

## The Buddha taught in Pali: A working hypothesis

*Stefan Karpik*<sup>1</sup>

### Abstract

The Theravada tradition claims that the Buddha taught in Pali. This conflicts with most current scholarship. Yet insights from linguistics and close reading of sources suggest that the Theravada account has not been disproved, that it could be correct, and that it even represents a stronger hypothesis than the current consensus. Instead of authorising translation of his teaching into dialects, the Buddha promoted a fixed transmission and the use of standard language. That the Buddha spoke Māgadhī is a late tradition; Tipiṭaka commentaries instead defined *Māgadhābhāsā*, ‘Magadha language’, as *Ariyaka*, ‘Aryan’, the canonical term for the Indo-Aryan language. Pali has the expected features of a natural standard language and can be seen as a precursor of Epigraphic Prakrit. This working hypothesis suggests a bolder stance for Pali studies of claiming that Pali is in all probability the formal language of the Buddha.

**NOTE:** Unless otherwise stated, translations are by this author, and Pali quotations are from the Myanmar Tipiṭaka in the Digital Pali Reader (DPR), available at <https://pali.sirimangalo.org/>.

---

<sup>1</sup>My thanks for helpful suggestions from Devon Pali Study Group members, Geoff Bamford and James Whelan, and for translations of German by Sally Roberts.

## 1. The issue

The Theravada Buddhist tradition asserts that the Buddha taught in the language that we now call Pali (§3.2). The model used here to promote that case is analogous to the situation of modern British English, which includes mutually intelligible dialects, such as Cockney, Yorkshire, Black Country etc., alongside a standard language, BBC English or Queen’s English, which predominates in the south-east, but is spread across England.<sup>2</sup> BBC English is a sociolect, i.e. a language of a social class, in this case of the well-educated, as well as a south-east dialect. Similarly, in 5<sup>th</sup> century BCE Northern India, the main language, *Ariyaka*, ‘Aryan’, had at least three mutually intelligible dialects and one standard language, whose contemporaneous names we do not have. The dialects are now called western, north-western or Gāndhārī and eastern or Māgadhī; *Ariyaka*, which includes these dialects, is now called Indo-Aryan or Middle Indo-Aryan (MIA). The standard language, based on the western dialect and also a sociolect of the well-educated, is controversially called here Pali, in reference to Buddhist texts, and Epigraphic Prakrit, in reference to post-Aśokan inscriptions. In addition, there were two religious sociolects, Sanskrit of Brahmanism and Ardha-Māgadhī of Jainism; 500 years later, Sanskrit started to replace Pali/ Epigraphic Prakrit as the standard language and sociolect of the educated.

The perspective of this paper is that the academic consensus has misidentified the language of the Buddha with Māgadhī. As the Buddha also taught outside Magadha, he is assumed to have spoken other dialects. Opinions vary whether these dialects were close to each other, so Pali is said to have been *translated*,

---

<sup>2</sup>The reader may notice a switch from *dialect* to *language*. There is no precise distinction as linguists and Indologists do not have agreed definitions of the two. Southworth (2005: 35) states: “The distinction between different languages and different dialects cannot be made on the basis of objective linguistic criteria. In real life political, historical, and other factors enter in. Thus, mutually unintelligible spoken varieties of Chinese are treated as dialects of the same language because they share a single writing system and belong to the same political entity and share a cultural tradition, while Hindi and Urdu, which are mutually intelligible in many colloquial forms, are treated as distinct languages using the same criteria (different scripts, different political entities, partly distinct cultural traditions).”

However, a line can be drawn between *accent* and *dialect*. Beal (2010: 25) states “... typologies of English dialects have tended to be constructed according to phonological criteria; strictly speaking, they are typologies of accents rather than dialects.” She then goes on to describe dialects by variations of morphology, syntax and lexis. Here variation in phonology only is called *accent*, *dialect* is always geographical and *sociolect* is non-geographical.

*transposed*, or *transformed* from these dialects. The result is that Pali is generally claimed to be an artificial, ecclesiastical language, instead of the natural standard language advocated here. The consensus is called here the ‘Multiple Oral Transmission Theory’ (MOTT) because the original teachings were allegedly in several dialects. This paper dissents and advocates a Single Oral Transmission Theory (SOTT) of the Buddha teaching in Pali.

This dissent joins a trend of renewed scrutiny of the MOTT: Ruegg (2000), Levman (2008/9) and Gombrich (2009: 146-8) have offered more plausible interpretations to the passage alleged to authorise local dialects (§4); Cousins (2013: 100-101) has claimed there were not enough monastics to maintain multiple oral canons in different dialects; Gombrich (2018: 69ff) has argued that the Buddha may have preached in Pali.

## **2. The implausibility of oral translation**

The MOTT implies oral translation (leaving aside complete fabrication). There is no record of the original transmissions allegedly translated into Pali. There is, however, evidence of an oral transmission of Pali in Sri Lanka in the 1<sup>st</sup> century BCE (Norman 1983: 10-11). By rejecting oral translation, we will be left with no realistic alternative to identifying Pali as the Buddha’s language.

### **2.1 The impracticality of translation**

Norman (1983: 3ff) describes the position that I reject:

Although there is some doubt about the interpretation of the phrase the Buddha used when asked if it was permissible to translate his sermons [*sakāya niruttiyā*, see §4], it is generally agreed that he did not preach in Sanskrit, but employed the dialect or language of the area where he was preaching. We must assume that his sermons and utterances were remembered by his followers and his audiences as they heard them. In the course of time, during his lifetime and after his death, collections must have been made of his words and translations or redactions of these must have been made as the need arose, either because the collections were being taken into an area where a different dialect or language was spoken, or because as time went by his words became less intelligible as their language became more archaic. As Buddhism became established in various parts of North India, there must have

been an attempt made to render all the holdings of any particular vihāra, which were probably still in various dialects as they had been remembered, roughly homogeneous in language, although we must bear in mind the fact that, as the dialects of North India had probably not diverged greatly from each other in the fourth and third centuries BC, absolute perfection of “translation” was not essential. The anomalous forms in Pāli ... probably represent the remnants of recensions in other dialects, which had not been completely translated.

Why, if dialects had not diverged greatly, would the Buddha have varied his own dialect from place to place? I know of no English speaker who deliberately in all seriousness switches to a local accent when travelling in foreign English-speaking countries. Switching accent would distract from the intended message by producing a comical or satirical effect both now, in English, and then, in Indo-Aryan.<sup>3</sup> Furthermore, the sheer volume of Early Buddhist texts (EBTs), amounting to over 5000 pages<sup>4</sup>, would preclude oral translation.<sup>5</sup> Cousins (2013: 100-1,107) rejected translation when arguing there were not enough monastics to maintain multiple oral canons. Creating a sufficiently fixed oral translation for group recitation would be an additional difficulty as the translator and reciter would need to agree and remember the precise wording and revisions. An oral translation of this scale and of this accuracy would be unparalleled in human history, let alone Indian history, yet

---

<sup>3</sup>Bechert (1980: 14) thought imitation of another dialect would sound ridiculous only if there were a standard language; he denied this was the case because of the language variation in Aśokan inscriptions. Although there is an argument for a standard language (§5), that is irrelevant: comedians entertain with international imitations of accents from Britain, USA, Australia etc. although there is no standard language common to all English-speaking nations.

<sup>4</sup>The first four Nikāyas amount to 5,512 pages in the PTS editions as follows: D I (252), D II (357), D III (292), M I (524), M II (266), M III (302), S I (240), S II (285), S III (278), S IV (403), S V (478), A I (299), A II (257), A III (452), A IV (466), A V (361). Sujato & Brahmali (2015: 9-10) include the above in their definition of EBTs as well as the *pāṭimokkha* (55) and parts of the following: Dhṃ (60), It (124), Sn (223), Th (115), Thi (174), Ud (94). Thus, 5,000 pages for the EBTs is a conservative figure.

<sup>5</sup>Norman in the above quotation put *translate* in inverted commas. In a later work, Norman (1989: 374fn) suggests *transform* or *transpose* might be more appropriate than *translate* and refers to simple sound, morphological or lexical changes without offering examples. However, it appears that an accumulation of these changes would amount to *translation* in a conventional sense; Norman is therefore trying to have it both ways, both *translate* and *transpose*.

this is the MOTT's implicit claim. Not only is Norman's account inherently implausible, but also there is textual evidence against it.

## 2.2 The ideal of a fixed transmission

In the suttas, monks are required to learn the teachings *byañjanena*, 'to the syllable' or, in modern parlance, to the letter.<sup>6</sup> In response to disputes among Jains over accurate transmission of their texts, the Buddha stressed the importance of precise communal recitation in the *Pāsādika Sutta* (D III.127, DN 29.17). Wynne (2004: 115) translates:

Therefore, Cunda, as regards the teachings I have taught to you through understanding, meeting together again and again, (comparing) meaning with meaning (*atthena attham*), (comparing) letter with letter (*byañjanena byañjanam*), you should recite communally and not argue, so that the holy life will be long lasting and endure long...

To judge the authenticity of a text, the Buddha in the *Mahāparinibbāna Sutta* (D II.124, DN 16.4.8) urges reference to the four Great Authorities (*mahāpadesa*). They are the Buddha, a group of monks led by a senior monk, a settled group of monks or a single senior monk. A teaching is authentic if it matches teachings from the Great Authorities to the word and syllable/letter. Wynne (2004: 103) comments:

...the 'words and letters' (*padavyañjanāni*) of the teaching under consideration were to be 'learnt correctly' (*sādhukam uggahetvā*) before judgement was passed. If attention was to be paid to the words and letters of proposed teachings, it implies that the content of what was known as 'Sutta' was also transmitted by paying a similar attention to its words and letters, i.e. that it was transmitted word for word. The passage therefore shows that the accuracy with which a body of literature called 'Sutta' was meant to be transmitted was very high, down to the letter.

---

<sup>6</sup>It is problematic to translate *vyañjana/byañjana* as 'letter' in an oral context; the Rhys Davids (1910: 134) offer 'syllable'; Woodward (1979: 52) has 'letter', as does Walshe (1987: 111), and MW confirms both as a possible translation. Von Hinüber (1987: 119-120) suggests 'sound' and 'syllable'. However, some accommodate the MOTT, mistakenly in my view, with vague translations that allow the possibility of translation: Walshe (1987: 255, 432) has 'expression'; Ñānamoli & Bodhi (1995: 848) 'phrasing', Bodhi (2012: 150) 'phrasing' or (2012: 538) 'formulation'.

To resolve disputes over the meaning or the syllable/letter of the dhamma (*atthato...vyañjanato*), the Buddha lays down procedures in the *Kinti Sutta* (M II.239, MN 103.5). These consist of approaching the most amenable monk on each side and attempting to reconcile differences in the text and accepting hurt feelings in the process. In these passages (and elsewhere), Wynne (2004: 97) finds evidence for “a relatively fixed oral transmission of early Buddhist literature”. He cited them to disprove improvisation in the Suttas, but they also disprove translation. For, if several dialects were employed from the outset, as Norman suggests, translating them would be altering the syllable/letter. The only way the Buddha’s requirements could have been met was to use the same standard language across different regions from the outset (see §5).

There are eighteen suttas in the first four Nikāyas which affirm the importance of learning teachings to the syllable: (1) D II.124 *Mahāparinibbāna*, (2) D III.127 *Pāsādika*, (3) M I.213 *Mahāgosinga*, (4) M II.239 *Kinti*, (5) S IV.379 *Khemā*, (6) S IV.394 *Moggallāna*, (7) S IV.397 *Vacchagotta*, (8) A I.59 *Adhikaraṇavagga*, (9) A I.69 *Samacittavagga*, (10) A II.168 *Mahāpadesa*, (11) A III.178 *Saddhammasamma* [3], (12) A III.201 *Khippanisanti*, (13) A IV.140 *Vinayadhara* [2], (14) A IV.142 *Vinayadharasobhana* [2], (15) A V.71 *Ubbāhikā*, (16) A V.80 *Kusināra*, (17) A V.201 *Thera* and (18) A V.32 *Saññā*. This weight of evidence lies in plain sight and is tantamount to forbidding translation.<sup>7</sup> Yet none of the works cited in this paper address this evidence against translation. Nor do any address the futility of translation, the next topic.

### 2.3 The mutual intelligibility of Indo-Aryan

There are three sources of evidence that the varieties<sup>8</sup> of *Ariyaka*, or Middle Indo-Aryan (MIA), were mutually comprehensible in the Buddha’s day. Firstly, the Vinaya (Vin III.27-28) prescribes the language to be used when a monk disrobes. Von Hinüber (1977: 239) translates:

“He declares his resignation in Aryan to a non-Aryan and the latter does not understand: his resignation from the community is not valid.” (see §3.2 for Pāli and further analysis)

---

<sup>7</sup>Translation was not explicitly forbidden, but I claim it was never considered as a possibility in the EBTs.

<sup>8</sup>‘Variety’ is a term used in sociolinguistics to denote any language type, without committing to whether it is a language, dialect, accent or sociolect.

Similarly, if the resignation is in a non-Aryan language to an Aryan speaker, the resignation is not valid. The commentary gives Tamil and Telugu as examples of non-Aryan languages. This demonstrates that *Ariyaka*, Aryan, was conceptualised as a single language and that the Vinaya was developed on the basis that the varieties of Indo-Aryan were mutually comprehensible.

The second source is the *Araṇavibhaṅga Sutta*, ‘Exposition of Non-Conflict’ (M III.234-5, MN 139) in which the Buddha gives seven words for ‘bowl’ as an example of a conflict over terminology. This is analogous to some British and American English terminology:

|                     |                 |                  |                        |
|---------------------|-----------------|------------------|------------------------|
| American equivalent | British English | American English | British equivalent     |
| side walk           | <b>pavement</b> | <b>pavement</b>  | road                   |
| underpass           | <b>subway</b>   | <b>subway</b>    | underground<br>(train) |

There are many such examples<sup>9</sup> in the two English varieties, which, remarkably, do not destroy mutual intelligibility. Indeed, conflicts over meaning can only happen among mutually intelligible varieties.

The third source is the Aśokan inscriptions of ca. 250BCE, one and a half centuries after the Buddha’s death. They are found from Afghanistan to the east coast of India and from the Himalayan foothills to Karnataka. Norman (1980: 70) argues:

The way in which the dialects of Middle Indo-Aryan continued to diverge from the norm of Sanskrit after 250 B.C. suggests that before 250 B.C. the dialects were less divergent than Sanskrit. We may, therefore, be certain that in the sixth and fifth centuries B.C. the dialects of Indo-Aryan spoken in North India were morphologically and phonologically closer together than the dialects spoken at the time of Aśoka, and, as is well known, the latter are in themselves not very dissimilar.

---

<sup>9</sup>E.g. (British first) alternative/ alternate; angry/ mad (or pissed); biscuit/ cookie; braces/ suspenders; breastfeed/ nurse; crisps/ chips; handbag/ purse; homely/ homey; jumper/ sweater; parting (in hair)/ part; petrol/ gas; pissed/ drunk; trousers/ pants; waistcoat/ vest. In those examples, one or both of the pairs has a different meaning in the other variety. There are also synonymous pairs, e.g. flat/ apartment; trainers/ sneakers etc., which are much less likely to produce disagreement or confusion.

