃvi*); the variation in orthography of writing a pure nasal (*-m-* or a *niggahīta ṃ-*) before the labial consonant is only a minor point of spelling. Usually there is no *niggahīta* if the nasal is of the same class as the consonant (which *-m-* and *-b-* are); the consonant *-ṃ-* before *-v-* would usually have the *niggahīta* underdot; and sometimes both the nasal and *niggahīta* is omitted, due to oral or written transmission anomalies (as in Gāndhārī, Brough 1962: §48, “sporadic weakening or loss of nasal before voiced consonants”), or because

<sup>45</sup> The Sn editor (p. 77) takes *vecchāmi* as derived from the root *vyadh*, but that form would be *vacchāmi* (*vatsyāmi* in OI).

the vowel was nasalized anyways, but not so written (Fussman 1989: 478).<sup>46</sup> The variant of intervocalic -y- with *Kāyaṃviya* is probably just a mistake, as intervocalic -r- never changes to -y-.

This leaves the question of the derivation and meaning of the word. There are three suggestions (PMA (*Pāli-Maynamā Abhidhān*) = *kārambhā* = *upāsāka* = Indian sarsaparilla; OI *kārambhā* “*Asparagus racemosus*, fennel”; OI *kārambhaka*, “a kind of *Achyranthes*”), for what the tree/shrub and/or forest might represent, none of them more compelling than the other. I cannot find any reference to the *upāsakā* plant in the canon or commentary or in any dictionary except the Burmese, and they do not give any reference; in addition the PMA provides the wrong spelling *upāsakā* as a Buddhist lay devote (male), it should read *upāsaka*, and the female *upāsikā* (609).

The word itself appears to be derived from the Munda word *karam* which is a popular festival involving cutting branches off the *karam* tree (*Adina cordifolia*, another possible meaning of *kārambh-*, but at least one with some phonetic correspondence).<sup>47</sup> See Bodding 1929–36: vol. 3: 451–52 for a detailed write-up on the festival, where the tree plays the central role. There is no difference in Munda between *kāram* and *kāram*, as vowel length is not phonemic. The word exists today in Santali, Kharia, and pre-Mundari (*karam* = “a kind of tree”) and there are similarities in other Munda languages, like Bondo *kumbi* “tree”; Korku *kumbi* “variety of tree”; Ho *karam* and *kuumba*, idem, and a near exact correspondence in Juang *kərambə* “a kind of large tree”; see also Juang *gombari* “gambari tree” and *qumburi* “tree” noting Kuiper’s observation of the “many Munda synonyms with varying initial gutturals, dentals and labials” which he calls “rhyme words” (1948a: 7). Note that some words preserve the nasal and others omit it; also aspiration is not phonemic in Proto Munda (PM). There are also some echoes of the AA affiliation of the word in Mon-Khmer (Shorto 2006: §1570 and §935) *\*[d]ker* “tree, plant” and *\*krwaŋ*, “a kind of spice-yielding tree.” Witzel (1999: 8) also considers the word to be AA in origin

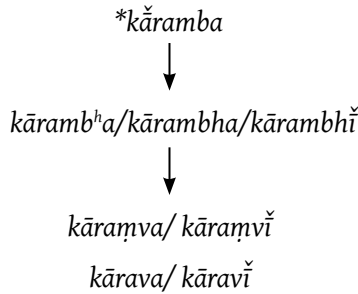
<sup>46</sup> “Toute syllabe ouverte lourde comporte des vibrations nasales, que sa voyelle soit une voyelle étymologiquement brève et nasalisée ou que sa voyelle soit une voyelle étymologiquement longue, étymologiquement non nasalisée ...” (“Every heavy, open syllable was nasalized whether the vowel was etymologically short and nasalized or whether the vowel was etymologically long and not etymologically nasalized.”)

<sup>47</sup> Dhammika 2015: 63 says that the *kāra* tree is the “curry leaf tree.” It occurs in the *Jātakas*.

but he is referring to the word *karambha* with two short -ă-s (*karambha*, meaning “groats”), which, however, must be related to *kārambhă*.

In summary then, the word is an unknown native tree with a Munda pedigree, imported into OI and MI and subject to varied changes according to its interpretation by MI speakers. Since neither aspiration nor vowel length are phonemic in PM (nor does PM have a letter -v-) the earliest form of the word was probably *kāramba* to which aspiration was added (*kāramb<sup>h</sup>a*/*kārambha*), and the long vowel confirmed because of accent; the word was then subjected to consonantal weakening (-bh- > -v-, *kāramva*), vowel nasalization (*kāraṃva*), or loss of nasalization (*kārava*), and variant vocalic interpretation (*kāravī*). This accounts for most of the variation that we find above.

This may be charted as follows, with the vertical order representing a very rough relative timeline:



The change of -b- > -bh- > -v- is diachronic; the (de-)nasalization and vowel changes cannot be dated.

#### 4.6 *Theragāthā* 22

##### Diachronic change over time

-ḍ- > -ḷ- > -l-

-t- > -ṭ-

##### Sanskritization

-ṭ- > -t-, -ṭ- > -b- > -bh-

##### Different interpretation of an underlying *koiné*

*kaḷiyā* > *kaḷibā* > *kaḷibhā*, *kaliyā* > *kilitā*

##### Foreign word borrowing or influence

< Dravidian *kali*, *kaḷi*, *kaṭi*

##### Orthographic variation

-l- and -ḷ-

*Nīlā sugīvā sikhino, morā Kāraṃviyaṃ abhinadanti.  
te sīta-vāta-kīḷitā* (vars. *kīlitā*, *-kalitā*, *-kaḷibhā*, *-kadditā*; comm.: *saṅjāta-kalitā*, *kilitā*, *kilī*, *-kalī*, *kaddita-kalitā*), *suttaṃ jhāyaṃ* (vars. *jhānaṃ*, *jhāyimaṃ*) *nibodhentī ti*.

“Blue, with beautiful necks, the crested peacocks call in Karamvī; urged on by the cool breeze they awaken the sleeper to meditation” (Norman 3).

In *pāda* c of the verse, there are six variants for the word *kīḷitā*: *kīḷitā*, *kilitā*, *kīlitā*, *kalitā*, *kadditā*, and *kaḷibhā*. The commentary has the additional variants *-kilī* and *-kalī*.<sup>48</sup>

These all have different meanings:

**var. #1 *kīḷitā*** “having played, having sported” (p.p.); “play, sport, amusement” (noun) < Pāli *kīḷati*, ~ OI < *krīḍati*/*krīḷati* < root *krīḍ*/*krīḷ*, “to play, to sport” (RV, earliest occurrence), but perhaps from the causative *kīḷeti* > *kīḷitā* (same p.p.p. as non-caus.) “made to play, made to sport.” *kīḷitā* is a variant with short -ṭ-.

<sup>48</sup> Mūla: Burmese, *kīḷitā* (Be), *kīlitā*, *kaḷibhā*, *kadditā*; Sinhalese, *kaddita-kalitā* (Ce), *kalitā*, *kadditā* (comm.); Thai, *kalitā* (Se), *kīlitā*, *kīnitā* (Mc). Commentary: Ce *-kilī*, *-kalī*; Se *kīḷitā*. Be *kilitā*.

**var. #2 *kīlitā*** “having played” < Pāli *kīlati* (Cone vol. 1: 691, “is joyful, plays?” [? in original]), OI *kīlati*, not attested in the literature, just in the dictionaries.

**var. #3 *kīlitā*** “bound, tied” < Pāli *kīlati* “to bind, fasten” OI, idem, not attested in Pāli or OI literature, just in the dictionaries.

**var. #4 *kalitā*** “impelled, driven, urged, made, formed, furnished, divided, sounded indistinctly” < Pāli *kalati*, OI < idem “to impel, incite urge on” (MBh). Compare OI *kaḍ/kaḍita*, “to be confused or disturbed by pleasure or pain” (change of -ḍ- > -ḷ- > -l- per Pischel §240).

**var. #5 *kadditā*** “made an unpleasant noise” < Pāli *kaddati*, OI < *kardati* Dhātup only, “rumble, caw, make any unpleasant noise.” Not attested except in dictionaries.

Compare also *kaḍḍati/kaḷati* “is elated”: OI *kaḍati* “to be confused, disturbed by pleasure or pain; to be elated or intoxicated” and OI *kaḍḍati* “to be harsh or severe”; Dhātup only for all.

**var. #6 *kaḷibhā* ?** “appearing in anger” < *bhā* “to shine, appear, show, manifest” + *kali* “anger, strife, discord.” Possibly an *upapada-tappurisa* (like *loka-vid*, “knowing the world”).

Note that the PTS editors (Oldenberg & Pischel 1883/2006: 4) have *kalitā* in the main text, so they have decided that was the original (or at least a better) reading. The Burmese has *kīlitā* and only lists two variants (*kaddita-kalitā* and *kalitā*).

So the possible translations are:

1. “Caused to play by the cool wind ...” *kīlitā*, *kilitā*. vars. #1, #2.
2. “Bound by the cool wind ...” *kīlitā* (with var. *kīlitā* with vocalic -l-). var. #3.
3. “Urged by the cool wind ...” *kalitā*. var. #4.  
(Norman 1969/95: 3, “urged on by the cool breeze”; Rhys Davids (1913: 27) has “by cool and humid winds made musical” taking her cue from the commentary *madhura-vassitam vassantā* (Th-a 1, 82<sup>11</sup>), translated as “musical call”; lit.: “uttering a sweet call”).
4. “Having made an unpleasant noise (or causative, “an unpleasant noise having caused to be made” (*kaddayitā*, *kadditā*) because of the cool wind (or along with the cool wind) ...” *kadditā* < *kaddeti*. var. #5.
5. “Appearing in anger because of/along with the cool wind ...” *kaḷibhā* (and should be *kalibhā* with regular l-). var. #6.
6. In addition Mc has the var. *kīnitā* (presumably v.l. for *kīnitā* < *kīṇāti* “to buy, purchase”), “Bought by the cool wind ...” which makes no sense in this context; also the commentary has two variants *sañjāta-kilī ti* and *sañjāta-kalī ti*. *kali* means “distress” which is the opposite of *madhura vassitam* (no. 3 above), so this may be a mistaken shortening for *sañjāta-kalitā* (“sweet sound urged on, produced by the cool wind”). *kili* is a clicking sound which is presumably made by the peacocks (“the sweet sound of clicking ...”).