*A fortiori* the *Ariyaka* of the Buddha's day showed even less difference. (Some disagree with Norman, see discussion at §5.3). Therefore, it is a reasonable hypothesis that the Indo-Aryan dialects were mutually intelligible from the start of the Buddha's ministry, ca. 450BCE, to ca. 50BCE, when Pali texts were written down in Sri Lanka. Certainly, the inscriptions of this time period show vocabulary, syntax and morphology recognisable to a Pali or Sanskrit reader, but with much variation of pronunciation. This is a modest claim as the period of mutual intelligibility can be extended to the 1<sup>st</sup> to 4<sup>th</sup> centuries CE according to Salomon (1998: 85):

...it is questionable whether the MIA dialects of the time were really so different; from the available literary and inscriptional data, it would appear that they were not yet so widely divergent as to present major difficulties of communication.

If the dialects were mutually intelligible before they were written and afterwards, there was no necessity for the heroic effort of oral translation of this vast literature. However, although mutual intelligibility eliminated the need for translation, it created a problem for the ideal of fixed transmission: unintended change.

## §2.4 The reality of 'fixed' transmission

The Buddha would have known of the precise transmission of the Vedic texts, so precise that Witzel (1997: 258) claimed "They must be regarded as *tape recordings*, made during the Vedic period and transmitted orally, and usually without the change of a single word." Some techniques for this precision are given by Scharfe (2002: 248). They include *samhitāpāṭha*, 'continuous' recitation; *padapāṭha*, 'word by word', where compounds are resolved and *kramapāṭha*, 'step' recitation, where, if four words are represented by *abcd*, the recitation order is *ab bc cd*. Scharfe (2002: 241) also attests to accompanying head and hand movements by the students of modern Vedic schools to reinforce their learning.

Yet these methods were not applied to the Buddhist scriptures. Though effective, Vedic techniques involved some reputational damage: Scharfe (2002: 27) translates Mahābhārata V.130,6 (= XII.10,1) "O king, like the mind—dulled by the (constant recital of) Veda sections—of a dim-witted unintelligent Vedic scholar, your (mind) focuses only on morality." Anālayo (2009) points out that



Vedic techniques were taught to boys from age 8 when they did not fully understand what they had learnt; on the other hand Buddhists generally learnt the texts only as fully-ordained adults, aged at least 20; by then they would have understood the texts, which paradoxically would interfere with perfect recall. Moreover, most of the texts were in prose, which would offer no mnemonic help from metre. For Buddhist prose texts, different techniques, including standard pericopes and the waxing syllable principle were used, as detailed by Allon (1997). Thus, Buddhist oral transmission would not have had the same accuracy as Vedic.

English speakers intending to repeat *a huge luxury yacht* accurately might say:

|                   |                      |                          |
|-------------------|----------------------|--------------------------|
| (BBC English)     | a hyuge lukshury yot | /ə hju:ʒ lʌkʃəri jɒt/ or |
| (Estuary English) | a yuge lugjury yoʔ   | /ə ju:ʒ lʌgʒəri jɒʔ/.    |

They may sincerely agree that they have repeated the phrase accurately despite the unintentional differences. Even if the speakers noticed the differences and chose to get as close to standard English as possible, they still might not obtain a perfect repetition. It is difficult to lose one's accent and even a standard accent is not unitary. On a single page of an English dictionary<sup>10</sup>, the following variations are noted: /ɪ'buljənt/, /ɪ'bʊljənt/ for 'ebullient'; /ɛk'sentrik/, /ɪk'sentrik/ for 'eccentric'; /'ɛkdɪsɪs/, /ɛk'daɪsɪs/ for 'ecdysis'; /'ɛʃələn/, /'eɪʃələn/ for 'echelon'; /ɪ'kɪnədə:m/, /ɪ'kɪnədʊdə:m/, /'ɛkɪnədə:m/ and /'ɛkɪnədʊdə:m/ for 'echinoderm'. One page is unrepresentative, but demonstrates some possible phonetic variation even within the narrow dictionary standard. *A fortiori*, in ancient India without dictionaries, there would have been greater flexibility on correct pronunciation.

Similarly, despite the ideal of fixed transmission, phonetic variation inevitably crept into the Buddhist texts. Von Hinüber (1987: 104-9) cites the Vinaya commentary (Vin-a 1399-1400), which lists substitutions which are acceptable for reciters of the Pali suttas: *d* for *t*, *c* for *j*, *y* for *k* and vice versa. For legal sangha proceedings, however, these changes are not allowed. Another class of substitutions is allowed, but disapproved of, in legal proceedings: confusing long and short vowels, inserting or omitting sandhi incorrectly and confusing heavy and light syllables. (It seems the ancients were more flexible on the Law of Morae than grammars suggest.) There is also a category that voids legal proceedings: confusing aspirates and unaspirates, also omitting or

<sup>10</sup> *The Concise Oxford Dictionary, Ninth Edition* 1995 p.428.

inserting nasals incorrectly. Perhaps the allowable changes were the concern of the *Kinti Sutta*'s advice not to get into a dispute: "The meaning agrees, but there is a difference in the syllable/letter. This is a trivial matter, namely, a syllable/letter. Let not the venerable ones get into a dispute on a trivial matter."<sup>11</sup>

We thus have a hybrid of a transmission which was fixed so far as the available techniques allowed, but nonetheless was subject to phonetic variation. In addition, the transmission would have been recited by speakers of several varieties who would accidentally introduce their idiosyncrasies, which could become the norm if they were common enough. Inevitably, involuntarily and largely unconsciously the sounds and morphology of the transmission would shift across geographical areas and across centuries through natural variation and transmission errors.<sup>12</sup> Thus, the variants evident in Pali can all be accounted for by the model of a single, somewhat fluid, oral transmission. As Cousins (2013: 107) states, "It is important to appreciate that a chanted text simply evolves in linguistic form with the passage of time as the language itself evolves. There is no need for any process of translation."

## 2.5 Single transmission compared to translation

The evidence points to translation being impracticable, discouraged and unnecessary. This applies also to variants of *translation*, such as *transposition* and *transformation*, which are all intentional processes in this context. The evidence suggests a single, unintentionally fluid, oral transmission from the Buddha. As we have evidence of an oral transmission in Pali, we have evidence the Buddha taught in Pali.

On the analogy that Shakespeare spoke English although his original pronunciation, vocabulary and grammar differ from modern English, the Buddha's teaching language must have differed somewhat from the Pali now recorded, but it is still Pali. The alleged 'Māgadhisms' in Pali, which are claimed as evidence of translation, are to be expected occasionally in a fluid single transmission across India and will be discussed further in §3. In §5 the case will be strengthened for identifying the single transmission with Pali.

---

<sup>11</sup>M II 240, DPR M 3.1.3, MN 103 *atthato hi kho sameti, byañjanato nānaṃ. appamattakaṃ kho panetaṃ yadidaṃ — byañjanaṃ. māyasmanto appamattake vivādaṃ āpajjithā'ī.*

<sup>12</sup>For example, Gombrich (2018: 80-1) offers what he believes to be different dialect forms, including *tattha* or *tatra*, 'there', and *-e* or *-asmin* or *-amhi* as endings for *-a* declension locative singular endings.

### 3. The Buddha did not teach in Māgadhī

The basis of the MOTT claim of translation of the EBTs is the argument that the Buddha spoke Māgadhī, not Pali. Although I have not seen a detailed exposition, the argument seems to be:

- the Buddha had a special connection with Magadha, where he attained enlightenment (§3.1);
- the Pali commentaries claim the Buddha spoke Māgadhī (§3.2-5);
- the eastern Aśokan inscriptions, which occur throughout the entire area where the Buddha taught, are uniformly in a single dialect which is a form of Māgadhī (§3.6);
- Pali is predominantly a western dialect, and therefore a change, whether intentional or unintentional, from Māgadhī to Pali must have occurred (§3.7);
- the change was incomplete, and some original Māgadhisms are found in Pali (§3.8).

This apparently formidable argument is now examined.

#### 3.1 The Buddha was a Kosalan, not a Magadhan

Buswell & Lopez (2014: 491, Magadha) exemplify a constant and misleading refrain: “Magadha has been described as the birthplace of Buddhism, and its language, the language of the Buddha.”<sup>13</sup> Though Uruvelā in Magadha was where the Buddha became enlightened, the Buddha’s primary connection was with Kosala. Edgerton (1953: 3) points out that the Buddha’s home (Kapilivastu), one of his favourite dwelling places (Śrāvastī), the scene of his first sermon (Sarnath), and the place of his death (Kuśināgarī) were all outside Magadha; in fact, they were all, except for Kuśināgarī, part of the Kosalan state in the time of the Buddha. Arguably, the birthplace of Buddhism, where the Buddha first taught, was Sarnath, in the province of Kāśi, in the state of Kosala. Furthermore,

---

<sup>13</sup>Cf. Thomas (1927: 13), “The home of Buddhism lies in what is now South Behar, west of Bengal and south of the Ganges. This was the country of the Magadhas with the capital at Rājagaha (Rajgir).”; also von Hinüber (1983: 2), “...Theravāda Buddhism, which holds the view that the language of Theravāda is Māgadhī, the language of the province of Magadha, present-day Bihar, where the Buddha lived and taught”.

King Pasenadi of Kosala refers to the Buddha as a Kosalan,<sup>14</sup> and the Buddha himself acknowledges he is of Kosalan ancestry.<sup>15</sup>

In the Buddha's day, King Pasenadi of Kosala and King Ajātasattu of Magadha had each defeated the other in battle (J II.237). Pasenadi finally took Ajātasattu prisoner and then made a peace sealed by giving his daughter in marriage (J II.404). Ajātasattu in turn prepared for attacks from King Pajjota of Avanti (M III.7) and also from the Vajjis (Vin I.228). Kosala was then at least an equal military power to Magadha, and their dialects are likely to have been equally prestigious and acceptable throughout the Ganges basin. The comic effect of imitation of another dialect and the difficulty of losing a childhood accent give reason to suppose that the Buddha did not teach in Māgadhī; in any case, imitation

<sup>14</sup> *bhagavāpi kosalo*, 'the blessed one is a Kosalan'. M II.124, DPR M 2.4.9.374, MN 89.

<sup>15</sup> *kosalesu nīketino*, 'they have a home in Kosala', Sn 422, said by the Buddha of his countrymen, *janapado*. This takes *nīketino* as a nominative plural in apposition to singular *janapado*, influenced by the plural *kosalesu*. However, the commentary takes *nīketino* as genitive singular, meaning 'belonging to one [unspecified] who has a home in Kosala'. It states ***kosalesu nīketinoti bhananto navakarājabhāvaṃ paṭikkhipati. navakarājā hi nīketīti na vuccati***, DPR Sn-a *dutiyo* 3.1. Levman (2013: 157-8) translates: "Saying 'indigenous among the Kosalans' (*kosalesu nīketino*), he rejects the new kingship. A new king is not to be called 'indigenous'." This implies that the Buddha was not Kosalan and his clan, the Sakyans, were unwilling vassals of the Kosalans. In contrast, Bodhi (2017: 867) translates with the opposite meaning "'native to the Kosalans'. Saying this, he rejects its rule by a subordinate ruler; for a subordinate ruler is not said to be native." For *navakarājā*, Levman appears to refer to Pasenadi, King of Kosala, and Bodhi to King Suddhodana, the Buddha's father. Cone (2010: 515) has 'new', 'a novice' for this passage; Bodhi is apparently stretching 'novice' to 'junior' and 'subordinate'.

I prefer Bodhi's 'native to' for the text, but Levman's reading of the commentary. For I think the Pali commentary is also "under the influence of the evolving Buddha legend", as Bodhi (2017: 1468) suggests of the Chinese parallel, which drops any mention of Kosala and turns the Buddha's father into a great king. Similarly, the Pali commentary is reinterpreting the vassal status of Suddhodana to harmonise with legends of his being a great king; this entails implying the father was a Sakyan, not a Kosalan, and could not be both.

The Buddha appeared to have a close connection with the king of Kosala. Walshe (1987: 409) gives the Buddha's view of relations as harmonious (*Aggañña Sutta*, DN 27, D III.83-4): "Now the Sakyans are vassals of the King of Kosala. They offer him humble service and salute him, rise and do him homage and pay him fitting service. And just as the Sakyans offer the King humble service, ... so does the King offer humble service to the Tathāgata, ..." Walshe (1987: 409) adds "... thinking: 'If the ascetic Gotama is well-born, I (Pasenadi) am ill-born ...'", but Anālayo (2016: 35) argues from the Chinese Āgama that Walshe is translating a mistaken reading *nanu ... ti*, 'certainly', and the Burmese edition *na naṃ ... ti*, '... not thinking...', is correct. This is a third instance, in addition to the Pali and Chinese commentaries on Sn 422, of a later source editing away the inconvenient, inferior vassal status of the Buddha's father.

was unnecessary among mutually intelligible varieties of Indo-Aryan.

Nor was Magadha the Buddha's adopted country. Norman (2002: 137) speculates "Most of his [the Buddha's] teaching life was spent in Magadha. We can therefore deduce that on some occasions, at least, he used the Magadhan dialect of the time, which we can call Old Māgadhī." However, Norman's claim is undermined by the evidence in tables 1 and 2. They show, in what is possibly a complete count of the Buddha's locations in the first four Nikāyas, that he was recorded in Kosala more often than in Magadha. In table 1 he is recorded in Kosalan locations (including Sakyans) in 78% of the count and in Magadhan locations (including Aṅgas) in 12% of the count; in table 2, which excludes the main cities, the gap narrows to 41% Kosala and 20% Magadha, largely because Sāvattihī is excluded.<sup>16</sup> This evidence supports Warder (1970: 207), who states "In fact, the Buddha spent relatively little time in Magadha, teaching in at least half a dozen other states."

---

<sup>16</sup>Table 2 was made because Schopen (1997) argues from the Kṣudrakavastu of the Mūlasarvāstivāda-vinaya that the settings attributed to the six great cities - Śrāvastī, Sāketā, Vaiśālī, Vārāṇasī, Rājagṛha and Campā (with the Pali tradition substituting Kosambī for Vesālī) - could have been made up. The Kṣudrakavastu offers procedures whereby, if the setting or actor of a sutta was forgotten, they could be substituted by any one of the great cities or an equivalent main person - Pasenadi, if the sutta were about a king, Anāthapiṇḍaka, if a layman etc.; choosing these last two would entail their city, Śrāvastī, becoming the setting.

Accordingly, Table 1 reflects the view that the choice of these particular cities is in itself significant even if they were a default in case of memory failure. Table 2 reflects the view that settings outside the great cities offer a more accurate perspective; it includes unions to show the relative significance in the Buddha's life of Magadha and Kosala.

**Table 1: Locations of the Buddha in the first four Nikāyas**

| People    | Place (number of suttas)<br>[DPPN spelling different from the<br>Myanmar Tipiṭaka]<br>Locatives, such as <i>kosalesu</i> , show settings<br>with no other location marker.   | Total | %  |
|-----------|--|-------|----|
| Kosalans  | Sāvattḥī (1089), <i>kosalesu</i> (10),<br>Icchānaṅgala (7), Sāketa (5), Sālā (4),<br>Naḷakapāna (3), Ukkaṭṭhā (2), Uruññā<br>[Ujuññā] (2), Ayuḷjhā [Ayojjhā]*<br>(2), Ekasālā (1), Opāsāda [Opasāda]<br>(1), Cañcalikappa [Caṇḍalakappa]<br>(1), Daṇḍakappa (1), Nagaravinda<br>(1), Manasākāṭa (1), Venāgapura (1),<br>Veḷudvāra (1), Saṅkavā [Paṅkadhā] (1),<br>Sālavatikā (1), Setabyā [Setavyā] (1). | 1135  | 75 |
| Magadhans | Rājagaha (133), Nālandā <sup>†</sup> (9)/ Nālandā<br>(1), Uruvelā (9), Gayā (3), Andhakavinda<br>(2), Ambalaṭṭhikā (2), Ambasaṇḍā<br>(1), Ekanālā (1), Kallavāḷaputta<br>[Kallavālamutta] (1), Khāṇumata (1),<br>Pañcasālā (1), Pāṭaligāma (1), Mātulā (1),<br><i>magadhesu</i> (1).   | 167   | 11 |
| Vajjis    | Vesālī (48), Nātika [Nādikā] (10)/<br>Ñātika (3), Hatthigāma (3), Mithilā (2),<br>Koṭigāma (2), Ukkacelā [Ukkācelā] (2),<br>Gosiṅgasālavanadāya (2), Veḷuvagāmaka<br>[Beluvagāmaka] (2), Bhaṇḍagāma (2),<br>Bhoganagara [Bhogagāmanagara](2),<br>Ambagāma (1), Jambugāma (1).  | 80    | 5  |
| Sakyans   | Kapilavatthu (31), Devadaha (3), Silāvati<br>(2), Nagaraka [Naṅgaraka/ Sakkhara/<br>Sakkara] (1), Medāḷupa [Medataḷumpa]<br>(1), Khomadussa (1), Cātumā (1),<br>Sāmagāmaka [Sāmagama] (1), <i>sakkesu</i> (1).   | 42    | 3  |
| Vaṃsas    | Kosambī (15), Bālakaloṅakāra (1).  | 16    | 1  |

|  |  |      |     |
|--|--|------|-----|
| Aṅgas  | Campā (8), Āpaṇa (4), Assapura (2), Bhaddiya (1).  | 15   | 1   |
| Mallas   | Kusinārā (6), Uruvelakappa (3), Pāvā (3), Anupiya (1).   | 13   | 1   |
| Kāsis  | Bārāṇasi (11), Kīṭāgiri (1)  | 12   | 1   |
| Kurus  | Kammāsadhamma (8), Thullakoṭṭhika [Thullakoṭṭhita] (1)   | 9    | 1   |
| Koliyas  | Haliddavasana (2), Uttara (1), Kakkarapatta (1), Pajjanika [Sajjanela] (1), Sāmuga [Sāpūga] (1). | 6    | -   |
| Bhaggas  | Susumāragira [Suṃsumāragiri] (6)   | 6    | -   |
| Sumbhas  | Sedaka (2)/ Setaka (1)   | 3    | -   |
| Kālāmas  | Kesamutta [Kesaputta] (1)  | 1    | -   |
| Cetiyans   | <i>cetīsu</i> (1)  | 1    | -   |
| Sūrasenas  | Madhurā (1)  | 1    | -   |
| Not known  | Aḷāvi (4), Verañjā (3), Gajaṅgalā [Kajaṅgala] (1), Cālikā (1)                                    | 9    | 1   |
| Āvantikas, Assakas, Kambojas, Gandhāras, Thūlus [Khūlus/ Bumus], Pañcālas, Būlis, Macchas, Moriyas, Vaṅgas, Sunāparantans. | No record for the Buddha   | —    | —   |
| Total of records of the Buddha's location counted in the sample.   |  | 1516 | 100 |

\*Ayojjhā is given instead of Kosambī at S IV.179 in the PTS edition, but neither city is on the Ganges, as the text states. Kosambī is on the Yamunā and Ayojjhā on the Sarayū; this supports the view that the transmission became corrupt and a great city name, Kosambī, was selected as being nearest the Ganges. Ayojjhā is counted instead of Kosambī on the assumption that the name of the Sarayū was forgotten and a famous river name, the Ganges, was substituted, followed by another substitution of another famous city name.

†Nālandā is given in the DPPN entry for Kosala, but this is a misunderstanding of the ambiguous S IV.323, SN 42.9: 'for some time, the Blessed One was journeying around in Kosala with a large group of monks and arrived at Nālanda; right there in Nālandā the Blessed One stayed at Pāvārika's Mango Grove.' Here the Buddha has crossed the Ganges from Kosala to Magadha. The Pāvārikambavana is in Nālandā in Magadha, a stop on the Buddha's final journey from Rājagaha to Kusinārā (D II.81).

**Table 2: Locations of the Buddha in the first four Nikāyas**

*Excluding the great cities and showing unions*

| People  | Table 1 | Deduct great city locations  | Total | %  | Union                | Total      | %          |
|---|---------|------------------------------|-------|----|----------------------|------------|------------|
| Kosalans  | 1135    | Sāvattthī (1089), Sāketa (5) | 41    | 20 | KOSALA and Dominion  | 85         | 41         |
| Sakyans   | 42      | —                            | 42    | 20 |                      |            |            |
| Kāsis   | 12      | Bārāṇasi (11)                | 1     | 1  |                      |            |            |
| Kālāmas   | 1       | —                            | 1     | 1  |                      |            |            |
| Magadhans   | 167     | Rājagaha (133)               | 34    | 16 | MAGADHA and Dominion | 41         | 20         |
| Aṅgas   | 15      | Campā (8)                    | 7     | 3  |                      |            |            |
| Vajjis  | 80      | Vesālī (48)                  | 32    | 15 | VAJJIAN Confederacy  | 32         | 15         |
| Vaṃsas  | 16      | Kosambī (15)                 | 1     | 1  | VAṂSA and Dominion   | 7          | 3          |
| Bhaggas   | 6       | —                            | 6     | 3  |                      |            |            |
| Mallas  | 13      | —                            | 13    | 6  | INDEPENDENTS         | 42         | 21         |
| Kurus   | 9       | —                            | 9     | 4  |                      |            |            |
| Koliyas   | 6       | —                            | 6     | 3  |                      |            |            |
| Sumbhas   | 3       | —                            | 3     | 1  |                      |            |            |
| Cetiyans  | 1       | —                            | 1     | 1  |                      |            |            |
| Sūrasenas   | 1       | —                            | 1     | 1  |                      |            |            |
| Not known   | 9       | —                            | 9     | 4  |                      |            |            |
| <b>Total of records of the Buddha's location excluding great cities</b> |         |                              |       |    |                      | <b>207</b> | <b>100</b> |

**Notes to tables 1 and 2**

Table 1 was made with this procedure: noting the place names of a country in the DPPN, such as *Nālanda* in *Magadha*; searching the first four Nikāyas with the DPR for terms such as *nālanda* etc.; scanning search results to discard those with the Buddha absent, dead or merely mentioning another place;\* cross-checking for alternative place name spellings, such as *Nāḷanda*, with the terms *magadhānaṃ*, *magadhesu* etc.; scanning the indexes of the Wisdom



publications translation of each Nikāya for missing place names and countries; finally, collating the results.

Problems included: the Myanmar Tipiṭaka names are often different from the DPPN, e.g. Saṅkavā for Paṅkadhā and Kesamutta for Kesaputta; sometimes more than one setting is given in the same sutta, and each was counted; where the Buddha is travelling between two settlements, both are included on the assumption that he went to both; where the Buddha is claimed to have teleported by psychic power, he is treated as if travelling on foot, and both locations are counted; sometimes a setting can have several names, but only one place name per setting was counted;<sup>†</sup> significant differences were found in the count of cities in SN to those of Ireland (1976: 105) and Gokhale (1982: 11), so the DPR count was used as it was the lowest and least skewed the results towards Sāvattḥī.<sup>‡</sup>

Table 1 may be incomplete: some settings may have been missed due to alternative spellings or to instances of the Buddha and place name not mentioned in the same paragraph. Hence, these results are presented as a possibly complete sample. If incomplete, there is no reason to suppose that any one country would be disproportionately affected, so inferences from the relative frequency of each country's count should be secure.

### 3.2 *Māgadhābhāsā* means *Ariyaka*, Indo-Aryan

Childers (1875: vii) states “The true or geographical name of the Pali language is Māgadhī, ‘Magadhese language’, or Magadhābhāsā, ‘language of the Magadha people’”. He claims to be following tradition in this understanding, but I find no evidence in the early texts to support this view. As a computer search will confirm, nowhere in the Pali canon or its commentaries is there any reference to Māgadhī. Norman (1980: 63) states “Nowhere, to my knowledge, does Buddhaghosa state that the language of the canon in his day was Māgadhī.” MOTT advocates may claim that this is splitting hairs, and the references in the

---

\*E.g., at different locations the Buddha recalls staying at Senānigama in identical passages in MN 26, MN 36, MN 85 and MN 100, so Senānigama is not counted.

†For example, at S V.152 only Veḷuvagāmaka is counted in *bhagavā vesāliyaṃ viharati veḷuvagāmake*, as elsewhere (at D II.9.8) Veḷuvagāmaka is treated as a separate settlement with no mention of Vesālī.

‡Ireland gives the count for Sāvattḥī in SN as 2,091, Gokhale as 736. Neither author gave the editions they were using, so the figures cannot be reconciled with my count of 993 from the DPR, which amalgamates many SN suttas. In the other three Nikāyas, there was no significant difference in counts.

commentaries (but nowhere in the Tipiṭaka) to *Māgadhabhāsā* and *māgadhiko vohāro* amount to the same thing. To disprove that claim we return in more detail to the treatment by von Hinüber (1977: 239) of the disrobing procedure quoted above (§2.3):

He declares his resignation in Aryan to a non-Aryan and the latter does not understand: his resignation from the community is not valid.<sup>17</sup>

[Von Hinüber’s translation of the commentary:] Here *Ariyaka* means Aryan language (i.e.) the language of Magadha. *Milakkhaka* means any non-Aryan language (such as) that of Andha, Damiḷa, etc.<sup>18</sup>

[Von Hinüber comments] According to this passage, then, the Dravidian languages of the south, Telugu and Tamil, are contrasted with the language of Magadha, that is to say, of the Buddha.

Thus, where the Vinaya text uses the term *Ariyaka*, its commentary uses *Māgadhabhāsā*. Von Hinüber’s point is that the term ‘Pali’ does not appear in these passages, so he does not make these further inferences: (1) if the commentary and text are read together, it is clear that *Māgadhabhāsā* is defined as *Ariyaka*; (2) *Māgadhabhāsā/Ariyaka* is in contrast to Dravidian languages like Telugu, not to other Aryan dialects like Kosalī and Māgadhī; (3) the commentaries are not defining *Māgadhabhāsā/Ariyaka* as Māgadhī; (4) functionally, *Māgadhabhāsā/Ariyaka* operates as a term for all Aryan dialects. It cannot be that disrobing could only be done by affecting a Māgadhī accent or dialect, for there was not even a standard formula for disrobing<sup>19</sup>. Equating *Māgadhabhāsā* with Māgadhī is a fundamental error, and ‘Magadha’ in the commentaries must, in the context of *Māgadhabhāsā*, mean the area where Indo-Aryan is spoken. This view is confirmed by the Abhidhamma Vibhaṅga commentary which von Hinüber (1977: 240) translates:

<sup>17</sup>Vin III.27-28, DPR Vin Pārā 1.1.54: *ariyakena milakkhassa santike sikkhaṃ paccakkhāti, so ca na paṭivijānāti, apaccakkhātā hoti sikkhā. milakkhakena ariyakassa santike sikkhaṃ paccakkhāti, so ca na paṭivijānāti, apaccakkhātā hoti sikkhā.*

<sup>18</sup>Vin-a I.255, DPR Vin-a Pārā 1.1 sikkhāpaccakkhānavibhaṅgavaṇṇanā: “*tattha ariyakamaṃ nāma ariyavohāro, māgadhabhāsā. milakkhakaṃ nāma yo koci anariyako andhadamiḷādi.*”

<sup>19</sup>Vin III.26-28

The mother is a Damiḷi, the father an Andhaka. If their newborn child hears its mother speak first, it will speak the Damiḷa language. If it hears its father speak first, it will speak the Andhaka language. But if it hears the language of neither of them, it will speak the Māgadha language. For even someone born in an uninhabited forest where there is no-one else who speaks at all, even he, by his own nature, begins to speak, and it will be the Māgadha language that he speaks.....Only in this Māgadha language, rightly called the language of Brahmā, the language of the Aryans, it alone does not change. When the Completely Enlightened One entrusted the Buddha-word as contained in the Tipiṭaka to the tradition, he did so only in the Māgadha language<sup>20</sup>

Reading together these passages from the Vinaya and two commentaries<sup>21</sup> affirms that the Buddha spoke not Māgadhi, but *Ariyaka*, the canonical term for the language of Indo-Aryan speaking monks, a language assumed to be equivalent to Pali in the Theravādin consciousness. For Pali means ‘text’ or ‘language of the texts’ and, from the context, must be what the two commentaries meant by *Māgadhabhāsā*.

### 3.3 The name change from *Ariyaka* to *Māgadhabhāsā*

In considering when and why the name change from *Ariyaka* to *Māgadhabhāsā* was made, two things are immediately apparent: firstly, it is a later development than the EBTs because *Māgadhabhāsā* is found only in the commentaries; secondly, it relates to a period when Magadha had eclipsed Kosala in prominence. The most likely answer to “when” is during the Mauryan empire, especially

<sup>20</sup>Vibh-a 387-8, DPR Vibh-a 15.1.1.718, *Mātā damiḷi, pitā andhako. Tesam jāto dārako sace mātukathaṃ paṭhamam suṇāti, damiḷabhāsam bhāsissati; sace pitukathaṃ paṭhamam suṇāti, andhakabhāsam bhāsissati. Ubhinnampi pana kathaṃ assuṇanto māgadhabhāsam bhāsissati. Yopi agāmake mahārañṇe nibbatto, tattha añño kathento nāma natthi, sopi attano dhammatāya vacanaṃ samuṭṭhāpentō māgadhabhāsameva bhāsissati... Ayam evekā yathābhucca-brahmavohārāriyavohārasaṅkhātā māgadhabhāsā na parivattati. Sammāsambuddhopi tepiṭakam buddhavacanaṃ tantim āropentō māgadhabhāsāya eva āropesi.* This is a part of a long gloss, arguably a digression, on *niruttapaṭisambhidā* Vibh 297, No. 731, DPR Vibh 15.2.2.731.

<sup>21</sup>At Vin-a vi 1214, there is a third commentary, discussed at §4.6, which has *māgadhiko vohāro*, the ‘Magadhan language’, and which I claim refers back to the disrobing commentary. These three are the only Tipiṭaka commentaries where the Buddha is claimed to have spoken the ‘Magadha language’.

during the reign of Aśoka. The answer to “why” is that not only was Aśoka a Buddhist, but the entire Indo-Aryan speaking world was for the first time united under a single ruler, king of Magadha and emperor of almost all the Indian sub-continent.

That ‘Magadha’ became the name of the new empire is suggested by the Bairāt edict which starts “Priyadasi, King of Magadha”.<sup>22</sup> This minor rock edict is in Jaipur District, Rajasthan, 921 kilometres from the then capital of Magadha, Pāṭaliputra. Magadha, the kingdom in the remote Ganges Plain, would hardly impress locals half-way across India, but ‘Magadha’, an empire including them, would have been significant.<sup>23</sup> They may have counted themselves as Magadhans, just as the Sakyans were also Kosalans and the Licchavis were also Vajjis. What else was this new empire to be called but Magadha? It would be natural to acknowledge this extraordinary political development by calling the Indo-Aryan language ‘*Māgadhabhāsā*’, a practice sure to be approved of by the government. It would make sense for a standard language used across all geographical areas of the empire, i.e. Pali (§5), to have that title, if only as a technical term among Buddhists.