Now the commentary, itself with many variants—some of which Oldenberg and Pischel call “nonsense”—seems to be clear on one point: that the peacocks are disturbed by the roar of the storm clouds (*megha-gajjitam*) and they in turn raise a noise (*abhinadanti*), drowning out the other birds and waking up those sleeping.

*pāvusa-kāle megha-gajitaṃ sutvā: ke-kā-saddaṃ karontā utu-sampadā-siddhena sare haṃsādike abhibhavatā viya nadanti.* Th-a 1, 827-9, “Having heard the roar of the clouds at the time of the rainy season, uttering a *kekā* (cry of a peacock) sound, because of the weather, they roar as if drowning out the sounds of the swans and others.

All the translations except nos. 2 and 6 (“bound by...”; and “bought by...”) would work, whether one takes the peacocks as simply urged by the roar of the storm, taking it in a positive sense as Rhys Davids does (“roused to make music”), and as *kīlītā* seems to suggest (“roused to play”), or whether one takes it in a negative way, as nos. 4 and 5 do (and *-kali* in no. 6)—the peacocks are disturbed and/or angry because their tranquility has been broken by the sound emanating from the clouds. The word *sīta* itself (*sīta-vata-*, OI *śīta*, “cold, cool, chilly, frigid”) suggests an unpleasant awakening.

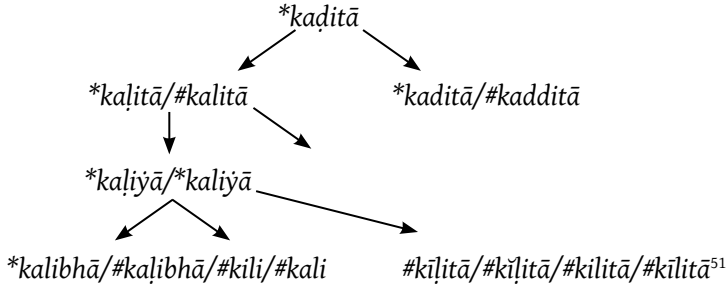
In trying to establish the original (or at least earliest) reading, one may hypothesize, then, that it was ambiguous (in terms of being semantically positive or negative), suggesting the verb *kaḍ/kaḍati* (“to be confused, disturbed by pleasure or pain; to be elated or intoxicated”) p.p. *kaḍitā* as the original reading, or its MI cognate *kaḷitā*; OI *-ḍ-* regularly changes to MI *-ḷ-* (Pischel §240/244; Geiger §38.6;) and often this is written as *-l-*, or *kalitā*, the variant chosen by the PTS editors. This alternation of *-ḷ-* and *-l-* is not just an orthographic issue, but is also due to the phonetics of Dravidian languages. In any case, *kalitā* with normal *-l-* primarily means “impelled” whereas *kaḷitā* has the meaning of “to be confused, disturbed by pleasure or pain; to be elated or intoxicated.” The change of *kaḷitā/kalitā* > *kīlītā/kilitā* arises with a change of *-a-* > *-ə-* (interpreted as *-i-*).<sup>49</sup> One or more of these variants would presumably substitute the intervocalic *-t-* for a *-y-* glide (Pischel §186, e.g. *kaḷiṃya/kaliṃya*), which, with the substitution of an intervocalic *-b-* and its aspiration<sup>50</sup> would result in *kaḷibhā/kalibhā*, the sixth variant. Variant #5 *kaddita* (“an unpleasant noise caused to be made”) may have been derived from the original *kaḍitā*

<sup>49</sup> Pischel §101 *-a-* > *-i-* (schwā, *-ə-*) in syllables before the accent. This assumes that *kīlīta* (“caused to play”) is interpreted as a causative verb *kīlēti*, p.p. *kīlīta* with the accent on the second syllable (which is always the case with a causative verb, cp Whitney §1041).

<sup>50</sup> My working assumption is that bilingual Pāli speakers whose first language was Dravidian (and who were not yet skilled in MI) were probably unsure if a stop was aspirated or not, because it was not phonemic in Dravidian languages.

form with the mixing of dental and retroflex consonants (Geiger §41.3; Pischel §218, 225) because of dialect ambiguities, or the influence of Dravidian verb forms (see below).

So a possible sequence may be hypothesized as:



(Those marked with \* are reconstructed, those marked # are attested variants in the *mūla* or *aṭṭhakathā* transmissions.)

For the reader who finds this reconstruction overly complex, the following Dravidian analogies may help to convince him/her; for Dravidian has three very similar verbs and near-homonymic forms which appear to have influenced the Pāli and which confirm the priority of the *kaḍi-* (*kaṭi-*) and *kaḷi-* (*kali-*) forms.

Potential Dravidian correlates:

1. *kaḷi* “to be intoxicated, be in rut, exult, rejoice, be proud” (DED #1374) < root *kaḷ* “liquor, honey, toddy”. Past participle *kaḷitta*. In Old Tamil (OT) *taṇ-vaḷi* (= *sīta-vata*) *kaḷitta*, “rejoicing because of the cool wind.” This is close to the meaning of *kaḍita* and *kaḷita*, “being elated or intoxicated by the cool wind.”

<sup>51</sup> This derivation chart works well in terms of the diachronic, phonological changes. Stefan Karpik (pers. comm.) suggests that *kīlitā* should be the earliest reading, as it is an early, attested Vedic form. It is a possibility, but requires a different interpretation of the passage’s meaning. The Dravidian correspondences also reinforce the interpretation above (with *\*kaḍitā* as the earliest reading); against *\*kaḍitā* as the earliest reading is the fact that it is an unattested Dhātup verb (< OI *kaḍ*, *kaḍati*). Nevertheless it works well semantically, provides a cogent diachronic source for all the variants, and is closest to the postulated Dravidian correlates.



2. *kaṭi* “to reprove, rebuke, chide” (DED #1126; pronounced *kaḍi* in Dravidian) < root *kaṭu* “to be angry”. Past participle *kaṭinta*, earlier, pre-Tolkāppiyam (Tolk) form *kaṭita* (Paramasivam 1979: 4 note 4; pronounced *kaḍita*).<sup>52</sup> Old Tamil *taṇ-vaḷi kaṭita*, “having rebuked the cold wind.” This is close to the meaning of #4 and #5 where the peacock is angry with the wind for disturbing it and responds noisily.
3. *kali* “to sound, clamour, roar; to flourish; to appear; to rejoice” (missing in DED but appears in Sangam literature, e.g. *Aiṅkuruṇūru* (Aiṅk) 65–1 *kalitta āmpal*, “flourishing waterlilies”). Past participle *kalitta*, Old Tamil *taṇ-vaḷi kalitta*, “clamouring/roaring because of the cold wind or “appearing/rejoicing because of the cold wind.” This verb has either a positive or negative meaning. The present participle of *kali* is *kalippa* which may be the source of the strange form *kaḷibhā* (var. #6 above).

So in Dravidian we have three near homonyms *kaḷitta*, *kaṭi(n)ta* and *kalitta* whose meaning fits the context and appears to have influenced the meaning (and confusion of meaning) in the Pāli. It confirms that the original or earliest transmission was probably a root beginning with either *kaḍi-* (*kaṭi-*) or *kali-/kaḷi-*. All of these are old roots in Dravidian,<sup>53</sup> whereas IA words like *kadditā/kaditā/kilitā* are either from the lexicography only (unattested except for the Dhātup), or quite young (e.g. *kalitā*, MBh, later than their appearance in Th). The reconstruction suggests that the Th 22 is either a translation from a poem originally written in Dravidian or, at the very least, that the large number of variants arose because of these three Dravidian roots, for which a bilingual Dravidian speaker tried to find an equivalent in Pāli, with the results that we see, variant after variant, each with a different meaning, but all apparently going back to the Dravidian source. In the end it is probably impossible to

<sup>52</sup> The past marker *-nt-* occurs in class 4 and class 12 verbs; the past marker *t(t)-* or *-ṭ-* occurs in classes 1, 2, 6, 7, 9, and 11, confirming its priority. See Krishnamurti 2003: §5.3 where the nasal is an optional element, (N)P signalling intransitive, and geminate (N)PP signalling transitive verbs.

<sup>53</sup> *kaḷi* is widespread amongst all the Dravidian languages. *kalippu* is the Old Tamil word for “intoxication, pride, delight” (cp *kaḷibhā* above). For Sangam appearances of *kaḷi* and derivatives, see Anon., *Index des mots*, vol. 2 1968: 486–87. For *kaṭi* and derivatives 423–24. For *kali* 464–65.

reconstruct the exact sequence of events in this *gāthā*; but it does appear very probable that a Dravidian influence is at least in part responsible for all the variant confusion.

The cry of a peacock (a word of Dravidian origin, cp Tamil *mayil*; Levman 2021a: 171–73) is particularly loud and grating; I would hardly describe it as “musical” as Rhys Davids suggests. Based on all the above I would translate:

“Disturbed by the chilly wind, the crested, blue-necked peacocks screamed raucously in the *Kārambhī* forest, waking those sleeping to meditation.”

This translation uses var. #5 above, *kaddita*, which includes the sense of its correlate *kaḍita/kaḷita* and incorporates some of the meaning of all three Dravidian terms (*kali/kaḷi/kaḍi*), transferring some of the sense of these words to *abhinandanti* in *pāda* b.