Norman (1980: 66f) notes that the Buddhist (and Jain) tradition thought *Māgadhabhāsā* (*addhamāgadhabhāsā*, ‘Half-Māgadhi’ in the Jain tradition) was the root of all languages. He believes this idea of language development grew up during and because of the Mauryan empire; for ‘Māgadhi’, which Norman takes as the language of the eastern Aśokan inscriptions, would also include variants of that dialect elsewhere in India. Norman (1983: 3) has a further explanation:

It is also possible that the prestige attaching to Magadha, and by implication to Māgadhi, during the time of the Mauryan kings, and also by the way in which the Māgadhi of the original Aśokan edicts was everywhere in India “translated” into the local dialect or language, led to the taking over by the Buddhists, at about the

---

<sup>22</sup>Hultzsch (1925: 172) gives the original as “*Priyadasi lājā Māgadhe saṃghaṃ abhivādetūnaṃ āhā apādhātanaṃ cha phāsuvihālatam chā*”. He translates “the Māgadha King Priyadarśin, having saluted the Saṃgha, hopes they are both well and comfortable.”, but it could be translated ‘King Piyadassi, having saluted the saṃgha in Magadha...’

<sup>23</sup>Compare the politically incorrect use of *England* to include Wales, Scotland and Northern Ireland. However, it is still correct to call their common language *English* after the dominant area rather than *British*; this somewhat parallels the use of *Māgadhabhāsā* after the dominant area, Magadha, instead of *Ariyaka*.

time of the council which the Theravadin tradition reports was held during the reign of Aśoka, of the idea that their “ruler” too employed such a language.

All these explanations are simultaneously possible and, crucially for the argument here, all agree that *Māgadhabhāsā* was not a synonym for Māgadhī. Furthermore, those restricting *Māgadhabhāsā* to Māgadhī have problems: Buddhaghosa, who allegedly wrote both commentaries quoted in §3.2, must have known some Māgadhī, yet he described Pali as *Māgadhabhāsā*.<sup>24</sup> This was not just Buddhaghosa’s idiosyncrasy. In the 12<sup>th</sup> century, King Vijayabāhu II of Sri Lanka wrote a letter to a Burmese king in *Māgadhabhāsā*, according to Cūlavam̐sa 80.6. Again, this must mean he wrote in Pali, not Māgadhī, which was hardly appropriate for international diplomacy at that time.

*Māgadhabhāsā* in the sense of ‘Pali’ was an anachronistic term for Buddhaghosa’s time, the 5<sup>th</sup> century CE, when Indo-Aryan had fragmented into several dialects. He must have been recording a tradition, but he was also engaging in propaganda, as the next section explains.

### 3.4 *Māgadhabhāsā* and diglossic competition

A positive pull towards the term *Māgadhabhāsā* during the Mauryan empire was likely, but so was resistance towards reclaiming *Ariyaka* after the collapse of this empire. For *Ariyaka* included a domineering variety, Sanskrit, one of whose grammarians, Patañjali (ca. 150BCE), criticised usages found in Pali and Ardha-Māgadhī as substandard (*apabhraṃśāḥ*, Pischel §8). With increasing Sanskritisation of inscriptions from the 1<sup>st</sup> century CE, Sanskrit replaced Epigraphic Prakrit as the H-language<sup>25</sup> of India during the 1<sup>st</sup> millennium CE. Sanskrit had long called itself, ‘the perfected language’ and ‘the language of the Gods’.<sup>26</sup> Deshpande (1979: 1-2) cites the Ṛgveda as claiming the Aryan language was spoken by the gods (RV 8.100.11) and was itself a goddess (RV 10.125.5-

<sup>24</sup>Norman (1980: 64) argued, unconvincingly in my view, that Buddhaghosa followed tradition in equating Pali, the Buddha’s language, with Māgadhī although he knew this was strictly incorrect.

<sup>25</sup>‘H-language’ is the H(igh)-language, or formal language, in diglossia, where formal and informal speech become separate varieties, e.g. in German-speaking Switzerland, where the H-language is standard German and the L(ow)-language is Swiss German.

<sup>26</sup>Pollock (2006: 44-5) finds the earliest evidence for ‘the perfected language’ in the *Vālmiki Rāmāyaṇa*, before the Common Era, and for ‘language of the Gods’ in Daṇḍin’s 7<sup>th</sup> century *Kāvyaḍarśa*. I assume both usages were current earlier than recorded.

6), so the Pali tradition may have wished to downplay *Ariyaka*, which included Sanskrit and its overbearing ideology. Instead, in an act of what Deshpande (1979: 40f) called “sociolinguistic self-defence”, *Māgadhabhāsā* was said by the Buddhists (and Ardha-Māgadhī was said by the Jains) to be ‘the root language of all people’.<sup>27</sup> It was a universal language for both Brahma Gods and men, not an elitist godly one like Sanskrit. The quote from the Vibhaṅga commentary (§3.2) implies this root language was also that of the forest-dwelling noble savage and, for any repelled by such a basic language, that commentary reminds us that the Buddha himself spoke it.

Pali commentaries may also have implied with *Māgadhabhāsā* that their language was that of the *cakravartin*, the Wheel-turning Emperor. Sujato & Brahmali (2015: 29-30) suggest that the *cakravartin* myth is a Buddhist version of the Brahmanical horse sacrifice, in that the *cakravartin* and the instigator of the horse sacrifice both ruled from sea to sea by conquest, though by non-violent conquest in the Buddhist case. Aśoka in RE13 X (Kālśī) echoes this legend by stating “they (his sons and grandsons) should regard conquest by dhamma as the only (true) conquest.”<sup>28</sup> Aśoka, the ruler of the first Magadhan (Mauryan) empire and of all *Ariyaka* speakers, was the first king in Indian history to rule from sea to sea. The term, *Māgadhabhāsā*, reinforces that parallel, hence its appearance in the commentaries attributed to Buddhaghosa, who was living during the Gupta empire, another Magadhan empire that also ran from sea to sea. This propaganda would counter pro-Brahmanical tendencies in society, evidenced by Samudragupta’s performance of the horse sacrifice in the 4<sup>th</sup> century CE (Knipe 2015: 9) and Kumāragupta’s in the 5<sup>th</sup> century (Agrawal 1989: 193).

### 3.5 *Māgadhabhāsā* changed meaning from *Ariyaka* to *Māgadhī*

After the Magadhan Empires, Magadha remained as an identifiable province, and so gradually *Māgadhabhāsā* shifted in meaning to the language of that smaller region. We have in Cūlavamsa 37.227ff the story of Revata, who asked Buddhaghosa to go to Sri Lanka to translate the Sinhalese commentaries into Pali, as India had no commentaries. Pali is there described as *Māgadhānaṃ nirutti*, ‘the language of the Magadhans’, and *Māgadhā nirutti*, ‘the Magadhan

<sup>27</sup>Cūlavamsa (PTS) 37. 244 *sabbesam mūla-bhāsāya Māgadhāya niruttīyā*. The same claim is made at Vism 441.34.

<sup>28</sup>Translation by Hultzsch (1925:49) of *tameva chā vijayaṃ manatu ye dhammavijaye*.

language'.<sup>29</sup> This leaves open whether *Ariyaka* or *Māgadhī* is meant, but other sources clearly mean the latter. Gornall (2014: 529-30) believes Moggallāna, the 12<sup>th</sup> century grammarian, was distancing his grammar from Sanskrit influenced models, such as Kaccāyana's Pali grammar, when writing:

Since grammar is manifold on account of the different (languages) such as Sanskrit etc. (sakkatādi), in order to distinguish my grammar, it is said 'Māgadha'. Māgadha words are those (words) that are understood in the Magadha region/among Magadhans. This (work) is a Māgadha grammar (lakkhaṇa) of those (words). It is said 'A grammar of Māgadha'.<sup>30</sup>

Gornall (2014: 530) comments: 'The reestablishment<sup>31</sup> of the Magadhan realm as the site of the Pāli language perhaps created a territory, albeit an imaginary one, on which the Lankan and Cōla saṅghas could stake their claim. The Pāli language was no longer a shared, transregional idiom but the site of a struggle between two competing monastic traditions [i.e. between the Moggallāna and Kaccāyana traditions].'

We also have what I believe to be an early use of the term *Māgadhī*, meaning Pali, in a poem added to a sub-commentary on the *sakāya niruttīyā* passage (see §4) in the 12<sup>th</sup> to 15<sup>th</sup> century handbook *Vinayālaṅkāraṭīkā* (34.46): *sā māgadhī mūlabhāsā; Narā yāyādīkappikā; Brahmāno cāssutālāpā; Sambuddhā cāpi bhāsare*. 'This Māgadhī is the root language. Men of whatever age, Brahma Gods who have not heard a word and fully enlightened ones speak it.' Around this time, the term *pālibhāsā*, meaning 'language of the texts', came into being according to Crosby (2004), who finds its first definitive use in the *Vinayatthasārasandīpanī* (12<sup>th</sup>-13<sup>th</sup> century, Sri Lanka). This uses *pālibhāsā*, 'the

<sup>29</sup>Cūlavamsa (PTS) 37.230: *Taṃ tattha gantvā sutvā tvaṃ Māgadhānaṃ niruttīyā parivatthehi*. Cūlavamsa (PTS) 37.244: *parivattesi sabbā pi Sīhalaṭṭhakathā tadā/ sabbesaṃ mūlabhāsāya Māgadhāya niruttīyā*.

<sup>30</sup>Gornall's translation of Moggallāna-pañcikā 1. 33–1: *saddalakkhaṇassāpi sakkatādibhedena bahuvīdhattā sakaṃ saddalakkhaṇaṃ visesayitum āha māgadhan ti magadhesu viditā māgadhā saddā tesam idaṃ lakkhaṇaṃ māgadhaṃ, idaṃ vuttaṃ hoti māgadhaṃ saddalakkhaṇaṃ ti*. [Text from Gornall]

<sup>31</sup>Gornall follows the MOTT misreading of *Māgadhābhāsā* (§3.2), and I argue 'establishment' is the correct term.

Pali language’, as a language name in contrast with *sīhaḷa*, ‘Sinhalese’. There may be a connection between these developments, as *pālibhāsā* may have been an alternative to *Māgadhabhāsā* for non-supporters of Moggallāna’s grammar who did not wish to imply Pali was Māgadhī.

Canonical usages such as *magadhakhetta*, ‘Magadhan field’, for the robe pattern based on the contours of a Magadhan rice field (Vin I.287) and ‘Magadhan’ for a type of garlic (Vin IV.259) referred to the province. The Vinaya commentary used both senses of ‘Magadha’, both ‘province’ and ‘empire’, but ‘province’ became the norm in the later tradition. For, by the time Childers (§3.2) and D’Alwis were consulting their Sri Lankan mentors, Māgadhī was understood by *Māgadhabhāsā*. D’Alwis (1863: xcvi) attests “the promiscuous use of the terms Pali and Magadh in Ceylon”. This conflation of the two terms has no authority from the canon or its commentaries. However, Buddhaghosa’s non-canonical use of *Māgadhabhāsā*, referring to an empire which did not exist when the EBTs were created, unintentionally misdirected a later Sri Lankan tradition towards equating Māgadhī and Pali. In turn modern scholars have amplified this misdirection and created the MOTT to explain the discrepancy between Pali and Māgadhī.

### 3.6 Eastern Aśokan inscriptions belie linguistic diversity

Norman (1980: 65) linked Māgadhī to the eastern Aśokan inscriptions, which are found throughout the area travelled by the Buddha. He gave as Māgadhī’s features: the nominative singular of short *-a* stems is *-e* instead of Pali *-o*; *l* occurs instead of Pali *r*; all sibilants become *ś* instead of Pali *s*. However, an east-west language division is unsatisfactory. Ardha-Māgadhī, which is supposedly eastern, has the same three features as Pali, although not to the same degree. Norman (1980: 68) argues that, although Aśoka’s Kalinga inscriptions in the geographical east have nominative singular endings in *-e*, the later Hāthīgumphā inscriptions of that area have ‘western’ endings in *-o*.

As Cousins (2013: 120) states:

The significant point is that the Eastern or Eastern-influenced dialect of all other Mauryan inscriptions in India cannot have been the local or ordinary spoken dialect of most people in the majority of the places where it is used. That this is so is indicated rather clearly by the fact that no post-Mauryan inscriptions in this dialect are extant.



We cannot, therefore, infer from the eastern Aśokan inscriptions that this language was universally spoken in the Ganges basin, either when they were inscribed, or in the Buddha’s day two centuries earlier.

### **3.7 Māgadhī could not have changed into Pali**

Intentional change has been ruled out in §2, but Cousins (2013: 121) believed natural language change could account for the difference between Māgadhī and Pali:

The language used in the Indian inscriptions of Aśoka was the state language of the kingdom of Magadha; it can only have been called the Māgadhā or Māgadhī language. I can see no reason to suppose that the administrative or cultural change which led to the adoption of some western dialect features would have required a change of name.

However, Pali cannot be said to have adopted “some” western dialect features; it is overwhelmingly ‘western’, with a very few ‘eastern’ features. A second difficulty is that one variety changing into another variety is not a phenomenon acknowledged in linguistics. In the case of mutual unintelligibility, Barnes (2010: 39) rejects the theory that Norn, an extinct Scandinavian language spoken in Shetland within the past three centuries, gradually changed into the closely related Scots: “...the imperceptible melting of one language into another they [other scholars] envisage seems to be without parallel.” If we assume the alternative, that Pali and Māgadhī were mutually intelligible, contact linguistics would predict dialect mixing or koine creation in newly settled areas and dialect-levelling in established areas, but never one variety almost turning into another. Moreover, the ideal of fixed transmission would limit change drastically. In no case could Māgadhī naturally change into Pali.

### **3.8 Alleged remaining Māgadhisms in Pali**

The MOTT narrative is that Māgadhī was changed into ‘western’ Pali, except for a few ‘eastern’ features, called Māgadhisms, which are allegedly residual frozen phonetics.

The Kathāvatthu is a challenge to the MOTT narrative that Māgadhisms signify translation. Norman (1979: 284) points out a frequent contrast many times in that work between a first speaker who uses mainly the nominative singular

-o ending (except for set phrases and repeating his interlocutor) and a second speaker who uses the -e ending. The differentiation between the speakers, the time of the Kathāvatthu in Aśoka's reign, when Māgadhī would have become a prestigious dialect, and the setting, in Aśoka's capital, Pāṭaliputra, in the heart of the province and empire of Magadha, all suggest that the Māgadhisms are original, native-language features of the second speaker. In the case of the first speaker, the intermittent use of the -e ending seems to be a contact Māgadhism in a process of accommodation.<sup>32</sup> I have yet to see a MOTT explanation of why the Kathāvatthu was so incompetently 'translated'. It makes better sense that this work, like other parts of the Tipiṭaka, is an accurate record of the original dialects, even to the point of recording contact Māgadhisms, a sociolinguistic feature probably unidentified by ancient grammarians.

Moreover, it is far from clear which features are 'eastern' and which 'western'. We should write of 'pre-eastern' and 'pre-western' features, if we accept the logic of Brough (1962: 115):

The classification of the Aśokan inscriptional dialects under the heading 'eastern' and 'western' appears to be due chiefly to extrapolation from linguistic evidence of a much later period; and the distribution of -e/-o could hardly be taken as defining an isogloss.

Pre-eastern features are found in the west. The pre-eastern nominative singular in -e, emblematic of Māgadhisms, is found on the opposite side of Aśoka's India, in modern Pakistan, at Shābāzgarhī. We have also seen the pre-eastern *lājā* instead of *rājā* in the west at Bairāt in Rajasthan (§3.3 fn). Thus, 'Māgadhisms' in the Aśokan era had a wider distribution than the province of Magadha, making it difficult to determine which are the real Māgadhisms in Pali.

Apparent Māgadhisms of two types could be present in Pali: one explained by spelling convention, the other by natural language change. An example of the first is Brough's suggestion that -e, as well as -o, was a genuine Gāndhārī form of the nominative, rather than a Māgadhism; it represented developments of a vowel of mixed quality that could not be adequately notated by either ending.<sup>33</sup>

---

<sup>32</sup>*Accommodation* in sociolinguistics refers to speakers adjusting their speech to one another, usually to minimise differences and demonstrate mutual understanding.

<sup>33</sup>Brough (1962: 115) explains "So far as we can tell from the writing, the vowel of the inflexional ending ultimately came to coincide with either the *o* or the *e* (according to period

This phenomenon could apply to Pali, as Gombrich (1994: xxvii) argues:

Before the texts were ever written down, it is not likely that their dialect was ever completely fixed, or even that the differences between the dialects were clearly conceptualised; it must have been a matter of reciting in what appeared like “regional accents”. In the last resort, Pāli was formed at the phonetic level by the spelling conventions which the first scribes chose to adopt.

As for natural language change, we know from recordings of 80-90 years ago that even in a standard accent, such as Received Pronunciation in British English, vowels change so much that the old accent sounds strange to modern ears. The Uniformitarian Principle, associated with the Sanskritist, Whitney (Hazen 2011: 30), is also promoted by the linguist, Labov (1972: 101): “the linguistic processes taking place around us are the same as those that have operated to produce the historical record.” We must presume a similar fluidity in Pali as in English in the course of oral transmission over several centuries.

I therefore take issue with the practice, found in standard works like Geiger (1916) and Lüders (1954) of automatically regarding all *-e* for *-o* forms as ‘Māgadhisms’. The vocative plural *bhikkhave* is an example. In a computer search of the first four Nikāyas, the vocative plural *bhikkhavo* is found only in sentence final position in prose and is the only verse form, whereas the supposed ‘Māgadhism’ *bhikkhave* is essentially an enclitic which occurs mid-sentence.<sup>34</sup> The difference could simply mark a final secondary stress in *bhikkhavo* by giving *-o* its full length and mark no final stress in the enclitic *bhikkhave*. Typically, *bhikkhave* serves as a pragmatic marker<sup>35</sup> to introduce or emphasise

---

and dialect) which represent the older diphthongs. This suggests the hypothesis that at an earlier stage of Indo-Aryan, the ending might have been a vowel of mixed quality: a front rounded vowel [œ:], or a back unrounded [ɤ:]. Either of these qualities would easily be understandable as a development from a rather close central vowel [ə:], which would be theoretically expected as a *sandhi*-variant beside *-ah*, *-as*, [əh, əs], if the phonetic differentiation of *a*, *ā* [ə, a:] had already taken place. If a dialect had still preserved an ending such as [œ:, ɤ:, or ə:] when it was first reduced to writing, there is no inherent difficulty in supposing that either of the two signs, *o* or *e*, might have been felt to be reasonably adequate notations.”

<sup>34</sup>There can be sentence-final use of *bhikkhave* if phrases like *āvuso bhikkhave*, and *taṃ kiṃ maññatha bhikkhave, api...* have the usual commas replaced by full-stops. However, I have not found unambiguous sentence-final *bhikkhave* in the Burmese or PTS editions.

<sup>35</sup>*Pragmatic marker* is a term in pragmatics to indicate words not part of the propositional content of a sentence, e.g. vocatives, conjunctions, disjuncts (*frankly, fortunately*) etc. Fraser

a topic. On the other hand, *bhikkhavo* is a pragmatic marker inviting an answer. For example, the pericope, *tatra kho bhagavā bhikkhū āmantesi — bhikkhavo ti. bhaddante ti te bhikkhū bhagavato paccassosum*, ‘right there the Blessed One addressed the monks: “Monks!” “Sir,” the monks answered the Blessed One.’<sup>36</sup> There is an analogous phonological change in English *the* to mark different functions: /ðə/ frequently refers to a subject already introduced, but /ði/ introduces a famous name, as in *I’m meeting the David Attenborough*, and is a pragmatic marker with the illocutionary force of ‘I expect you have heard of this person’. The alternative explanation of ‘Māgadhism’ is asserted in these standard works without any examples of an *-ave* form in any dialect but Pali and without examples of any Pali plural but *bhikkhave*.<sup>37</sup> It is more likely that the pronunciation of the original *bhikkhavo* changed in enclitic positions.<sup>38</sup>

---

(1996: 186) gives the example of a vocative, *my friend*, as a solidarity marker and *bhikkhave/o* may also have had that function.

<sup>36</sup>Giving the final syllable its full length may also be analogous to Pāṇini’s Aṣṭādhyāyī Rule 8.2.83 where a final long, high-pitched syllable is used as a *pratyabhivādana*, a response to a respectful greeting, except in the case of a Śūdra. The Buddha, having taken his seat, may have received bows from the audience and used *bhikkhavo* as a formal acknowledgement.

<sup>37</sup>Geiger (1916: §82) claims *bhikkhave* is “a ‘Māgadhism’ which has penetrated into the literary language from the popular speech”, but he did not have the advantage of modern computer searches; Lüders (1954: 13, §1) also asserts a Māgadhism. Their case would be more convincing if they had produced examples of an *-ave* ending in any variety, apart from the vocative plural in Pali. So far as I know, no such examples exist: the Aśokan inscriptions do not have *-ave* or *-avo*; Pischel §381 has only AMg *bhikkhavo* vocative plural and Pischel §378-381 on the *-u* declension offers no *-ave* forms for any case in any Prakrit, including Māgadhī. In Pali, only *bhikkhave* seems to have this *-ave* ending and only for the vocative plural.

Lüders et al. (1963: 70), writing of the Bharhut inscriptions of western India, go further than Geiger, claiming that the stem *bhikkhu* is also a Māgadhism: “When translating into the Western language, which we are used to call Pāli, not only numerous faults occurred, but at many places the Eastern forms have been retained. So, for instance, in the Eastern language the *ksh* of *śaiksha* and of *bhikkshu*, *bhikkshuṇi* became *kkh*, in the Western language, however, it became *chchh*. But *sekkha*, *bhikkhu*, *bhikkhunī* were taken over without change as technical expressions in the church language.” The index (Lüders et al. 1963: 198) shows 4 occurrences for *bhikkhunī* at Bharhut and 10 for *bhichhunī*. (Sāñchī inscriptions show both *bhikhu* and *bhichu*.) The fact that *bhichhunī* is only western does not mean *bhikkhunī* is only eastern. An alternative explanation is that in the West lenition of *bhikhu* and *bhikkhunī* was developing, but this was not reflected in the earlier and conservative Pali.

I consider these claims of Māgadhisms to be completely unfounded.

<sup>38</sup>Bechert (1991: 11f) claimed *bhikkhavo* was a later form which replaced *bhikkhave* in the above pericope, which was added later, and also in Sutta Nipāta, which was otherwise full of Māgadhisms. Anālayo (2011: 22) finds some ambiguous evidence for Bechert’s view, that the

The SOTT also predicts transmission Māgadhisms: that speakers of pre-eastern varieties would inadvertently colour the transmission of texts with their own accent. This could explain inconsistent use of pre-eastern features. Norman (1976: 118ff) points out that Māgadhisms are found only for three of six Ājīvika type ascetics, and not consistently within the three. For Pakudha’s doctrine, the varieties are mixed even in the same passage at D I.56: “*pathavikāyo, āpokāyo, tejokāyo, vāyokāyo, sukhe, dukkhe, jīve sattame*”, ‘The seven are the earth element, the water element, the fire element, the wind element, pleasure, pain and life’.

Thus it should not be claimed that the existence of the very few Māgadhisms in Pali ‘proves’ incompetent translation from Māgadhī. Original and contact Māgadhisms can be taken as proof of non-translation; elsewhere, apparent or transmission Māgadhisms are more plausible alternatives to the MOTT narrative.

### 3.9 Conclusion: Māgadhī is a false trail

Current scholarship has followed a false trail by accepting the later tradition that *Māgadhabhāsā* is local and equals Māgadhī, and ignoring the commentaries of Buddhaghosa that it is trans-regional and (Indo-) Aryan, meaning Pali. This has led to unfounded claims that the Buddha’s main connection was with Magadha, the province, that his language was a precursor to the Eastern Aśokan inscriptions and that Pali is an imperfect translation of Māgadhī.

Norman (§3.3), for example, saw that *Māgadhabhāsā* was not synonymous with Māgadhī, but did not extricate himself from the academic consensus that identified it with that specific dialect. The *Einstellung* effect may explain this consensus. Bilalić and McLeod (2014: 75ff) explain that the *Einstellung* effect in psychology refers to an obvious solution to a problem, in this case *Māgadhabhāsā* meaning Māgadhī, blocking access to a better solution, in this case *Māgadhabhāsā* meaning *Ariyaka* and Pali. We move on to that better solution in §5.

---

above pericope is absent from *Madhyama Āgama*, but is found in a tradition derived from it; however, he also believes a difference of emphasis is plausible. Manné (1990: 82) believes that pericopes are likely to be original features, which supports *bhikkhavo* as the original, as does the *-avo* ending in AMg. and Mg. In fact, *bhikkhavo* is the form used in poetry, even in enclitic positions, at Sn 280, Sn 385, D II.272, A IV.89, It 41 and Ap 299. The entire Be Khuddhaka Nikāya has no instances of *bhikkhave* in verse and I believe this holds true throughout the Pāli canon. This implies that the final syllable of *bhikkhave* was a shortened sound prosodically.

#### 4. Sakāya niruttiyā

Oberlies (2003: 166), who has provided a recent and clear explanation of the MOTT (see §6), has made the following claim:

The Buddha is reported to have said that his teachings should be given to the people not in Sanskrit, but in their own language.

Oberlies does not offer a reference for this point, but almost certainly, like Norman (§2.1), he is referring to the *sakāya niruttiyā* passage in the Cūlavagga of the Vinaya.<sup>39</sup> MOTT advocates will translate the passage broadly as Edgerton (1953: 1) does:

Two monks, brothers, brahmins by birth, of fine language and fine speech, came to the Buddha and said: Lord, here monks of miscellaneous origin (literally, of various names, clan-names, races or castes, and families) are corrupting (*dūsentī*) the Buddha's words (by repeating them in) their own dialects. Let us put them into Vedic. The Lord Buddha rebuked them; Deluded men, how can you say this? This will not lead to the conversion of the unconverted... And he delivered a sermon and commanded (all) the monks: You are not to put the Buddha's words into Vedic [*chandaso*]. Who does so would commit a sin. I authorise you, monks, to learn the Buddha's words each in his own [*sakāya*] dialect [*niruttiyā*]. [Pali wording added]

Edgerton claims support from Chinese translations of Vinaya sources.<sup>40</sup>

<sup>39</sup> Vin II.139, DPR Cv 5.285. *tena kho pana samayena yameḷakekuṭṭā nāma bhikkhū dve bhātikā honti brāhmaṇajātikā kalyāṇavācā kalyāṇavākkaraṇā. te yena bhagavā tenupasaṅkamimsu, upasaṅkamitvā bhagavantaṃ abhivādetvā ekamantaṃ nisīdīmsu. ekamantaṃ nisinnā kho te bhikkhū bhagavantaṃ etadavocum — “etarahi, bhante, bhikkhū nānānāmā nānāgottā nānājaccā nānākulā pabbajitā. te sakāya niruttiyā buddhavacanaṃ dūsentī. handa mayaṃ, bhante, buddhavacanaṃ chandaso āropemā” ti. vigarahi buddho bhagavā ... pe ... “kathaṅhi nāma tumhe, moghapurisā, evaṃ vakkhatha — ‘handa mayaṃ, bhante, buddhavacanaṃ chandaso āropemā’ ti. netāṃ, moghapurisā, appasannānaṃ vā pasādāya ... pe ... vigarahitvā ... pe ... dhammiṃ kathaṃ katvā bhikkhū āmantesi — “na, bhikkhave, buddhavacanaṃ chandaso āropetabbaṃ. yo āropeyya, āpatti dukkaṭassa. anujānāmi, bhikkhave, sakāya niruttiyā buddhavacanaṃ pariyāpuṇitun” ti.*

<sup>40</sup> Edgerton (1953: 2) offers three Chinese sources translated by Lin Li-Kouang:

**Mahīśāsaka Vinaya:** “...the two originally brahman brothers heard monks reciting the sūtras ‘incorrectly’. They ridiculed them, saying: ‘Tho’ they have long since become monks, they recited the sūtras in this fashion! not knowing masculine and feminine gender, nor singular and plural, nor

However, his reading cannot be correct:

1. This interpretation is in direct contradiction to the evidence in §2 that oral translation was impractical, discouraged and unnecessary. This includes 18 passages stressing the value of learning Suttas accurately, to the syllable.
2. The *Araṇavibhaṅga Sutta* (M 139) discourages the use of dialect words (§5.1).
3. Buddhaghosa's commentary on *sakāya niruttiyā* explains the Buddha's teaching is to be learnt in the Buddha's speech (§4.6).
4. The Vinaya text could have been *sakasakāya* instead of *sakāya* if 'each in their own dialect' was intended, but the text does not specify that meaning.
5. The two occurrences of *sakāya* are separated by seven sentences in the PTS edition, so there is no grammatical or logical need for them to have the same referent.

---

present, past and future, nor long or short sounds (vowels), nor (metrically) light and heavy sounds (syllables).’ When they appealed to the Buddha, he ordered that the texts be recited ‘according to the sounds of the regions, but taking care not to distort the meaning. It is forbidden to make of the Buddha’s words an “outside” (non-Buddhist, heretical) language.’”

**Dharmaguptaka Vinaya:** “... a monk... complained to the Buddha that ‘monks of different clans and bearing different names were ruining the sūtras’, and proposed ‘to arrange them according to the good language of the world’, that is, no doubt, Vedic or Sanskrit, the language of culture. In his rebuke the Buddha said it would ruin the sūtras to use ‘the language of heretics’, and that ‘it is allowed to recite and learn the Buddha’s sūtras according to the interpretation of the popular languages of (various) regions.’”

**Vinayamāṭṛkā:** (Affiliation unknown). The Buddha states “In my religion, fine language is not recognised. All I want is that meaning and reasoning be correct. You are to preach according to a pronunciation (lit. sound) which people can understand. Therefore it is proper to behave (sc. in the use of language) according to the countries.”

Lamotte (1958: 611-13) also offers a translation from the **Sarvāstivādin Vinaya:** Two brahmins who had been converted to Buddhism ... had recited the texts of the four heretical Vedas ... they recited the Buddhist sutras with the same intonations.... (one) reported the matter to the Buddha. The Buddha said to him: “Henceforth, whoever recites the Buddhist sutras with the intonations of heretical books will be committing a misdeed.” and **Mūlasarvāstivādin Vinaya:** The Buddha ... announced the following regulation: “... If bhikṣus recite the sutras with the intonation of the *chan t’o* (*chandās*), they will be guilty of the offence of transgressing the Dharma. However, if a regional pronunciation requires the intonation to be long, it is not wrong to do that.”