#### 4.7 *Theragāthā* 26, *Abhayatthera*

##### Orthographic variation

-ă- for -ā-, -b- for -v-, *saccha-* for *sacca-*

##### Different verb forms of the aorist

-(a)*vijjhim*, and *avyādhim*

##### Diachronic change/dialect change/bilingual influence

-t- > -dh- (*avyātsīt* > *avyādhi*)

##### Commentary misplaced in the *mūla*

*paccavyādhi*, *sacchavyādhim*, *saccappādī*

or

##### Oral transmission confusion

*pacca-* mis-heard as *sacca-*

*Sutvā subhāsitaṃ vācaṃ, buddhassa-adicca-bandhuno.  
paccavyādhim* (var. *paccabyādhim*, *sacchabyādhim*, *saccappādī*,  
*saccabyādi*, *paccavyādhim*, comm. *paccavyādhim*, *paccabādhim*,  
*saccavyādhi*, *paccabyadhim*) *hi nipuṇaṃ, vāla-aggamaṃ usunā yathā’ ti*.<sup>54</sup>

<sup>54</sup> *Mūla*: Burmese, *sacchabyādhim*, *saccabyādi*, *paccabyadhim* (Be), *paccavyādhim*. Sinhalese, *paccavyadhim* (Ce), *saccappādī*. Thai, *paccabyādhim* (Se). Comm.: Sinhalese, *paccabādhim*, *saccavyādhi*. Thai, Sinhalese, and Burmese, *paccabyadhin ti paṭivijjhim*, PTS *paccavyādhin ti* (var. *saccavyādhi*, *paccabādhim*) *paṭivijjhim*.

“Hearing the well-spoken utterance of the Buddha, the sun’s kinsman, I pierced the subtle thing indeed, as one pierces the tip of a hair with an arrow” (Norman, 4).

The verbs with their various meanings and derivation:

**var. #1 *paccavyādhim*** < verb *paṭivijjhati*

Note that the OI form *prativyadh* “to shoot against, hit, wound” (RV) has changed in meaning somewhat: cp Pāli “pierce penetrate, see into, comprehend” (Cone vol. 3: 129) and BHS *pratividhyati* “penetrate, reach, attain, comprehend.” The aor. 1<sup>st</sup> sing. of *paṭivijjhati* = *pratyavyādhim* aor. *prati* + *y* + *-avyādhī* (OI *-avyātsīt*, sigmatic aorist). This root *vyadh* has two aorist forms (Whitney §767), one based on an abbreviated form of the root (*vyadh-* > *vidhyate*, a class 4 verb), and one based on the root *vyadh-* itself: 1) aor. *avidhyi/-avidhyim* > Pāli *-(a)vijjhi/(a)vijjhim*, and 2) based on the root *vyadh*, *avyādhī/avyādhim* (*prati* + *y* + *avyādhim* in OI > *paccavyādhim* in Pāli; although per Fahs 1985: 369 this form of the aorist does not usually take the augment). The form *paccabyādhim* is simply a variant spelling with a short *-ā-* (an error apparently, although Cone gives it as a possible variant) and *-b-* substituting for *-v-*, fairly common, especially in the Burmese recensions.

**var. #2 *sacchabyādhim*** “I have penetrated the truth,” with *sacca* misspelled as *saccha*, and a short *-a-* which is normally a long *-ā-* in sandhi > *sacca-abyādhim* = *saccābyādhim*.

**var. #3 *saccappādīhi*** corr. to *-dhīhi*.

Presumably this is a spelling mistake for *sacca-apādi(m) hi* (“indeed I have attained truth” *saccāpādim*) < root *pad*, usually with the prefix *paṭi-*, aorist *paccapādi*, *paccupādi*, or *paṭipajji*. I am not clear on what the editors say in the footnotes “*saccappādīhi* (corrected into *-dhīhi*).” They (Oldenberg & Pischel) seems to be taking the *hi* particle as part of a 2<sup>nd</sup> pers. sing. imperative suffix of the verb *pra* + *dhi* “long for, strive after, look out, be on the watch,” *sacca-pradhīhi*, “strive after truth.” However, the verb is not attested in Pāli (although it does occur in the RV); since it is a class three verb (*dīdhīte*), the imperative form would be *didhīhi*, which presumably

could be shortened to *dhihi* before a prefix, *pra-dhihi*. However, the commentary specifically says the *hi* is merely a particle or a particle signifying cause (*nipāta-mattaṃ ... vā hetu-aṭṭhe nipāto* Th-a 1, 89<sup>1-3</sup>) and equates *paccavyādhim* = *paṭivijjhim* (“I have penetrated,” the alternate aorist form, see above var. #1). Plus the context, explained in the *kā uppatti* story, makes it clear that Abhaya is proclaiming his attainment of arhathood, so it should be first person. For treatment as a possible instrumental plural see below (discussion).

**var. #4 *saccabyādihi*** = *sacca-byādi hi*. The same as var. #2 above, with the *hi* particle joined to the verb, and with *sacca* spelled correctly, “Indeed he has penetrated the truth.” Should read *sacca-abyādi hi*, with the aspirate *-dh-* and augment *a-* for the aorist.

**var. #5 *paccavyādhimhi*** = *paccavyādhim hi*, same as var. #1 above, with the *hi* particle joined to the main verb. “Indeed I have penetrated the truth.” Should read *paccavyādhim hi*.

**var. #6 commentary: *paccavyādhin ti*** (PTS vars. *saccavyādhi*, *paccabādhim*) *paṭivijjhim*. The gloss *paṭivijjhim* is the alt. form of *prati + vyadh* (as above, var. #1) without the augment. *saccavyādhi* is a form of var. #4 above. *paccabādhim* is a different verb < *bādhati*, “oppress, afflict” *pratibādh* “to check, restrain, ward off, repel” which makes no sense in the context; it is just a mistake for *paccabyādhim*. Se, Ce and Be have *paccabyadhim* with a short *-ā-* which is also a mistake for *paccabyādhim/paccavyādhim*. The gloss *paṭivijjhim* also has a var. *paṭipajjhi(ṃ)* (“I have entered upon a path” < *paṭipajjati* ~ OI *pratipadyate*).

## Discussion

Here we have eight different variants of the main verb *paccavyādhim* (“I have penetrated, understood”) from the verb *paṭivijjhati*; the OI root (*prati + vyadh*, class 4) has a different meaning (“to shoot against, hit, wound”), as the meaning changed in Pāli and BHS. Whitney (§767) notes that the root is abbreviated to *vidh-* which is how it appears in BHS (*pratividhyati* “penetrate, reach, attain” = Pāli *paṭivijjhati*). There are two forms of the aorist here from both roots *vyadh-*

and *vidh-*. *paccavyādhim* is from the sigmatic 3<sup>rd</sup> person aorist *prati + avyātsīt* = *pratyavyātsīt* > *paccavyāti* > *paccavyādhī*, or *paccavyādhim* (1<sup>st</sup> person) and has the variant with a *-b-* instead of a *-v-* (usually shown this way in the Burmese), so *paccabyādhim*); note the change of *-t-* > *dh-*, which may be dialectal or more likely is caused by bilingual Dravidian speakers who automatically pronounced the intervocalic *-t-* as *-d-*, and add the aspirate which is an allophone of *-t<sup>h</sup>-*. The root with *vidh-* does not take an augment here (although it could) > *pratividhyi* > *paṭivijjhi* in Pāli or *paṭivijjhim* in the 1<sup>st</sup> person.

The commentary glosses the word *paccavyādhim* as *paṭivijjhim*, which is, as we have seen, just an alternate form. There are two variants of this ilk, *paṭipajji* the aorist of *paṭipajjati* “practice, follow a path” which would fit the context and is phonologically related and the verb *paccabādhinti*, I assume the *ti* is simply an end quote mark, i.e. *paccabādhim* < *pratibādhati* “to oppress, afflict” which makes no sense in the context, so it must be another variant of *paccavyādhim*, with the *-y-* missing; in other words, a spelling mistake.

This accounts for all the variants except the three starting with *sacca-/saccha*. The editors (p. xiii, note 1) call this a “blunder” which is, however, common to all three manuscripts (two Burmese and one Sinhalese), which they opine are derived from the same original, now lost. These are *sacchabyādhim* (ms A). *saccabyādhīhi* (ms B) and *saccappādīhi*, corrected by the editors to *saccappādīhi* (ms C).

The simplest way to rationalize these variants is to identify a separate tradition with *sacca + vyadh*, i.e. “I have penetrated the truth” (var. #2 above). The *hi* particle is then detached from the verb and treated as a separate emphatic (“certainly, truly”) with *saccabyādhī* and *saccappādīhi* treated as above (var. #3 and var. #4). The other possibility is to treat *saccappādīhi* as an imperative, discussed above (var. #3), which does not seem very cogent. A third possibility is to treat the latter two variants (*saccabyādhīhi* = *saccabyādhīhi* and *saccappādīhi*) as instrumental plural commentarial glosses, which have got mixed up in the *mūla* text.

In this last scenario, *sacca-byādhīhi* is a compound in the instrumental meaning “truth and disease/illness/sickness” (*satya + vyādhī* in OI); *sacca-ppādīhi* is also a *dvandva* in the instrumental plural meaning “with truth and great intelligence” (*satya + pradhī* in OI, verb, “strive after”; noun “great intelligence; *padhī* is not attested in Pāli), or it might be construed as a *tatpuruṣa/tappurisa* compound ending in a verb root (like *loka-vid*, “world-knower”), with the meaning “by those who strive after truth.” This would then be a commentary on *paccavyādhim* mistakenly written into the *mūla* text. The realization takes place for “those who

strive after truth” or “I realize by means of truth and great intelligence.” The same explanation goes for the first compound (*sacca-byādhīhi*): “I have realized (the Buddha’s teaching) by means of truth and suffering ...”, a reference to the four noble truths. The commentary also equates *paccavyādhim* with penetrating the subtle four noble truths (Th-a 1, 89<sup>4</sup>: *paccavyādhim nipuṇaṃ catu-saccaṃ*, the truth of cessation or just the four noble truths (*nirodha-saccaṃ*, *catu-saccaṃ eva vā*; 89<sup>2-3</sup>). It seems highly unlikely that all these were just spelling mistakes, occurring independently in three different manuscripts, copied blindly from a lost manuscripts and never corrected by any scribe; one could have made the mistake, but others would certainly have corrected it. However, since these are all hapax legomena in the canon (and they are not listed in the PMA), the argument that they are a mistake is certainly possible.