6. If the MOTT claims that the normal role of *saka*, of referring to the grammatical subject of a sentence, can be stretched in its second occurrence to implied subjects, then *a fortiori* it must also equally allow the translation, ‘I require you, monks, to recite the Buddha’s teachings in **our** language’, a translation which does include the grammatical subject (among others) and which undermines the MOTT reading.
7. The most up to date dictionary does not give *dialect* or *language* for *nirutti*. Cone (2010: 607) has only *explanation, interpretation, expression, form of words, way of speaking, alternative terminology, or gloss*. (The earlier PED does have *dialect*.)
8. It is doubtful that any of the five Chinese sources are referring to mutually unintelligible varieties that would require translation. Levman (2008/9: 42, 43) states that all five Chinese sources refer to pronunciation, and only one source, the Dharmaguptaka Vinaya, is about language, as well as pronunciation. He translates it as the Buddha allowing “the sounds and common language of the country to be used in learning the scriptures by recitation and explanation.”<sup>41</sup> Levman believes that the translator was not accurate elsewhere and may have wished to legitimise the translation into Chinese with *common language*. Both ‘sounds’ and ‘common language’ are outside the scope of *nirutti* as given by Cone. Perhaps a conscientious Chinese translator, uncertain of which nuance of *nirutti* was intended and checking other sources, indicated a range of possibilities.<sup>42</sup> Thus, this one equivocal Chinese source does not confirm the reading of ‘dialect’.
9. There are at least six alternative and more plausible readings, the first two take *sakāya* as Edgerton does, the last four follow Buddhaghosa in referring to the Buddha’s speech:

---

<sup>41</sup>Lin Li-Kouang in Edgerton omits ‘sounds’ from the Dharmaguptaka version, Brough (1980: 39-40) is ambiguous.

<sup>42</sup>Sincere, competent translators inevitably interpret their sources. For *sandīṭṭhiko dhammo*, literally ‘dhamma which is completely visible’, Bhikkhu Ñaṇamoli (1995: 358) has “Dhamma, which is visible here and now,” though ‘here’ and ‘now’ are nowhere in the original Pali; Bhikkhu Bodhi offers (2000: 98) “this Dhamma is directly visible” even though ‘directly’ is not a normal translation of the prefix *sam*; Gombrich (2018: 59) has the non-literal ‘practical’ for *sandīṭṭhiko*.



#### 4.1 Reading 1: Ruegg (*chandaso* = Vedic chant, *niruttiyā* = delivery)

Ruegg (2000: 306) follows Lévi (1915) in taking *nirutti* as meaning prosodic or phonological features. He states:

In keeping with the prosodic interpretation, the problematic expression *sakāya niruttiyā* seems more intelligible in its context if it is understood to refer to the individual reciter's manner of speaking (on both the prosodic and phonological levels). In other words, it is perhaps best understood as meaning not 'in one's own language/dialect' but, instead, 'in one's own speech' (i.e. vocal delivery).

As evidence, he cites the Cullavagga (Vin II.107), where a protracted melodic recitation style, *gītassarena*, is not allowed, but *sarabhañña*, 'plainchant', is allowed. This is consistent with the description of the Yamelakekuṭa brothers as *kalyāṇavācā*, which Ruegg translates as "whose voices were good, whose vocal delivery was good". This sense of *nirutti* is not given by Cone, but it is in line with all five Chinese versions.

#### 4.2 Reading 2: Gombrich (*chandaso* = Vedic chant, *niruttiyā* = gloss)

Gombrich (2009: 147) takes *chandaso* as recitation in a particular Vedic reciting style with pitch accents. As he accepts that monk learnt texts word for word, *sakāya niruttiyā* means 'their own mode of expression' and refers to explanatory glosses or paraphrases given in their own dialects.

#### 4.3 Reading 3: Geiger (*chandaso* = Vedic Sanskrit, *niruttiyā* = language)

Geiger (1916: 7) follows Buddhaghosa's commentary (§4.6) and translates, "I ordain the words of the Buddha to be learnt in his own language (in Māgadhī, the language used by the Buddha himself)." Geiger regarded the Buddha as not speaking a pure Māgadhī, but a form of popular speech. Edgerton (1953: 2) is well aware of the commentary and Geiger's view, but dismisses them on the grounds that Geiger is in a minority of scholars and that the Chinese sources are against him and Buddhaghosa. He does not address Geiger's arguments: that the comparison in question is with dialect and Vedic, not between non-Vedic dialects, that there is no *vo* to connect *sakāya* to the monks' *nirutti* and that memorisation of the founder's words is in accordance with Indian custom.

**4.4 Reading 4: Norman (*chandaso* = as desired, *niruttiyā* = gloss)**

Norman (1971: 331) paraphrases the Brahmin brothers: “let us translate into the various vernaculars to meet the various needs of these different people who cannot cope with the language of the Buddha’s words”. According to Norman, the Buddha then refused translation. He (1980: 63) thought that the Buddha varied terms or glosses to meet local needs without translating all his speech, and other reciters followed his practice. He consistently believed *sakāya* cannot have different referents, but decided (1992: 83) it referred to *buddhavacanam*, and not to the Buddha, as he previously thought.

**4.5 Reading 5: Levman (*chandaso* = Vedic Sanskrit/ chant, *niruttiyā* = name)**

Levman (2008/9: 39) makes a case for *nirutti* to mean ‘names’ in the sense of ‘technical terms’ and takes *dūśenti* to be the confusing of these terms. He translates, “Monks, I enjoin the Buddha’s words be learned with its (my) own names.” He has no difficulty with *sakāya* having different referents.

**4.6 Reading 6: Karpik (*chandaso* = verse, *niruttiyā* = way of speaking)**

Though it is something of a rarity in discussions of this passage, the obvious meaning of *chandaso* as ‘into verse’ must be considered.<sup>43</sup> I interpret the crucial sentence as ‘Monks, I require you to learn the Buddha’s words in my (his/ its) own way of speaking (i.e. prose or verse, whichever was the original).’ Or, to paraphrase, ‘Monks I require you to learn the Buddha’s words in the original.’

Thomas (1927: 254) takes *chandaso* as ‘metre’ and, on the basis of a tendency to versify late texts, notably the Parivāra of the Vinaya, he concludes “there was once an attempt to versify the Canon, and it was rejected, at least to the extent that the versifications were not allowed to take the place of the fundamental texts.” It must have been a daily frustration for brahmin converts to Buddhism to memorise suttas in prose when they had been trained from boyhood to use metre as a memory aid. This frustration must have been aired with the Buddha at some point, and the advantages of metre suggested to him, so it would be unsurprising to have a record of

---

<sup>43</sup>Thomas (1927: 254) has ‘metre’, but gives *sakāya niruttiyā* as ‘in its own grammar’. As I cannot see how *nirutti* can mean ‘grammar’, I have altered his reading to Cone’s ‘way of speaking’. Horner (1963: 194) has ‘into metrical form’.

that suggestion being made. Another advantage of metre, the brothers might think, would be reducing variation; for example, metre would eliminate variation in long and short vowels. Communal chanting, as encouraged in the *Pāsādika Sutta* (§2.2), in a variety of broad accents would offend them as lacking clarity. This is the minimum of what the complaint, *dūsentī*, ‘they are spoiling’, means and is sufficient to understand this passage.

Buddhaghosa’s commentary on this passage is given in full:<sup>44</sup>

‘Fine speech’ – pleasant sounding. ‘Let us elevate it into verse’ [*chandaso*] - let us put the words into the style of refined language like the Veda. ‘Own way of speaking’ [*sakāya niruttīyā*] – In this case, ‘own way of speaking’ is the previously stated<sup>45</sup> Magadhan language, [*māgadhiko vohāro*]<sup>46</sup> of the Perfectly Awakened One.

Buddhaghosa’s “style of refined language like the Veda” appears to be a periphrasis for ‘gāthā language’, which had archaic forms reminiscent of the Vedas. The Brahmin brothers would have been proposing the *gāthā* verse style, putting Buddhist prose into Buddhist verse. The Buddha refused and this refusal would be consistent with the prohibition of *gītassarena* (§4.1), which can be seen as a wish to make chanting nearer normal speech and less like song. *Māgadhiko vohāro*, ‘the Magadhan language’ refers to *Māgadhabhāsā* and *Ariyaka* (§3.2-5) and signals the ordinary, inclusiveness of the Buddha’s speech, the *mūlabhāsā*, ‘root language’, of the noble savage, a speech far from the exclusiveness imputed by MOTT advocates and to

<sup>44</sup>Vin-a VI.1214, DPR Vin-a Cv 5.285 *kalyāṇavākkaraṇāti madhurasaddā. chandaso āropemāti vedaṃ viya sakkatabhāsāya vācanāmaggaṃ āropema. sakāya niruttīyāti ettha sakā nirutti nāma sammāsambuddhena vuttappakāro māgadhiko vohāro.*

<sup>45</sup>‘Previously stated’ refers to the section on disrobing in Vin-a I.255, DPR Vin-a Pārā 1.1.54 (§3.2), the commentary on Vin III.27-28, DPR Vin Pārā 1.1.54 (§2.3). Norman (1980: 61) rejects Horner’s “the current Magadhese manner of speech according to the awakened one” and (1980: 62) offers “the Māgadhī terminology in the form spoken by the Enlightened one”. However, I stand by my translation, which is supported by the PED entry for *pakāra*.

<sup>46</sup>‘The term, Magadhan’ contrasts with ‘refined’ [*gāthā*] language like the Vedas. Buddhaghosa thought the brothers cared about aesthetics (‘fine sounding’, ‘refined language’) and wished his readers to be clear that the Buddha’s prose language was to be followed. By referring back to the disrobing passage, he is contrasting *Māgadhabhāsā* with non-Aryan languages and implying wide-spread, non-elitist, language of no particular variety.

which they rightly object<sup>47</sup>. Buddhaghosa was implying the Magadhan language, i.e. Pali, was truly a language for all.<sup>48</sup>

This theme of non-elitism fits in well with the five Chinese sources who all refer to the Buddha allowing local sounds. Although the Buddha recommended standard language, he would not, like the Vedic reciters, seek to fix pronunciation to ensure the efficacy of ritual, but had to allow to some degree regional accents in the delivery of the standard language, given that Vedic techniques of transmission were not used. Doubtless, monks would try to tone down their local accent, if any, to conform to the western standard during recitation of texts (§5.6), but not everyone can adapt their accent fully. Hence, the transmission would have been somewhat phonetically fluid (§2.4).

That the Buddha was not elitist over pronunciation and did not want his prose elevated into elitist verse is my reading of the Pali Vinaya passage, its commentary and also of the Chinese sources.

#### 4.7 *Sakāya niruttiyā*: a red herring

In conclusion, the meaning of the expression *sakāya niruttiyā* remains controversial, but provides no solid evidence for the MOTT. The MOTT reading, exemplified by Edgerton, is merely one of seven interpretations and it is by far the worst for the reasons given above. I leave the reader to decide which of the other six readings is best as that is not relevant to this argument.

### 5. Inferring the Buddha's teaching language

If we accept the evidence for the ideal of a fixed transmission (§2.2), this leaves

---

<sup>47</sup>MOTT advocates might mistranslate “Let us put it into Vedic’ [*chandaso*] – let us put our manner of speech into the Sanskrit language like the Veda. Here ‘own dialect’ [*sakāya nirutti*] is the manner of speech, the Māgadhī usage [*māgadhiko vohāro*] of the Perfectly Awakened One” and thus reject this gloss.

<sup>48</sup>This addresses misreadings of Buddhaghosa, of which Law (1933: xii-xiii) is a forthright example: “It is beyond our comprehension how Buddhaghosa went so far as to suggest that by the term *sakāyanirutti*, the Buddha meant his own medium of instruction and nothing but Māgadhika or the Māgadhī dialect. Nothing would have been more distant from the intention of a rational thinker like the Buddha than to commit himself to such an opinion which is irrational, dogmatic and erroneous. He could not have done so without doing violence to his position as a *sammādiṭṭhika* and *vibhajjavādin*. To give out that Māgadhī is the only correct form of speech for the promulgation of his teachings and every other dialect would be an incorrect form is a *micchādiṭṭhi* or erroneous opinion which the Buddha would ever fight shy of. Buddhaghosa has misled us all.”

as the only possibility a single variety suitable for use across all Indo-Aryan-speaking areas, i.e., a standard language.

### 5.1 Textual evidence for a standard language

In the *Araṇavibhaṅga Sutta* [MN 139, M III.234-5], already discussed in the context of mutual intelligibility of varieties (§2.3), the Buddha recommended standard vocabulary. Horner (1959: 282) translates:

When it is said: ‘One should not affect the dialect of the countryside, one should not deviate from recognised parlance,’ in reference to what is it said? And what, monks, is affectation of the dialect of the countryside and what is departure from recognised parlance? In different districts they know (the different words): Pāti ... Patta ... Vittha ... Sarāva ... Dhāropa ... Poṇa ... Pisīla. Thus, as they know the word as this or that in these various districts so does a person, obstinately clinging to it and adhering to it, explain: ‘This indeed is the truth, all else is falsehood’. Thus, monks is affectation of the dialect of the countryside and departure from recognised parlance. And what, monks, is non-affectation of the dialect of the countryside and non-departure from recognised parlance? In this case, monks, they know (the different words) Pāti .... Patta ... Poṇa ... Pisīla, yet although they know the word as this or that in these various districts a person does not cling to it and explains: ‘These venerable ones definitely explain it thus.’ Thus, monks, is non-affectation of the dialect of the countryside and non-departure from recognised parlance. When it is said: ‘One should not affect the dialect of the countryside [Cone: ‘local terminology’], one should not deviate from recognised parlance,’ it is said in reference to this.<sup>49</sup>

<sup>49</sup>DPR M3.4.9, MN139 *Araṇavibhaṅgasutta*. 331. “*janapadaniruttiṃ nābhiniveseyya, samaññaṃ nātidhāveyyā’ti — iti kho paṇetaṃ vuttaṃ. kiñcetaṃ paṭicca vuttaṃ? kathañca, bhikkhave, janapadaniruttiyā ca abhiniveso hoti samaññāya ca atisāro? idha, bhikkhave, tadevekaccesu janapadesu ‘pā’ti sañjānanti, ‘pattan’ti sañjānanti, ‘vittan’ti sañjānanti, ‘sarāvan’ti sañjānanti ‘dhāropan’ti sañjānanti, ‘poṇan’ti sañjānanti, ‘pisīlavan’ti, sañjānanti. iti yathā yathā naṃ tesu tesu janapadesu sañjānanti tathā tathā thāmasā parāmāsā abhinivissa voharati — ‘idameva saccam, moghamaññaṃ’ti. evaṃ kho, bhikkhave, janapadaniruttiyā ca abhiniveso hoti samaññāya ca atisāro.*

332. “*kathañca, bhikkhave, janapadaniruttiyā ca anabhiniveso hoti samaññāya ca anatisāro?*

Ñāṇamoli & Bodhi (1995: 1084) concur: “One should not insist on local language, and one should not override normal usage.” Von Hinüber (1995: 190) states “The advice given in this paragraph from the *Majjhimanikāya* is clear: one should use standard language, what should be meant by *samaññā* here, and avoid dialects or colloquialisms.” Brough (1980: 40) and Norman (1980: 62), both MOTT advocates, take *janapadanirutti* as meaning local dialect, but do not address the corollary that *samaññā* therefore means standard dialect. Cone (2010: 204) takes a different tack and translates *janapadanirutti* as “local terminology”, thus implying that *samaññā* means standard terminology, which is confirmed by the commentary.<sup>50</sup>

This injunction to use *samaññā*, ‘standard vocabulary’, has been widely misread by MOTT advocates. One reason may be that the normally authoritative Lamotte (1958: 611) arrived at the opposite meaning and others have followed.<sup>51</sup>

---

*idha, bhikkhave, tadevekaccesu janapadesu ‘pātī’ti sañjānanti, ‘pattan’ti sañjānanti, ‘vittan’ti sañjānanti, ‘sarāvan’ti sañjānanti, ‘dhāropan’ti sañjānanti, ‘poṇan’ti sañjānanti, ‘pisīlavan’ti sañjānanti. iti yathā yathā naṃ tesu tesu janapadesu sañjānanti ‘idaṃ kira meāyasmanto sandhāya voharantī’ti tathā tathā voharati aparāmasaṃ. evaṃ kho, bhikkhave, janapadaniruttiyā ca anabhiniveso hoti, samaññāya ca anatisāro. ‘janapadaniruttiṃ nābhiniveseyya samaññaṃ nātīdhāveyyā’ti — iti yaṃ taṃ vuttaṃ, idametam paṭicca vuttaṃ.*

<sup>50</sup>The commentary states (DPR M-a 3.4.9.232): *samaññaṃ ti lokasamaññaṃ lokapaṇṇattim*, ‘*samaññā* is universal terminology, universal description.’. This implies a standard vocabulary. There is no comment on *janapadanirutti*.