**4.8 Theragāthā 31, Gahvara-tīriya-tthera** (with vars.: *Tahūrati-tīriya*, *Tahūra-tīriya*, *Gahva-tīriyo*, *Gavha-tīriyo*, *Gahava-tīriyo*; commentary: *Gaṅgā-tīre*, *Gavhara*, *Nahūrati-tīriya*, *Gahura-*, *Gahavara-*)<sup>55</sup>

#### Diachronic change

-hv- > -vh-

#### Foreign word borrowing

*Gahva-* < Dravidian *kaḥpa*, *kaḥuvu* “to wash, purify.” *Gaṅgā* < proto Mon-Khmer \**kaŋ* “transverse, to branch, stretch horizontally” < Dravidian *kavar*, “to separate into various channels, to bifurcate (as a tree or river)”

#### Oral transmission confusion

metathesis of *kaḥuvu* > *Gavuru*

#### Commentarial gloss in the *mūla*

*Gaṅgā-tīre*

*Phuṭṭho ḍaṇṣehi makasehi araññasmiṃ brahā-vane  
nāgo saṃgāma-sīse va sato tatra-adhivāsaye ’ti.*

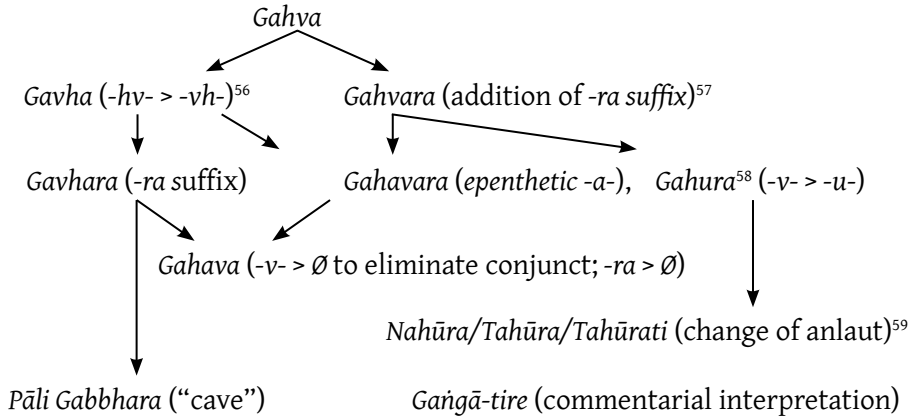
“Tormented by gnats and mosquitos in the forest, in the great wood, like an elephant in the van of the battle, one should endure there mindful” (Norman, 5).

<sup>55</sup> Burmese: *Gahvara-* (Be), *Gahva-*, *Tahūrati-*, *Tahūra*, *Gavha-tīriyo*, *Gaṅgā-tīre*. Sinhalese: *Gabbhara* (Ce), *Gahava-tīriyo*, *Gavhara-*, *Nahūrati-*. Thai: *Gahura-* (Se), *Gahavara-*.

The monk's name comes from the place where he lived, *Gahvara-tīra*, which sounds like the shore of the *Gahvara* river. *Gahvara* is a Sanskrit word meaning “cavern, harbour, bower” (TS, 4<sup>th</sup>–5<sup>th</sup> century BCE), “place, thicket, wood” (*Atharvaveda*, MBh, *Rāmāyaṇa*) and cavern (lex.); M1 vol. 1: 332 and M2 vol. 1: 481 connect it to *gahana*, “deep dense, impenetrable, abyss, depth” which they connect to *gambha* “depth” <IA root \**gambh*/\**gabh* “deep” with *-bh-* > *-h-*. The Pāli equivalent is *gabbhara* (“cave, cavern” Sn; AMg *gabbhiya* “hollow”). However, this root does not explain the *-hv-*/*-vh-* conjunct which is an integral part of the variants.

There are 11 variants, an astonishing number which points to a possible foreign source.

Purely on technical (diachronic change) grounds we can postulate the following (very approximate) time sequence:



<sup>56</sup> Woolner 1928/96: §54, *-hv-* > *-vh-* > *-bbh-* (e.g. Pāli *gabbhara*, “cave”). Some examples of a similar Prakrit change, *-hm-* > *-mh-* changes may be found in the Asokan edicts with the word *brāhmaṇa* > *bāmhaṇa* (Levman 2014: 362).

<sup>57</sup> Whitney §1188. For epenthetic insertion, see Campbell 2004: §2.7.2. Change of *-v-* > *-u-* is called *samprasāraṇa* (see MW) and is not time correlated; for possible change of *-ava-* > *-u-* (*Gahavara* > *Gahura*), see von Hinüber 2001: §139.

<sup>58</sup> There is a river called the Gaula/Gaura which *Gahura* may refer to. It is a branch of the Ganges, starting in the Lesser Himalayas, joins the Ramganga river and then the Ganges. The *Gahvara* is probably just another dialect name for it.

<sup>59</sup> Note Kuiper’s observation of the “many Munda synonyms with varying initial gutturals, dentals and labials” which he calls “rhyme words” (1948a: 7).

Since there is no known IE derivation for Gahvara, one must look to the indigenous languages. Here we find two possibilities in Dravidian and Munda. The more compelling source is Dravidian where the phonetically equivalent word to Gahva is found in Kui a central Dravidian language (*kahpa*<sup>60</sup> “to lave, anoint, wash the face”) with cognates in all the other language branches (DED #1369): Tamil *kaḷuvu* “wash, rinse, purify”; Telugu *kaḍḍu*, *kaḍuvu* “wash, scrub, bathe”; etc. The meaning is also appropriate for the name of a river. With a presence in PSD, PCD and PND the word has a very old pedigree extending back into the third millennium BCE per Southworth (2005: 195); since the first appearance of *gahvara* (in the meaning of “thicket, wood”) is the AV (perhaps 900 BCE; in the meaning of “deep” it first appears in the TS, perhaps 500 BCE), its Dravidian usage pre-dates OI by perhaps one thousand years. While it is impossible to reconstruct a PD root, the Kui correspondence is intriguing and the Tamil word *kaḷuvu* with metathesis > *Gavuru*<sup>61</sup> would provide another logical variant. Dravidian also has the root *kavar* (DED #1325), “to separate into various channels, to bifurcate (as a tree or river)” which is another possible source, also with a wide distribution. If OI *gahvara*/Pāli *gabbhara* are indeed derived from a native word (of which *kahpa* is one reflex), one would then have to re-evaluate the traditional association (as in DPD) of Pāli *gabbhara* with *gabbha*/*garbha* (“womb”). Mayrhofer uncertainly identifies OI *gahvara* with *gambha* (“depth”) and *gahanah* (“deep”; M1, vol. 1: 324, 332).

Since the Gangā river is one of the variants (or a commentarial interpretation), one wonders if there is a phonetic relation to the *Gahvara*. The name *Gaṅgā* is “probably a foreign (Austro-Asiatic?) river name” per M1 vol. 1: 313, for which cognates appear in Indo China and south China. Sanskritized (and etymologized) to *Gaṅgā* because of the similarity to the root *gam* (MW *Gaṅgā* = “swift goer”). The word conceals a pre-IE river name per M2 vol. 1: 457. Compare Munda names for river, *gāda*, *ga'ḍa*, *gaṛa* (Santali); *gada* or *gaḍa* (Korku); *gaḍaṭala* “river bed” (Korku); *gaḷa*/*gaḍa*: “river” (Ho); *gaṛ'ha*, *ga'ḍa*, *ga'ṭa* (Mundari); *ga'ṭa* (Ho); all with the meaning “river.” Compare also AA cognates in Shorto 2006: 496, Vietnamese *ngánh* “branch of a river” < Khmer *ka:ŋ* “to spread,” proto-MK (Mon-Khmer) *\*kaŋ*; *\*kaaŋ*; *\*kaiŋ* ]; *\*kiaŋ*; *\*kaik* (& *\*kaak*?) “transverse, to branch, stretch horizontally.”

<sup>60</sup> “Intervocalic \*-p- > -v- is almost universal in Dravidian” (Zvelebil 1990: §1.7.7). Although the -v- here is not intervocalic, other cognates of the word (like Tamil *kaḷuvu* “to wash”; Telugu *kaḍuvu* occur with the intervocalic -v-. PD had no voiced stops so g- was an allophone of k-.

<sup>61</sup> The letter -ḷ- in Dravidian is transliterated as ṇ, ḍ, ḷ, y, and r per Caldwell (1875: 59). See n. 40



#### 4.9 *Theragāthā* 32, *Suppiyatthera*

##### Oral transmission confusion

ad hoc glosses on words not understood  
confusion of roots *ni + me*, *ni + mā*, *ni + mi*

##### Diachronic language change

simplification of OI *dhātus* and inflections.

##### Bilingual speakers

lack of knowledge of Pāli

##### Orthographic variation

various spelling “mistakes”

*Ajaraṃ jīramānena tappamānena nibbutiṃ  
nimmissaṃ* (vars. *nimmisaṃ*, *nimiyaṃ*, *nimissaṃ*, *nirāmisāṃ*  
“corrected to *nimissaṃ* per eds.”, *nimineyyaṃ*, *nibbana-mīyanti*,  
*nimiyan ti*, *namiya*, *niyaṃ*)<sup>62</sup> *paramaṃ santiṃ yoga-kkhemaṃ  
anuttaraṃ ti*.

“I shall exchange the ageing for agelessness, the burning for quenching, for the highest peace, for unsurpassed rest-from-exertion” (Norman, 5).

There are five different forms here: *nimiyaṃ*, *nimmissaṃ*, *nimissaṃ*, *nirāmisāṃ* and *nimineyyaṃ*, and three more in the commentary *nibbana-mīyanti*, *namiya*, *niyaṃ*.