<sup>51</sup>Levman (2014: 110) confirms that the mistake is Lamotte’s. Lamotte (1958: 611) cites *janapadaniruttiyā ca abhiniveso*, ‘affectation of the dialect of the countryside’ and *samaññāya ca atisāro*, ‘departure from recognised parlance’. Then he mistakes this description of how conflict arises for the recommended conduct. As four pages earlier Lamotte (1958: 607) had claimed “The main achievement of the sects was to have put the word of the Buddha into the vernacular”, I take this error as confirmation bias. He may have been influenced by Law (1933: xvi) who offers a paraphrase which also reverses the meaning: “Now a man of a particular locality, when he is in other localities where different names of the same thing are in vogue, knowing that in different localities different names of the same thing are used conventionally by the gentlemen, uses different names in different localities without any attachment to his own local form.” Wimalawansa & Perera (1976: 5) also misunderstand: “... it is clear that the Buddha did not pay undue attention to language. When several terms are used to describe one thing in different states or regions, one must not stick to one particular term only.” Piyasīlo’s (1996: 146) misreading is more nuanced: “... it is said the ‘middle way’ is not to insist unduly on his own provincial dialect and at the same time not to diverge from general or recognised language...for instance, a different word is used for ‘bowl’: *pāti, patta, vittha, sarāva, dhāropa, poṇa, pisīla*, and that each one considers his word the only correct one, but that in the interest of peace, it is best for each one to use the word currently at hand.”; unfortunately for Piyasīlo’s reading, the ‘middle way’ applies only to the first of the six recommendations for avoiding conflict, that of avoiding both sensuality and self-mortification.

Norman (1988: 12fn) too seems to have misunderstood and has used this very passage as part of an argument that the Buddha must have used local dialects.

This passage contradicts the MOTT claim that the Buddha and his disciples taught in local dialects. Furthermore, although it does not unambiguously refer to a standard language, it strongly implies it, as von Hinüber realised.<sup>52</sup> For it is hard to see how a standard vocabulary arises independent of a standard language. When combined with the evidence of the 18 suttas that affirm learning dhamma to the syllable (§2.2), this is strong circumstantial evidence of the use of a standard language across all dialect areas for transmission of texts.

## 5.2 The inferred characteristics of the Buddha's teaching language

If a standard form of Indo-Aryan existed in the Buddha's day, we can infer its probable characteristics: a western variety, evidenced in inscriptions and showing pre-Aśokan features. All these features are evident in Pali.

### 5.2.1 A western variety

A western variety becoming the standard in a dialect continuum spanning the north-west, west and east of India is what one might expect. For there is a tendency for the central variety to become the standard. This was the case in medieval England, as noted by John of Trevisa in 1385,<sup>53</sup> and in Germany, as

---

As these four readings claim support from the standard MOTT misreading of *sakāya niruttīyā* as authority for translation (§4), I attribute these mistakes to confirmation bias.

<sup>52</sup>However, von Hinüber dismissed this passage in favour of the MOTT misinterpretation of the *sakāya niruttīyā* passage (§4). Von Hinüber (1995: 190f) questioned the fact of multiple terms by arguing that only *patta* and *pāti* are precise synonyms and the other words refer to different objects. He gives no references and the significance of this point is not clear; he may think only synonymous pairs present difficulties. However, I believe the passage refers, not to synonyms, but to semantic divergence where the same word applies to different objects in different dialects, as in *pavement*, *subway* etc. in British and American English (§2.3). If true, compared to English, semantic divergence was considerable in MIA dialects.

<sup>53</sup>Freeborn (2006: 94) modernises John of Trevisa:

“... also concerning the Saxon tongue that is divided into three and has barely survived among a few uneducated men (there) is great wonder, for men of the east with men of the west, as it were under the same part of heaven, agree more in (their) pronunciation than men of the north with men of the south. Therefore it is that Mercians, who are men of middle England, as it were partners of the extremes, understand better the languages on either side, Northern and Southern, than Northerners and Southerners understand each other.”

(The East Midland variety became standard English in the fifteenth century.)

noted by Luther in the 16<sup>th</sup> century.<sup>54</sup> Of course there are exceptions,<sup>55</sup> but not India. Patañjali claims that the inhabitants of the Āryāvarta, a central region north of the Vindhyas, are the normative speakers of Sanskrit (Deshpande 1986: 316-7). Epigraphic Prakrit was a central-western variety (§5.2.2). Later, the normative speakers of Prakrit were from the Deccan, another central area due south. Bubenik (1996: 12) states:

Māhārāṣṭrī, according to the grammarians, was the Prākṛit par excellence, the ‘standard’ Prākṛit. While the grammarians describe its features, in the case of other Prākṛits they mention only how they deviate from the ‘standard’ Prākṛit. According to Daṇḍin (6<sup>th</sup> c. A.D.), Māhārāṣṭrī was the most ‘excellent’ Prakrit. It was based on the living tongue of the north-western part of the Deccan (along the river Godāvarī)...

Śroṇa Kuṭīkaṇṇa (Pali Soṇa Kuṭīkaṇṇa), a native of Avanti in modern Madhya Pradesh or else of Aparānta(ka) further west, was praised by the Buddha for his recitation of the whole Aṭṭhakavagga.<sup>56</sup> The Buddha also regarded him as foremost of monks with a fine speaking voice.<sup>57</sup> The Buddha definitely praises the delivery of the Aṭṭhakavagga recitation<sup>58</sup>, but his western accent and dialect were also valued,<sup>59</sup> for eastern varieties were not prestigious.

<sup>54</sup>Russ (1994: 13) translates Martin Luther:

“I haven’t any certain, special language of my own but I use the common German language so that both North and South Germans can understand me. I speak according to the Saxon Chancery, which all the princes and kings in Germany follow; all Imperial Towns, princely courts write according to the Saxon Chancery of our prince, therefore that is the most common German language.” Saxony is in the centre of Germany.

<sup>55</sup>For example, from the French Revolution onwards, the government promoted the non-central Parisian French as the standard language.

<sup>56</sup>Vin I.196, DPR Mv 5.258 *sabbāneva aṭṭhakavaggikāni sarena abhāsi* and Ud 59, DPR Ud 5.6.18 *aṭṭhakavaggikāni sabbāneva sarena abhaṇi*.

<sup>57</sup>A I.24, DPR A 1.14.2.206, AN 1.206 *etadaggaṃ, bhikkhave, mama sāvakānaṃ bhikkhūnaṃ kalyāṇavākkaraṇānaṃ yadidaṃ soṇo kuṭīkaṇṇo*. ‘Foremost, monks, of my monk disciples with fine speaking voices is Soṇa Kuṭīkaṇṇa.’

<sup>58</sup>Horner (1962: 264) translates the Pali (Vin I.196-7): “Good, it is good, monk, that by you, monk, the Divisions of Eights are well learnt, well reflected upon, well attended to, and that you are endowed with lovely speech, distinct, without hoarseness, so as to make the meaning clear.” Tatelman (2005: 89) translates the Sanskrit version “Excellent! Excellent, Shrona! Mellifluous is the Dharma you have spoken and presented.”

<sup>59</sup>‘Accent’ is possible for *sarena* and is confirmed by Lévi (1915: 404fn.): “une prononciation



Bronkhorst (2007: 7-8) points out the language of the east had low status. He cites the Śatapatha Brāhmaṇa (3.2.1.23) giving the speech of demons as *he'lavo he 'lavaḥ*; according to Patañjali, this stands for the Māgadhī *he'lavo he 'layaḥ*; in Sanskrit, it corresponds to *he'rayo he 'rayaḥ*, 'Hail Friends!' Hock (1991: 1) translates this passage: "The Asuras, deprived of (correct) speech, saying *he lavo*, *he lavaḥ*, were defeated. This is the unintelligible speech which they uttered at that time. Who speaks thus is a barbarian. Therefore a brahmin should not speak like a barbarian, for that is the speech of the Asuras."

Bronkhorst (2007: 9) follows up with the Jaiminīya Brāhmaṇa (1.337-38) which relates how a certain Brahmin called Brahmadata Caikitāneya was appointed *Purohita*, 'head-priest', by the king of the Kosalans, Brahmadata Prāsenajita, but left claiming the king's son spoke like an easterner and could not be understood.<sup>60</sup> The Brahmin's unwillingness to adapt attests to the low status of the eastern variety. We can infer that in the late Vedic period Kosala must have been a transition zone where a western variety predominated, though an eastern variety was becoming more accepted by the younger generation. From his analysis of Vedic dialects, Witzel (1989: 226) concluded

---

spéciale, un accent locale", 'a special pronunciation, a local accent' Lévi offers parallel passages from: Divyāvādāna from the Mula-Sarvāstivādin Vinaya (1915: 404) "*avec le timbre du pays d' Aparantaka*", 'in the accent of the country of Aparantaka'; from the Sarvāstivādin Vinaya (1915: 407) "*avec la prononciation du pays d' Avanti*", 'in the pronunciation of the country of Avanti'; finally, from the Mahīśāsaka Vinaya (1915: 409) with the Buddha enquiring "*Les hommes de ce royaume prononcent-ils tous ainsi?*", 'Do all the men of this kingdom pronounce in this way?' Tatelman (2005: 88-89) confirms 'accent' at Divyāvādāna 1.159 "in the accent of Ashmaparantaka, but with the proper intonation" for *asmāparāntikayā svaraguptikayā*.

However, the Theravadins appear to have a different tradition on *sarena*, taking it as 'delivery' instead of 'accent': Ud-a and Vin-t offer *suttussāraṇasarena abhāsi*, *sarabhaññavasena kathesīti attho*, 'he spoke the suttas in a high-pitched voice, the significance is because he recited in plainchant'; Vin-a makes no comment. Rhys Davids & Oldenberg (1882: 37) give "intoned"; Horner (1962: 264) translates *sarena* as "from memory", which would only be correct if reciting from a written text were the norm in the Buddha's day; Masefield (1994: 105) has "melodically", but this is specifically forbidden at Vin II.107. The commentary to A IV.63 on Velukaṇṭakī Nandamāta's recitation of the Pārāyana gives *sarena bhāsaṭṭi ... madhurena sarabhaññena bhāsaṭṭi*, '**she intoned** ... she recited in a pleasing plainchant'.

<sup>60</sup>In case the reader believes this undermines my claim of mutual intelligibility (§2.3), allow me to quote Trudgill (2010: 182): "It is well known that mutual intelligibility can increase with exposure; it can depend on the **willingness** of the parties involved to attempt to understand each other and to make themselves understood; it may not be equal in both directions; and so on."

“The Kosala land, occupied by the Kāṇvas, Baudhāyanīyas, and Śāṇḍīlyas, however, is in many ways a transitional area (usually with a strong Western influence) ...”<sup>61</sup>

The Sakyans were possibly considered rough spoken.<sup>62</sup> Certainly one brahmin regarded them, and presumably their language, as lacking prestige and a negative evaluation of the Buddha’s people in a Buddhist text must be significant. Whatever, the truth of matter, the Buddha is likely to have been educated, as a chief’s son, to speak a standard western dialect suitable for conversation with members of the Kosalan governing elite, including royalty.<sup>63</sup> The Buddha’s praise of the westerner Soṇa Kuṭṭikaṇṇa’s accent makes most sense if there existed such a standard variety.

Although eastern Aśokan inscriptions are in the court language of the east, it is uncontroversial to say those of Girnar in Kāthiāwār in the west are close to Pali. Lamotte (1958: 626) wrote:

... we can conclude that neither Māgadhī nor Ardhamagadhī constitutes the linguistic basis of Pālī, and that the cradle of the latter - if we can speak of cradle for such a composite language - is to be sought amongst the Western Prākritis, in the area of Avanti extending into Kāthiāwār.

---

<sup>61</sup>I assume this linguistic situation continued through the Buddha’s time until changed by the Mauryan empire. The bias against eastern varieties must have abated then, but returned afterwards, for eastern varieties disappeared from the inscriptional record and Māgadhī had a low status in Sanskrit drama.

<sup>62</sup>The Brahmin Ambaṭṭha calls Sakyans *pharusā* (*pharusā, bho gotama, sakyajāti* D 1.90, DPR D 1.3.264, DN 3); this could mean ‘rough-spoken’ without the suffix *-vāca*. However, the commentary gives *kharā*, ‘rough’ for *pharusā*.

<sup>63</sup>Wynne is quoted in Gombrich (2018: 82) as suggesting that the Buddha’s dialect would have been a western-type Kosalī as he was based in Sāvattthī, Kosala’s capital, midway between Delhi and Magadha. This is compatible with my argument that the Buddha would not have spoken Māgadhī (§3.1)

However, the linguistic situation of Kosala could have been complex. Given the evidence above that Kosala was a transitional area in the late Vedic period, we cannot assume the dialect of Sāvattthī was the same as that of Kapilavattu, the Buddha’s home town, some 100km east and potentially across an isogloss demarcating a more eastern variety. The Sāvattthī dialect may itself have been a mixture of eastern and western types. We cannot even be sure the Buddha was a native Indo-Aryan speaker; Levman (2013: 157) infers that he may have spoken a Munda language. Moreover, the Kosalan elite, including the Buddha, may not have spoken a local dialect, but a standard sociolect. Finally, the Buddha may have spoken both a local dialect and a standard sociolect.

This, of course, is the home region of Soṇa Kuṭikaṇṇa. What is controversial, however, is the claim here that the Buddha spoke this (non-composite) language.

### 5.2.2 *A variety evidenced in inscriptions*

Salomon (1998: 77) states:

Like the eastern dialect under Aśoka, the central-western dialect of the post-Mauryan era was used far beyond what must have been its original homeland. Thus we find inscriptions in this standard epigraphic Prakrit as far afield as Orissa in the east, for instance, in the Hāthīgumphā inscription ..., while in the south it is abundantly attested in inscriptions from such sites as Nāgārjunakoṇḍa and Amarāvati. This central-western MIA dialect was, in fact, virtually the sole language in epigraphic use in the period in question, and therefore seems, like Pāli, to have developed into something like a northern Indian lingua franca, at least for epigraphic purposes, in the last two centuries B.C. This is not to say that the inscriptions in this dialect, which Senart called “Monumental Prakrit”, are totally devoid of local variations. ... But all in all, the standard epigraphic or “Monumental” Prakrit can be treated as essentially a single language whose use spread far beyond its place of origin, and which should not be taken to represent the local vernacular of every region and period where it appears.