Burmese has *nimiyaṃ* in the *mūla* while the PTS has *nimmissaṃ* (Sinhalese, var. *nimissaṃ*) < *ni + me*, “to exchange”; future is *ni + meṣyāmi* (OI) = *ni-messāmi* in Pāli (“I will exchange”), with an alternate future *ni-missāmi* or *ni-missaṃ*. The PTS takes *nimissāmi* as the future of the verb *mināti* (“to measure”; *nimināti* “to exchange, to barter; to change”), which is not an extant form in OI, where the form would be *nimāti* as a class 2 or *nimimīte* as a class 3 or *nimāyate* as a class 4 verb based on the root *mā* or *niminoti*, *niminate* based on the root *mi*. It derives *nimissāmi* from the roots *mā* and *mi*, a cross between the two; there is no form *ni-miṣyati* in OI, the future of *mā* would be *nimāṣyati* and the future of *mi* is

<sup>62</sup> Burmese: *nimiyaṃ* (Be), *niyaṃ*. Sinhalese: *nirāmisāṃ* > *nimissaṃ*, *nimmissaṃ* (Ce), *nibbana-mīyanti*, *nimiyan ti*, *namiya*, *niyaṃ* (comm.). Thai: *nirāmisāṃ* (Se). Unknown *nimineyyaṃ* (? var. in Be).

*nimeṣyati. ni-missati* is an artificial, simplified form to match the root vowel. The form with the double *mm-* is probably just a mistake (see below).

The Burmese commentary takes *nimiyaṃ* as an optative glossing it with *parivatteyyaṃ cetāpeyyaṃ* (both meaning “to exchange” Th-a 1, 98<sup>13-14</sup>). The ending *-yaṃ* (OI *-yām*) is first person thematic ending; the “correct” OI form would be *nimayeyam/nimayeya* for the optative of *mi* (class 1), or *nimāyām* for the optative of *mā* (as a 2<sup>nd</sup> class verb; Pāli *nimineyyaṃ* or *nimine*).

The Thai *nirāmisam* is viewed as a mistake by the PTS editors, who correct it to *nimissam*. It is actually an adjective form with the meaning *nir-āmisā* (“not worldly”), to which a first person ending has been added. One might consider it a regular form of a denominative *nirāmisaya*, 3<sup>rd</sup> person *nirāmisayati*, 1<sup>st</sup> person *nirāmisayāmi/nirāmisayaṃ* with *-aya-* > *-ā-* (von Hinüber 2001: §142) > *-ā-* before *niggahīta* > *nirāmisam*. “I spiritualize old age into agelessness, asceticism into extinguishment ...” (lit.: “I spiritualize agelessness by means of old age ...”). Nevertheless it is a hapax legomenon in the canon (as a verb form) which suggests it is a mistake.

*nimineyyaṃ* is the first person optative of the verb *nimināti*. This is formed as if the stem were *min-* with regular optative endings. The OI endings are as above.

There are three roots here: *ni + me* “to exchange”; *ni + mā* “to measure, adjust”; *ni + mi* “to perceive, notice, understand”; *mi* also has the meaning of “measure” which is presumably how it was combined with *mā* to form *nimināti* (as above). It shows several simplifications typical of Pāli’s treatment of the complex OI verbal forms: 1) the future form *nimessaṃ* regularized to *nimissam* so the *-i-* matches the root form *mināti*; 2) same with *nimiyaṃ* where the form *nimāyām* in OI is changed to *nimiyaṃ* (with *-ā-* > *-i-* to match the root; the *-ā-* in the final syllable is automatically shortened before a *niggahīta*; 3) the formation of the optative *nimineyyaṃ* is as if the root were *min-*. Altogether we have two optatives (*nimiyaṃ, nimineyyaṃ*) a future *nimissaṃ* and an indicative *nirāmisam*. If we take future variant *nimmissam* as more than just a mistake, then it is derived from the verb *nir + mā* which means “to build, to construct, create” and whose OI future is *nirmāṣyati* or *nimmāssati* (*nimmissāti* with change of *-ā-* > *-i-* consistent with the above), 1<sup>st</sup> person *nimmissam*, “I will create agelessness with old age, extinguishment with asceticism ...” or “I will measure agelessness with old age ...” The meaning holds good, but “exchange” is still the better reading and so confirmed by the commentary. It was probably just a mistake for *nimissaṃ* in the oral transmission where the scribe imagined he/she heard a closed first syllable (*nim-missaṃ*) rather than an open one *ni-missaṃ*).

For the commentarial variations: *nibbana-mīyanti*. The verb is 3<sup>rd</sup> person plural, passive of *mināti*, “to measure” so would mean that agelessness (*ajaram*) and extinction (*nibbutim*) are measured by means of ageing and burning which is not particularly coherent, especially since *ajaram* and *nibbutim* are in the accusative here and should be in the nominative for this kind of structure. The word *nibbana* is an adjective meaning “desireless” without a modificand, unless *nibbanā* was the word intended (“those who are without desire”), so the meaning might be: “Those without desire are measured by ageing (turning into) agelessness, by burning (turning into) extinction”; I suppose that is possible, but not very likely. The word *niyaṃ* (Pāli *niya* = “one’s own”) is presumably a mistake for *nimiyaṃ* with *-im-* > *Ø*; if an adjective it would modify *paramaṃ santiṃ* (“one’s own highest peace”), but then the *gāthā* has no verb and the metre is short. The word *namiya* looks like yet another verb form (< *nam*, “to bow down”), employing the OI optative form *nameya* (“with ageing, I should turn towards agelessness ...”), misspelled as Pāli *namiya* (which would ordinarily be *name* or *nameyyaṃ* in Pāli).

How did this happen where one verb has diversified into so many different forms? First, one must assume that it took place in the oral tradition; if the words had been written down, presumably tradents would have been able to look up the words and clarify any confusions. Many or all of the variations were probably ad hoc glosses on an original word which was not understood; the verb *nimināti* is very rare in the canon (PTS lists four instances in the *Jātakas* and one in *Mil*; Cone lists the same and a few from *Cariyāpiṭaka*), so presumably it was not understood—especially if the tradent’s first language was not IA but Dravidian or Munda—and attempts were made to understand what was meant. Its occurrence here in Th was probably its first occurrence anywhere in IA as the OI form *ni + me* (“to exchange”) first occurs in the MBh, which likely postdates Th. The verb *ni + mā* (“to measure”) does occur in the RV but the meaning is off. The word *nirāmisā* probably got included because of its phonological similarity and because it was fairly common in the canon and had the right sense for the context (“spiritual”). The commentarial variations look like further attempts to interpret the meaning, as noted above. The exact sequence is impossible to reconstruct but it does provide some insight into the vagaries of the oral transmission of the *Buddhadhamma*.

END OF PART I.

Part II will be forthcoming in JOCBS 2025.

**ACKNOWLEDGEMENTS**

Thanks to editors Aleix Ruiz-Falqués and Alex Wynne for their important suggestions and comments on this paper which have helped to improve it. Thanks also to Stefan Karpik for reviewing the article and for his suggestions, even though we still disagree on a number of points.

**ABBREVIATIONS**

Aiñk	<i>Aiñkurunūru</i>
AMg	Ardha-Māgadhi
AA	Austro-Asiatic (of which Munda is a sub-branch)
AV	<i>Atharvaveda</i>
Be	<i>Chaṭṭha Saṅgāyana</i> Burmese recension
BHSD/G	<i>Buddhist Hybrid Sanskrit Dictionary/Grammar</i> (Edgerton 1953/1998)
BJT	<i>Buddha Jayanti Tipiṭaka</i>
comm.	commentary
C	consonant
Ce	<i>Buddha Jayanti</i> Sinhalese recension
CP	<i>Collected Papers</i> (Norman)
DED	<i>Dravidian Etymological Dictionary</i> (Burrow and Emeneau 1984)
DPD	Digital Pāli Dictionary <a href="https://github.com/digitalpalidictionary/digitalpalidictionary.github.io">Home - Digital Pāli Dictionary (digitalpalidictionary.github.io)</a>
Geiger	1916/2005
IA	Indo-Aryan
ka.	Cambodian var.
MBh	<i>Mahābhārata</i> (4th century BCE–4th century CE)
M1	Mayrhofer 1956–76
M2	Mayrhofer 1992–96
Mc	<i>Mahācuḷā</i> Thai recension (1963)
MI	Middle Indic
MED	<i>Munda Etymological Dictionary</i> (Stampe, D.)

MKED	<i>Mon-Khmer Etymological Dictionary</i> (Stampe, D.)
MW	Monier-Williams, Sanskrit English Dictionary
ms(s)	Manuscript(s)
N	nasal
non-IA	non-Indo-Aryan
Norman	Norman 1969/95
OI	Old Indic (Vedic)
OT	Old Tamil
P	plosive
(p.) p.p.	(passive) past participle
PCD	Proto Central Dravidian
PD	Proto Dravidian
Pischel	Pischel 1900/1981
PM	Proto Munda
PMA	<i>Pāli-Maynamā' Abhidhān'</i> (Burmese-Burmese Pāli dictionary)
PND	Proto North Dravidian
PSD	Proto South Dravidian (for PSD1 and PSD2, see Southworth 50)
PTS	Pali Text Society
OT	Old Tamil
RV	Rig Veda
Sadd	<i>Saddanīti</i> (Smith 1928–54/2001)
Se	Thai <i>Syāmaratṭha</i> recension
Sī	Sinhalese recension
Sn	<i>Sutta Nipāta</i> (4 <sup>th</sup> –5 <sup>th</sup> centuries BCE)
Southworth	Southworth 2005
Spk	<i>Samantapāsādikā</i>
ṭ	<i>ṭikā</i>
Th	<i>Theragāthā</i>
Th-a	<i>Theragāthā aṭṭhakathā</i> ( <i>Paramattha-Dīpani</i> )
TL	Tamil Lexicon <a href="http://tamil.uchicago.edu">Tamil Lexicon (uchicago.edu)</a>

Tolk	<i>Tolkāppiyam</i> (mid-1st millennium BCE)
TS	<i>Taittiriya Saṃhitā</i>
var(s).	variation(s), variant reading(s)
v.l(l).	<i>varia lectio</i> , ( <i>variae lectiones</i> ), variant reading(s)
Whitney	Whitney 1879/2000
~	Alongside, corresponding to, allophone, compare to

## WORKS CITED

### A Note on *Theragāthā* recensions

The principal reference is the Oldenberg and Pischel edition of the *Thera-* and *Therī-gāthā*, originally published in 1883, reprinted with additions by Norman and Alsdorf in 2006. Oldenberg and Pischel made use of four different manuscripts of Th and Th-a (*Paramattha-Dīpanī*) in Burmese and Sinhalese (p. xii). Woodward’s edition of the commentary, published in three volumes (1940/1995, 1952/1977 and 1959/2004), has many variants and has also been used; it draws on five Burmese and Sinhalese manuscripts. In this article, Burmese readings are referred to as “Burmese” except for the reading in the *mūla*-text of the *Chaṭṭha Saṅgāyana* recension from 1956 which is labelled “Be.” Thai readings and variants (not found in the above) are from the *Syāmaratṭha* 1927 edition and the reading in the *mūla* is labelled “Se” (available at [https://drive.google.com/drive/folders/1gvEv87JkQijpSmBM1jjpuJ1\\_XIOpJYBV](https://drive.google.com/drive/folders/1gvEv87JkQijpSmBM1jjpuJ1_XIOpJYBV)). I have also consulted the 1963 *Mahācūḷā* edition (Mc) of the Thai canon (which, however, has very few variants not mentioned in the other recensions). Sinhalese readings are drawn from all the above and are so labelled except for the reading from the *mūla* of the Buddha Jayanti edition (1957–1993) which is labelled “Ce”; a roman version of this exists at the Sri Lanka Tripitaka Project ([Sri Lanka Tripitaka Project: Pali Tipitaka Source Texts \(accesstoinight.org\)](http://sri.lanka.tripitaka.org)), which is not a 100% accurate transcription of the Sinhalese original, but in the process of being revised/corrected. The Ce readings in this paper have been taken from a romanized version of the Buddha Jayanti edition (original PDFs available at <http://dr.lib.sjp.ac.lk/handle/123456789/2123>), which were transcribed by Ven. Bodhirasa for the author, to whom I offer my sincere thanks. The commentary has very few other readings not in the other *mūla* texts or commentaries. Khom Cambodian variants where available in the above

are so listed. Variants from other parts of the canon in parallel passages are included where relevant. PTS *mūla* readings are not usually listed, where they are the same as one of the other recensions. Variants are usually not repeated; if they occur in two or more recensions, they are only given in the first one listed, unless they occur as the *mūla* text. “Comm.” usually refers to the PTS edition of Th-a, which has the most variants; occasionally the Burmese, Thai and Sinhalese commentaries have variants not covered by the PTS and they are included.

### Primary Sources

Oldenberg, H. and R. Pischel. 1883/2006. *The Thera- and Therī-gāthā (Stanzas ascribed to Elders of the Buddhist order of Recluses)*. Second edition with Appendices by K. R. Norman and L. Alsdorf, 2006. Lancaster: Pāli Text Society.

Woodward, F. L. 1940/1995–1959/2004. *Paramattha-Dīpanī Theragāthā-Aṭṭhakathā. The Commentary of Dhammapālācariya*. Three volumes. Oxford: Pāli Text Society.

Mahācūḷā edition of the Thai canon 1963. No editor given. Bangkok: Mahācūḷālaṅk araṇarājavidyālayena.

Buddha Jayanti edition of the Sinhalese canon (1957–1993), available in Sinhalese at <http://dr.lib.sjp.ac.lk/handle/123456789/2123>; Volume 22 contains the *Theragāthā*. The Sri Lanka Tripitaka Project. ([Sri Lanka Tripitaka Project: Pali Tipitaka Source Texts \(accessstoinight.org\)](http://sri.lanka.tripitaka.org)), is a romanization of the original. This paper used a romanized version of the Sinhalese prepared by Ven. Bodhirasa for the author; it differs somewhat from the Sri Lanka Tripitaka Project which has some typos.

Syāmaratṭha edition of the Thai canon 1927. Volume 26 of the *mūla* contains the *Theragāthā*. Volumes 32 and 33 of the commentary contain the *Paramattha-Dīpanī*. Edited by Somdej Phra Ariyavaṃsāgatañña (Pussadevamahāthera). Bangkok: Mahāmakut Rajavidyalaya.

## Secondary Sources

- Anālayo, Bh. 2009. "The Vicissitudes of Memory and Early Buddhist Oral Transmission." *Canadian Journal of Buddhist Studies*. Vol. 5: 5–20.
- . 2012. "The Historical Value of the Pāli Discourses." *Indo-Iranian Journal* 55: 223–53.
- . 2022. *Early Buddhist Oral Tradition, Textual Formation and Transmission*. Somerville: Wisdom Publications.
- Anon. 1967–1970. *Index des mots de la littérature tamoule ancienne*. Vols. 1, 2, 3. Pondicherry: Institut Français d'Indologie.
- Bechert, H. 1968. "Some Remarks on the Kaṭhina Rite." *Journal of the Bihar Research Society* 54: 319–29.
- Bloch, J. 1950. *Les Inscriptions d'Asoka*. Paris: Société d'Édition "Les Belles Lettres."
- Bloomfield, M. and F. Edgerton. 1932/1979. *Vedic Variants*. Vol. 2, Phonetics. Philadelphia: Linguistic Society of America.
- Bodding, P. O. 1929–36. *A Santal Dictionary*. Oslo: I kommisjon hos J. Dybwad.
- Bodhi, Bh. 2000. *The Connected Discourses of the Buddha, A Translation of the Saṃyutta Nikāya*. Boston: Wisdom Publications.
- . 2012. *The Numerical Discourses of the Buddha. A Translation of the Aṅguttara Nikāya*. Boston: Wisdom Publications.
- Böhtlingk, O. and R. Roth. 1855–75/1990. *Sanskrit Wörterbuch*. New Delhi: Motilal Banarsidass Publishers Pvt. Ltd.
- Brough, J. 1962. *The Gāndhārī Dharmapada*. Delhi: Motilal Banarsidass Publishers.
- Bühler, G. 1894. "Votive Inscriptions from the Sanchi Stūpa." In Jas Burgess, ed., *Epigraphia Indica*. Vol. 2: 87–115. Delhi: Motilal Banarsidass.
- Burrow, T. 1943. "Dravidian Studies III." *Bulletin of the School of Oriental and African Studies*. 11(1): 122–39.
- . 1946. "Loan-words in Sanskrit." *Transactions of the Philological Society*. Vol. 46: 1–30.
- . 1948. "Dravidian Studies VII. Further Dravidian words in Sanskrit." *Bulletin of the School of Oriental and African Studies*, University of London. Vol. 12, 2: 365–96.
- Burrow, T. and M. B. Emeneau. 1984. *A Dravidian Etymological Dictionary*. Second Edition. Oxford: Clarendon Press.



- Caldwell, R. 1875. *A Comparative Grammar of the Dravidian or South-Indian Family of Languages*. London: Trübner & Co., Ludgate Hill.
- Campbell, L. 2004. *Historical Linguistics*. Cambridge: The MIT Press.
- Clark, C. 2015. *A Study Of The Apadāna, including an Edition and Annotated Translation of the Second, Third and Fourth Chapters*. PhD thesis, University of Sydney. <https://ses.library.usyd.edu.au/handle/2123/13438>
- Cone, M. 2001–2020. *A Dictionary of Pāli*. Bristol: Pali Text Society.
- Deshpande, M. M. 1979. “Genesis of Ṛgvedic Retroflexion: A Historical and Sociolinguistic Investigation”. In M. M. Deshpande and P. E. Hook eds., *Aryan and non-Aryan in India*: 235–315. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.
- Dhammika, S. 2015. *Nature and the Environment in Early Buddhism*. Singapore: Buddha Dhamma Mandala Society.
- Dharmakāya Institute 2013. *Dīghanikāya. Volume 1. Sīlakkhandhavagga*. Thailand: Dhammachai Institute, Dharmakāya Foundation.
- Drewes, D. 2017. “The Idea of the Historical Buddha.” *Journal of the International Association of Buddhist Studies* 40: 1–25.
- . 2022. “The Buddha and the Buddhism that Never Was.” Paper presented at the XIXth Congress of the International Association of Buddhist Studies, Seoul, August 15–19.
- Dutt, N. 1947. *Gilgit Manuscripts*. Vol. 3, Part 1. Calcutta: Calcutta Oriental Press.
- Edgerton, F. 1930. “Dialectic Phonetics in the Veda: Evidence from the Vedic Variants.” in *Studies in Honor of Hermann Collitz*. Baltimore: John Hopkins Press.
- . 1953/1998. *Buddhist Hybrid Sanskrit Grammar and Dictionary*. Delhi: Motilal Banarsidass Publishers.
- Emeneau, M. B. 1954. “Linguistic Prehistory of India.” *Proceedings of the American Philosophical Society*. Vol.98 No. 4: 282–92.
- . 1956. “India as a Linguistic Area.” *Language* 32: 3–16. Also available in Emeneau 1980: 126–166.
- . 1969. “Onomatopoeics in the Indian Linguistic Area.” *Language*. Vol. 45 No. 2, Part 1: 274–99.
- . 1970. *Dravidian Comparative Phonology, A Sketch*. Tamil Nadu: Annamalai University.

- . 1974. The Indian linguistic area revisited. *International Journal of Dravidian Linguistics*. Vol. 3: 92–134.
- . 1980. *Language and Linguistic Area*. Stanford: Stanford University Press. A collection of essays edited by A. S. Dil published between 1954–1974.
- Endo, T. 2013. *Studies in Pāli Commentarial Literature. Sources, Controversies and Insights*. Hong Kong: Centre of Buddhist Studies The University of Hong Kong.
- Fahs, A. 1985. *Grammatik des Pāli*. Leipzig: VEB Verlag Enzyklopädie.
- Fussman, G. 1989. “Gāndhārī écrite, Gāndhārī parlée.” In C. Caillat ed., *Dialectes dans les littératures indo-aryennes*: 433–501. Paris: Collège de France, Institut de Civilisation indienne.
- Geiger, W. 1916/2004. *Pāli Literatur und Sprache*. Strassburg: Verlag from Karl. J. Trübner. English translation in Geiger, W. 2004. *Pāli Literature and Language*, translated by Batakrishna Ghosh. New Delhi: Munshiram Manoharlal Publishers. Originally published in 1943.
- . 1916/2005. A Pāli Grammar, translated into English by Batakrishna Ghosh, revised and edited by K. R. Norman. Oxford: The Pali Text Society.
- Ghosh, A. 2008. “Santali.” In G. D. S. Anderson, ed., *The Munda Languages*: 11–98. New York: Routledge Taylor & Francis Group.
- Hallisey, C. 2024. “Seeing Shadows in the Shade. The Narrative Quality of Textual Variants and the History of Buddhism in South and South-east Asia.” In Glenn W. Most, ed., *Variants and Variants in Classical Textual Cultures*: 85–108. Boston/Berlin: Walter de Gruyter.
- Hare, E. M. 1935/2006. *The Book of the Gradual Sayings (Aṅguttara-Nikāya) or More-Numbered Suttas*, Volume 4. Delhi: Motilal Banarsidass Publishers.
- Hinüber, O. v. 1982. “Pāli as an Artificial Language.” *Indologica Taurinensia* 10: 133–140. Also published in Harry Falk und Walter Slaje eds., *Kleine Schriften, Teil 1* (2009): 451–58. Wiesbaden, Harrassowitz Verlag.
- . 1983. “The Oldest Literary Language of Buddhism.” *Saeculum* 34 (1983): 1–9. Also published in *Selected Papers on Pāli Studies* (2005): 177–94. Oxford: The Pali Text Society.
- . 1991. *The Oldest Pāli Manuscript, Four Folios of the Vinaya-Piṭaka from the National Archives, Kathmandu*. Mainz, Stuttgart: Akademie der Wissenschaften und der Literatur; Franz Steiner Verlag.
- . 1996/2008. *A Handbook of Pāli Literature*. New Delhi: Munshiram Manoharlal Publishers Pvt. Ltd.

- . 1996. “Linguistic Considerations on the Date of the Buddha.” In H. Bechert ed., *When did the Buddha live? The Controversy on the Dating of the Historical Buddha: 185–94*. Delhi: Sri Satguru Publications.
- . 2001. *Das Ältere Mittelindisch im Überblick*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- . 2006. “Hoary past and hazy memory. On the History of early Buddhist texts.” *Journal of the International Association of Buddhist Studies*. 29(2): 193–210.
- . 2015. “Languages: Indic.” In J. Silk, R. Bowring and V. Eltschinger eds., *Brill’s Encyclopedia of Buddhism Online*. Leiden, The Netherlands; Boston: Koninklijke Brill NV.
- . 2019. “The Buddha as a Historical Person.” *Journal of the International Association of Buddhist Studies*. Vol. 42: 231–64.
- . 2001. *Das Ältere Mittelindisch im Überblick*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Horner, I. B. 1938–66 / 2001–07. *Book of the Discipline*. Six volumes. London: Pali Text Society.
- Hultzsch, E. 1924/1969. *Inscriptions of Asoka, New Edition*. Delhi: Indological Book House.
- Hu-von Hinüber, H. 1994. *Das Poṣadhavastu Vorschriften für die buddhistische Beichtfeier im Vinaya der Mūlasarvāstivādins*. Hamburg: Dr. Inge Wezler Verlag für Orientalistische Fachpublikationen.
- Jones, J. J. 1949–1956. *Mahāvastu*. Vols. 1, 2, 3. London: Luzac and Co.
- Keown, D. 2003/2004. *Oxford Dictionary of Buddhism*. Oxford/New York: Oxford University Press.
- Krishnamurti, Bh. 2003. *The Dravidian Languages*. Cambridge: Cambridge University Press.
- Kuiper, F. B. J. 1939. “Indoiranica.” *Acta Orientalia Ediderunt Societates Orientales Batava Danica Norvegica*. Vol. 17: 17–64.
- . 1948a. *Proto-Munda Words in Sanskrit*. Amsterdam: N.V. Noord-Hollandsche Uitgevers Maatschappij.
- . 1948b. “Munda and Indonesian.” *Orientalia Neerlandica, A Volume of Oriental Studies*. 372–401. Leiden: A. W. Sijthoff’s Uitgeversmaatschappij N. V.
- . 1954. “Two Rigvedic Loan-words.” In *Sprachgeschichte und Wortbedeutung; Festschrift Albert Debrunner*. 241–50. Bern: Francke.

- Kusalasaya, K. 1965/ 2005. *Buddhism in Thailand Its Past and Its Present*. Kandy, Sri Lanka: Buddhist Publication Society.
- Lamotte, É. 1958/1988. *History of Indian Buddhism: From the origins to the Śāka Era*. Sara Webb-Boin, trans. Louvain-laNeuve: Université Catholique de Louvain Institut Orientaliste.
- Levman, B. G. 2010. “Asokan phonology and the Language of the Earliest Buddhist Tradition.” *Canadian Journal of Buddhist Studies*. Vol. 6: 59–88.
- . 2011. “The *muṇḍa/muṇḍaka* crux.” What does the word mean? *Canadian Journal of Buddhist Studies*. Vol. 7: 45–78. Also in Levman 2021a: 210–35.
- . 2013. “Cultural Remnants of the Indigenous Peoples in the Buddhist Scriptures.” *Buddhist Studies Review*. Vol. 30, No.2: 145–80.
- . 2014. “Linguistic Ambiguities, the Transmissional Process, and the Earliest Recoverable Language of Buddhism.” PhD diss., University of Toronto, Department for the Study of Religion. Available from Dissertations & Theses @ University of Toronto; ProQuest Dissertations & Theses Global. (1566171666).
- . 2016a. “The Language of Early Buddhism.” *Journal of South Asian Languages and Linguistics*. Vol. 3, No. 1: 1–41.
- . 2016b. “Towards a Critical Edition of the *Tipiṭaka*.” *Ñāṇasamvara Journal of Buddhist Studies*. Vol. 1: 83–120.
- . 2017. “Language Theory, Phonology, and Etymology in Buddhism and their relationship to Brahmanism,” *Buddhist Studies Review*. Vol. 34, No. 1: 25–51.
- . 2019a. “The language the Buddha spoke.” *Journal of the Oxford Centre for Buddhist Studies*. Vol. 17: 63–105.
- . 2019b. “The Historical Buddha: Response to Drewes.” *Canadian Journal of Buddhist Studies*. Vol. 14: 25–56. [https://thecjbs.org/wp-content/uploads/2019/07/Levman\\_The-Buddha-as-An-Historical-Figure.pdf](https://thecjbs.org/wp-content/uploads/2019/07/Levman_The-Buddha-as-An-Historical-Figure.pdf)
- . 2020. “Sanskritization in Pāli.” *Journal of South Asian Languages and Linguistics*. Vol. 7, No. 1: 105–49. [doi.org/10.1515/jsall-2021-2030](https://doi.org/10.1515/jsall-2021-2030).
- . 2021a: *Pāli and Buddhism Language and Lineage*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- . 2021b. “A Dravidian Poem Translated into Pāli? Apadāna-aṭṭhakathā/ Visuddhajanavilāsini (53413–53728, vv. 12–48).” *Buddhist Studies Review*. Vol. 38, No. 2: 169–223.

- . 2022. “The influence of Proto-Dravidian on Indo-Aryan Phonology, Morphology and Syntax.” Part 1. *International Journal of Dravidian Linguistics*. Vol. 51, No. 2: 11–65. Part 2 in *IJDL*, Vol. 52, No. 1.
- . 2023a. “Dravidian Buddhism.” *Buddhist Studies Review*. Vol. 38, No. 1: 59–114.
- . 2023b. “Descent with Variation.” *Journal of the Oxford Centre for Buddhist Studies*. Vol. 23: 1–40.
- . 2025a. *Dīgha Nikāya, a New Translation*. Alexander Wynne, ed. Bangkok: Royal Thai Translation Project. Forthcoming.
- . 2025b. “The Thai Pāli Tradition, A Prolegomenon, Part I.” *Ñāṇasaṃvara Centre for Buddhist Studies*. Forthcoming.
- . 2025c. “The Thai Pāli Tradition, A Prolegomenon, Part II.” *Ñāṇasaṃvara Centre for Buddhist Studies*. Forthcoming.
- Lüders, H. 1899. “Bemerkungen zu dem Kharoṣṭhī Manuscript des Dhammapada (MS. Dutreuil de Rhins).” *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, Philologisch-Historische Klasse Berlin*: 474–94. Also in *Kleine Schriften* 151–171.
- . 1954. *Beobachtungen über die Sprache des buddhistischen Urkanons*. Berlin: Akademie Verlag.
- . 1973. *Kleine Schriften*. Oskar von Hinüber, ed., Wiesbaden: Franz Steiner Verlag GMBH.
- Mallik, M. 1970. “Sanskritism in Pāli.” *Vishveshvaranand Indological Journal*: 103–08 (no volume given). Hoshiarpur: Vishveshvaranand Institute of Sanskrit and Indological Studies, Panjab University.
- Masica, C. P. 1976. *Defining a Linguistic Area*. Chicago: The University of Chicago Press.
- Mayrhofer, M. 1956–76. *Kurzfassstes etymologisches Wörterbuch des Altindischen. A Concise Etymological Sanskrit Dictionary*. Three volumes. Heidelberg: Carl Winter, Universitätsverlag.
- . 1992–96. *Etymologisches Wörterbuch des Altindischen*. Two volumes. Heidelberg: Carl Winter, Universitätsverlag.
- Mehendale, M. A. 1968. *Some aspects of Indo-Aryan linguistics*. Bombay: University of Bombay.
- Monier-Williams, Sir. M. 1899/1986. *A Sanskrit English Dictionary*. New Delhi: Motilal Banarsidass.

- Norman, K. R. 1965. "Middle Indo-Aryan Studies V." *Journal of the Oriental Institute (Baroda)*. Vol. 15: 113–17.
- . 1969/1995. *The Elders' Verses I Theragāthā*. Oxford: Pali Text Society.
- . 1980. "Four Etymologies from the Sabhiya-sutta." *Buddhist Studies in Honour of Walpola Rahula*, Somaratna Balasooriya (et al.) eds., 173–184. London: Gordon Fraser. Reprinted in Norman, K. R., *Collected Papers 2* (1991): 148–161. Oxford: Pali Text Society.
- . 1983. *Pāli Literature*. Wiesbaden: Otto Harrassowitz.
- . 1987. "An epithet of Nibbāna." In *Śramaṇa Vidyā: Studies in Buddhism (Prof. Jagannath Upadhyaya Commemoration Volume)*: 22–30. Also published in *Collected Papers 3* (1992): 183–89. Oxford: Pali Text Society.
- . 1989. The Pāli language and scriptures. In Tadeusz Skorupski ed., *The Buddhist Heritage*, 29–53. Tring, UK: Institute of Buddhist Studies. Reprinted in Norman, K. R. 1993. *Collected Papers 4*: 92–123. Oxford: The Pāli Text Society.
- . 1990. "Pāli philology and the Study of Buddhism." In T. Skorupski ed., *The Buddhist Forum*. Vol. 1. London: School of Oriental and African Studies: 31–39. Also published in *Collected Papers 4* (1993): 80–91. Oxford: Pali Text Society.
- . 1992/2006. *The Group of Discourses*. Lancaster: Pali Text Society.
- . 1997/2004. *The Word of the Doctrine (Dhammapada)*. Oxford: Pali Text Society.
- . 1997/2006. *A Philological Approach to Buddhism. The Bukkyo Dendo Kyokai Lectures 1994*. Lancaster: Pali Text Society.
- . 2002. "Pāli and the languages of early Buddhism." In Nicholas Sims-Williams ed., *Indo-Iranian Languages and Peoples*, Proceedings of the British Academy: 135–50. Oxford: Oxford University Press. Also published in *Collected Papers 8* (2007): 96–120. Oxford: Pali Text Society.
- Oldenberg, H. and R. Pischel 1883/2006. *The Thera- and Therī-gāthā (Stanzas Ascribed to Elders of the Buddhist Order of Recluses)*. Lancaster: Pali Text Society.
- Osada, T. 2008. "Mundari." In Gregory D. S. Anderson, ed., *The Munda languages*: 99–164. New York: Routledge Taylor & Francis Group.
- Pāli Myanmā Abhidhān'*. 1964–2014. Myanmar: Ministry of Religious Affairs Department.
- Pande, G. C. 1974. *Studies in the Origins of Buddhism*. Delhi: Motilal Banarsidass.
- Paragaramuthali, 1985–2007. *A Comprehensive Etymological Dictionary of the Tamil Language*. Eight volumes. Tamilnadu: Government of Tamilnadu. <http://www.tamilvu.org/ta/library-ldpam-ldpam00-html-ldpam00hom-244696>.

- Paramasivam, K. 1979. *Effectivity and causativity in Tamil*. Trivandrum: Dravidian Linguistics Association.
- Pischel, R. 1900/1981. *Comparative Grammar of the Prākṛit Languages*, Subhadra Jhā trans., Delhi: Motilal Banarsidass.
- Pokorny, J. 1959–69. *Indogermanisches etymologisches Wörterbuch*. Bern: Francke. Available online at [J. Pokorny's Indo-European Etymological Dictionary](#)
- Przyluski, J. 1929. “Emprunts Anaryens en Indo-Aryen.” *Bulletin de la Société Linguistique, Paris*. Vol. 30: 195–201.
- Rajams, V. S. 1992. *A Reference Grammar of Classical Tamil Poetry*. Philadelphia: Memoirs of the American Philosophical Society, Vol. 199.
- Rhys Davids, C. A. F. 1913. *Psalms of the Early Buddhists II. Psalms of the Brethern*. London: Henry Frowde, Oxford University Press.
- Rhys Davids, T. W. 1877. *Buddhism, being a sketch of the life and teachings of Gautama the Buddha*. London: Society for Promoting Christian Knowledge.
- . 1899/1923. *Dialogues of the Buddha*. London: Humphrey Milford, Oxford University Press.
- . 1911. *Buddhist India*. New York: Putnam.
- Salomon R. 1998. *Indian Epigraphy, A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the Other Indo-Aryan Languages*. New Delhi: Munshiram Manoharlal Publishers Pvt. Ltd.
- . 2000. *A Gāndhāri Version of the Rhinoceros Sūtra*. British Library Kharoṣṭhī Fragment 5B. Seattle and London: University of Washington Press.
- . 2024. “Forgotten Buddhist Literatures as Revealed by Manuscript Discoveries in India and Beyond.” *Philological Encounters* (published online ahead of print 2024). <https://doi.org/10.1163/24519197-bja10054>
- Shorto, H. 2006. *A Mon-Khmer comparative dictionary*. Paul Sidwell ed., Assisting editors, D. Cooper and C. Bauer. Canberra: Pacific Linguistics Research School of Pacific and Asian Studies The Australian National University.
- Sjoberg, A. F. 1992. “The impact of Dravidian on Indo-Aryan: An overview.” In Edgar C. Polomé & Werner Winter eds., *Reconstructing languages and cultures*: 507–29. Berlin: Mouton de Gruyter. Reprinted in Sjoberg, A. F., *Dravidian language and culture*, (2009): 55–80. Srinivasavanam: Dravidian University.
- Smith, H. 1928–54/2001. *Saddanīti, la grammaire palie d’Aggavaṃsa*. Vols. 1–4. Oxford: Pali Text Society.

- . 1952. “Le futur moyen indien et ses rythmes.” *Journal Asiatique*. Vol. 240: 169–83.
- Southworth, F. C. 2005. *Linguistic Archaeology of South Asia*. London and New York: RoutledgeCurzon.
- Stampe, D. *SEalang Munda Etymological Dictionary*. <http://www.sealang.net/munda/dictionary/>
- Stampe, D. *SEalang Mon-Khmer Etymological Dictionary*. <http://sealang.net/monkhmer/dictionary/>
- Study Group on Buddhist Sanskrit Literature. 2006. *Vimalakīrtinirdeśa*, A Sanskrit Edition Based upon the Manuscript Newly Found at the Potala Palace. Tokyo: The Institute for Comprehensive Studies of Buddhism, Taisho University, Taisho University Press.
- Sujato Bh. and Bh. Brahmalī. 2014. *The Authenticity of Early Buddhist Texts*. [https://bodhi-college.org/wp-content/uploads/Sujato-und-Brahmalī\\_2014\\_The-authenticity-of-Early-Buddhist-Texts.pdf](https://bodhi-college.org/wp-content/uploads/Sujato-und-Brahmalī_2014_The-authenticity-of-Early-Buddhist-Texts.pdf)
- Thomas, E. J. 1931. *The Life of Buddha as Legend and History*. London: Routledge & Kegan Paul.
- Tov, E. 1992/2001. *Textual Criticism of the Hebrew Bible*. Minneapolis: Fortress Press.
- Wackernagel, J. 1896–1954. *Altindische Grammatik*, Part I–III. Göttingen: Vandenhoeck und Ruprecht.
- Waldschmidt, E. 1979. “The *Varṇasatam*, An Eulogy of One Hundred Epitheta of Lord Buddha Spoken by the *Gṛhapati Upālī(n)*.” *Nachrichten der Akademie der Wissenschaften in Göttingen I. Philologisch-Historische Klasse*. Jahrgang 1979, Nr. 1: 3–19.
- Walters, J. S. 2018. *Legends of the Buddhist Saints*. *Apadānapāli*. Whitman College: Jonathan S. Walters and Whitman College.
- Warder, A. K. 1963/2001. *Introduction to Pali, Third Edition*. Oxford: Pali Text Society.
- . 1967. *Pali Metre, A Contribution to the History of Indian Literature*. London: Pali Text Society.
- Whitney, W. 1879/2000. *Sanskrit Grammar*. Delhi: Motilal Banarsidass Publishers.
- Witzel, M. 1999. “Substrate Languages in Old Indo-Aryan (Ṛgvedic, Middle and Late Vedic).” *Electronic Journal for Vedic Studies* Vol. 5: 1–67.
- . 2009. “South Asian Agricultural Terms in Old Indo-Aryan.” In Toshiki Osada ed., *Linguistics, Archaeology and Human Past in South Asia*: 79–100. New Delhi: Replika Press Pvt. Ltd.



- Woolner, A. C. 1926–28. “Prakritic and non-Aryan Strata in the Vocabulary of Sanskrit.” In *Sir Asutosh Memorial Volume*: 65–71. Patna: J. N. Samaddar.
- . 1927–32. *Ardha-Māgadhī Koṣa*. London: Probsthain and Co. Four Volumes.
- . 1928/96. *Introduction to Prakrit*. New Delhi: Motilal Banarsidass.
- Wynne, A. 2004. “The Oral Transmission of the Early Buddhist Literature.” *Journal of the International Association of Buddhist Studies*. Vol. 27, No.1: 97–129.
- . 2005. “The Historical Authenticity of Early Buddhist Literature”. In: *Vienna Journal of South Asian Studies*. Vol. 49: 35–70.
- . 2019. “Did the Buddha exist?” *Journal of the Oxford Centre for Buddhist Studies*. Vol. 16: 98–148.
- Zide, N. H. 2008. “Korku.” In Gregory D. S. Anderson, ed., *The Munda Languages*. London: Routledge.
- Zvelebil, K. V. 1990. *Dravidian Linguistics, An Introduction*. Pondicherry: Pondicherry Institute of Linguistics and Culture.