Not only did post-Mauryan inscriptions default to a roughly homogeneous western variety, some are remarkably similar to Pali. Therefore, Epigraphic Prakrit is taken here as a reflex of Pali. Barua (1926: 119) states of the 2<sup>nd</sup> century BCE Bharhut inscriptions in Madhya Pradesh “Barring the provincialisms, the language of the Bharhut railing can be regarded as a Pāli dialect.” This is hardly surprising in a Buddhist site. However, in Orissa in the Hāthīgumphā cave, there is a completely secular inscription of the 1<sup>st</sup> century BCE recording the rule of King Khāravela, who claims to respect all sects. Barua (1929: 157) says of it: “Leaving the spelling and pronunciation of a few words out of consideration, we can say that their language is Pāli, and nothing but Pāli.” Norman (1993: 87) concurs: “There is, in fact, very little difference between Pāli, shorn of its Māgadhisms and Sanskritisms, and the language of the Hāthīgumphā inscription.” Cousins (2013: 124-7) described Epigraphic Prakrit as ‘Old Pāli’, and of the much later inscriptions found at Devnimori in Gujarat (ca. 400CE) and Ratnagiri in Odisha (ca. 500 CE) von Hinüber (1985a: 197) states “... the language, although closer to Pāli than any

other surviving Middle Indic, is by no means identical with it. Thus the inscriptions should be classified linguistically as two new varieties of continental Pāli...”

This sketch of the geographical scope, non-sectarian character and longevity of Epigraphic Prakrit on the Indian mainland is offered as evidence that it had the status of a standard language. The similarities to Epigraphic Prakrit suggest that Pali also had the same status.<sup>64</sup>

### 5.2.3 *A variety predating the Aśokan inscriptions*

The antiquity of Pali is suggested by features paralleled in Vedic and absent from classical Sanskrit. Oberlies (2001: 7-14) offers:

1. 71 words found in Pali and Vedic but missing from classical Sanskrit;
2. Pali and Vedic use the aorist as the standard preterite, but Sanskrit does not;
3. Pali *-a* stem m. nom. pl. endings *-āse/-āso*, n. nom. pl. *-ā* and m., n. instr. pl. *-ehi* and *-ī* stem acc. sg. *-iyam* parallel Vedic *-āsaḥ*, *-ā*, *-ebhiḥ*, *-yam* and contrast with Sanskrit *-āḥ*, *-āni*, *aiḥ*, *-īm*;
4. the Pali dat./gen. sg. of the personal pronouns and the loc. sg. of *ta(d)* without *anusvāra*, *mayha*, *tuyha* and *tamhi*, parallel R̥gvedic *māhya*, *túbhya* and *yásmi/ sásmi* and contrast with Sanskrit *mahyam*, *tubhyam* and *tasmin*;
5. Pali and Vedic preserve suffixes missing from Sanskrit: the infinitive in *-tave*, the participle in *-āvi(n)*, the suffix *-ttana* forming abstract nouns;
6. 6 words in Pali and Vedic, but with a meaning lost by classical Sanskrit; e.g. *senā* ‘missile’ or the distinction between Pali *kasati*, ‘ploughs’, and *kassati*, ‘drags’ (Vedic *kṛṣāti* and *kárṣati*) fused into Sanskrit *karṣati*, ‘plough/ drag’.

---

<sup>64</sup>The silence of Indian grammarians on Epigraphic Prakrit and Pali gives the opposite impression that this variety was a marginal language. Von Hinüber (1985b: 72) explains the silence as Pali being subsumed under *Paiśacī*. It is also possible that Sanskrit grammars were ignoring the competition for the H-language by linking other Prakrits to Sanskrit.

The Vedic and Pali *-tvāna* absolutive is another example absent from classical Sanskrit.

Pali has early features which are not found in inscriptions.<sup>65</sup> In some cases, they are also clearly older than classical Sanskrit or Ardha-Māgadhī.

1. Barua (1929: 158) noted Pali is close to Vedic in retaining *l* rather than adopting Sanskrit *ḍ*. However, the Hāthīgumphā inscriptions (H) follow Sanskrit (Skt) and Ardha-Māgadhī (AMg). Pali *kīlikā*, ‘sport’, (*krīla* in Vedic) becomes *krīḍā* (Skt), *kīḍiyā* (AMg) and *kīḍikā* (H). Pali *kīlitā*, ‘played’, (*krīlitā* Vedic) becomes *krīḍitā* (Skt), *kiḍḍā* (AMg) and *kīḍitā* (H). Pali *pīlā*, ‘pain’, (Vedic *pipīle*, ‘was pressed’) becomes *pīḍayati* (Skt), *pīḍā* (AMg) and *pīḍāpayati*, ‘oppress’, (H). Pali *talāka*, ‘reservoir’, becomes *taḍāka* (Skt), and *taḍāga* (Skt, AMg, H). Pali *veḷuriya*, ‘beryl’, becomes *vaiḍūrya* (Skt) and *veḍuriya* (AMg, H). Pali *kaḷāra* ‘tawny’ becomes *kaḍāra* (Skt, AMg, H).<sup>66</sup> Salomon (1998: 35-6) states of the Brāhmī script used for Tamil inscriptions of 2<sup>nd</sup> century BCE at the earliest “...it has four entirely new characters, interpreted as na, ra, ṛa and ḷa, which were evidently created in order to represent Dravidian phonemes not represented in standard (northern) Brāhmī.”. It appears that ḷa had not been needed in Epigraphic Prakrit because it had already developed into ḍa. For example, in the Aśokan Pillar Edict V, Pali *eḷaka*, ‘ram’, and *eḷakā*, ‘ewe’, become *eḍaka* and *eḍakā* (Skt *eḍaka*, *eḍakā*).<sup>67</sup>

<sup>65</sup> There is no database of Indian inscriptions to check with and these observations are based on a small sample.

<sup>66</sup> Pischel §240 states that in the literary Prakrits *ḍ* is generally *l*. This is the exact opposite of what I propose for Epigraphic Prakrit. However, I am speaking of historical sound changes proven by *l* vanishing, but *ḍ* remaining in modern Hindi, Urdu, Bhopuri, Magahi, Maithili, Nepali, Sindhi and Kashmiri (Cardona & Jain 2003). Pischel (§226) is referring to spelling conventions of North India which were broadened to the extent that *ṛa*, *ḍa* or *na* could become *la* or *ḷa*. These conventions were unlikely to apply to the Sri Lankan Pali tradition; modern Sinhala retains *ṭ*, *ḍ*, *n*, *l* and *ḷ*, as does Oriya, Panjabi, Gujarati, Marathi and Konkani. Cardona & Jain (2003) do not present a modern Indo-Aryan language that retains *l* but loses *ḍ*.

<sup>67</sup> However, Hultzsch (1925: 225) reads *eḷakā* at Toprā and Rāmpūrvā PE V C (*ḍ* at other sites) and *eḷake* at every site but Toprā PE V J, which has *ḍ*. He also reads *duḍī* ‘tortoise’ (Skt *duḍi/ḍuli*) for PE V B for Allahabad, but disregards Buhler’s *daḍī* for Toprā and offers *daḷī* instead along with *duḷī* for three other sites; there is no Pali equivalent. For Pali *pañnarasa* ‘fifteenth’ Hultzsch reads

2. Oberlies (2001: 99) states that in Pali *-vv-* as a result of *-vy-*, *-vr-* or *-rv-* is represented by *-bb-* in medial position.<sup>68</sup> The *-rv-* cluster has a unique development to *-bb-* in Pali and later to *-vv-*. Vedic, Sanskrit and Western Aśokan conserve *sarva*, ‘all’, Pali has *sabba*, which develops to Eastern Aśokan, Bharhut, Sañci and Hāthīgumphā *sava*<sup>69</sup> and Ardha-Māgadhī *savva* (with a separate development for North-western Aśokan *savra*). Similarly, Vedic and Sanskrit conserve *parvata*, ‘mountain’, Pali has *pabbata*, which develops into Eastern Aśokan, Bharhut and Hāthīgumphā *pavata* and Ardha-Māgadhī *pavvata*. The *-vy-* cluster has a unique development to *-bb-* in Pali and later to *-vv-* or *-viy-*. For the gerundive, Sanskrit and Western Aśokan conserve *-tavya*, Pali has *-tabba*, North-western Aśokan *-tave*, Ardha-Māgadhī *-yavva* and Eastern Aśokan *-taviye*. As Pischel §201 gives medial *-b-* changing to *-v-* in literary Prakrits, we can infer *-bb-* in Pali became *-vv-* in Epigraphic Prakrit and its absence from inscriptions makes *-bb-* pre-Aśokan.<sup>70</sup>

---

*paṃnaḷasaṃ* at Ararāj and Nandangarh PE V H, but *paṃnaḷasaye* at every site but Nandangarh PE V J. In each case, Hultsch is substituting *!* for Buhler’s *ḍ* on the basis of Buhler’s admission of some uncertainty on interpreting the shapes of worn characters and Lüders’ certainty that *!* was correct here; Lüders (1911: 1089) thought *!* was more widely used in Epigraphic Prakrit and Sanskrit than generally supposed and etymologies needed to be revised accordingly.

In conclusion, Hultsch and Lüders are contradicted by Salomon’s claim that early Brāhmī script had no *!*, and I defer to Salomon as he is more up-to-date; however, if a future critical edition of the Aśokan inscriptions should vindicate Hultsch, I would modify my argument to say that the change from *!* to *ḍ* was in progress at the time of Aśoka though, as *!* predominates in Pali, Pali has pre-Aśokan characteristics.

<sup>68</sup>By accepting this, I do not wish to imply that Pali is derived from Vedic or Sanskrit. I subscribe to the model following Wackernagel in Hock & Pandharipande (1978: 13) of parallel developments of Vedic and Prakrit with a pre-Vedic common ancestor.

<sup>69</sup>These are written in the Brāhmī script, which gives single letters for double consonants.

<sup>70</sup>The initial *by-* cluster is unique to Pali. Pali frequently also has *vy-* for the same word, but this is the sole form in other varieties. Pali *byañjana/ vyañjana* ‘syllable/ letter’, has equivalents *vyañjana* (Vedic and Sanskrit), *viyañjana* (Aś Sarnath) and *vayañjana* (Aś Rūpnāth). Pali *byatta/ vyatta* ‘clever’, equates to *vyakta* (Vedic and Sanskrit) and *viyata* (Aś Pillar Edicts). Pali *byappatha/ vyāvaṭa*, ‘busy’, equates to *vyāprta* (Sanskrit), *vyāpata* (Western Aśokan) and *viyapaṭa* (North-western and Eastern Aśokan).

It could be argued that the difference between *by-* and *vy-* is regional: Burmese and Thai manuscripts prefer the former, whereas Sinhalese manuscripts prefer the latter; therefore, while

3. The absolute ending *-tvā* is usually considered a Sanskritisation, on the dubious assumption that the Buddha used (Ardha-)Māgadhī *-(i)ttā*.<sup>71</sup> If, however, one acknowledges the Theravada tradition that Pali dates from the Buddha and the evidence above that Pali is in some respects closer to Vedic

---

*-bb-* is used across all regions, it might merely represent a consensus on orthography and not an original sound.

Against this argument are the following observations:

To my knowledge, the Burmese and Thai editions only have *by-*; this may be related to the lack of [v] or [v̄] in the Burmese or Thai phonetic inventory.

Sinhalese has both [b] and [v] and Sinhalese editions, as given by GRETEL, <http://gretel.sub.uni-goettingen.de/>, show the following distributions of *by-*/*vy-*: D I, 122/ 43; D II, 83/ 6; D III, 146/ 34; M I, 352/ 58; M II, 129/ 28; M III, 54/ 60; S I, 13/ 7; S II, 156/ 58; S III, 6/ 5; S IV, 7/ 276; S V 53/ 155; A I 32/ 130; A II 79/ 78; A III 68/ 184; A IV 21/ 92; A V 552/ 126. Only five examples, M III, S IV, S V, A III and A IV, show a preference for *vy-*, whereas the other eleven examples prefer *by-*. I claim that these results do not show any levelling or evidence of a later redaction, unlike the Thai and Burmese texts.

The argument for regional orthographic variation implies the doubly implausible scenario of (a) one orthographic convention, *v*, for inscriptions and a different convention, *b*, for manuscripts and of (b) an exemption from levelling in Sinhalese manuscripts for *by-/vy-*, but a complete levelling of *-vv-/-bb-* to *-bb-*.

The absence of *by-* from the inscriptional record suggests it is pre-Aśokan and confirms *-bb-* as also pre-Aśokan. The scenario envisaged here is that *b* is a genuine phonetic variant of *v*, well entrenched in the Buddha's day in medial position, but influenced in the *by-* cluster by contact with other varieties to become *vy-* at times. Later, /b/ had changed to /v/ in Epigraphic Prakrit (and in literary prakrits), but was conserved as *-bb-* and *by-* in the Sinhalese Pali tradition along with the *vy-* alternative, which would be pronounced [vj] or [v̄j]. The *vy-* alternative was levelled to *by-* in the Burmese and Thai Pali traditions, possibly because it was suspected of being a Sanskritism, possibly because [v̄] and [v̄j] are not in their phonetic inventories, *v* being pronounced [w], and because [w̄j] is difficult to pronounce.

<sup>71</sup>Norman (1997: 78) states "Writing down would have been an excellent opportunity for the homogenisation of forms – all absolutes in *-ttā* being changed to *-tvā* ...." However, Oberlies (2001: 265) gives eight endings for the Pali absolute, *-(i)tvā*, *-(i)tvāna(m)*, *-(i)tu(m)*, *-tūna*, *-(i)yā* (*-ccā*), *-(i)yāna(m)*, *-eyya* - and *am*. He does not include *-ttā*. Against Norman's view, it is unlikely that only the alleged *-ttā* ending would be completely 'Sanskritised', while other absolute forms remain unedited.

Von Hinüber (1982: 136) argued for an absolute in *-ttā* by comparing the agent noun *chettā* at Sn 343 in the PTS edition with the Burmese edition absolute *chetvā*; he suggested that the agent noun was mistaken for the alleged *-ttā* absolute and was edited in the Burmese edition; however, he has no definitive examples of *-ttā* as an absolute.

Norman and von Hinüber sought evidence of *-ttā* in Pali on the hypothesis that the Buddha spoke a kind of Māgadhī. I do not think that they succeeded.

than Sanskrit (including the Vedic and Pali *-tvāna* absolutive missing in Sanskrit), we can hypothesise that this absolutive is an original feature. Then we have a simple trajectory of the development from *-(i)tvāna* (Vedic and Pali) > *-(i)tvā* (Vedic, Sanskrit and Pali) > *-(i)ttā* (Ardha-Māgadhī) > *-(i)tā* (Hāthīgumphā) with separate developments for *-(i)tu* (Eastern and North-western Aśokan) and *-(i)tpā* (Western Aśokan). On the basis of Occam’s Razor, as advocated by Hock (1991: 538), taking *-tvā* as an original feature is preferable to the redundant Sanskritisation narrative and thus supports the Theravada tradition.

4. *Brāhmaṇa* ‘brahmin’ is also usually considered to be a Sanskritisation, but by taking it as a 5<sup>th</sup> century BCE feature we have a simpler and therefore a preferable development from *brāhmaṇa* (Vedic, Pali, and Sanskrit) > *brahmaṇa* (Western Aśokan) > *bram(h)ana* (Bharhut) > *bamhaṇa* (Hāthīgumphā) with separate developments, *bramaṇa* (North-western Aśokan), *baṃmaṇa* (Ardha-Māgadhī) and *baṃbhana* (Eastern Aśokan). Almost certainly brahmins called themselves *brāhmaṇāḥ* in every area of India, and this is likely to be a loanword in Pali before being naturalised in other varieties by the sound changes recorded in inscriptions.<sup>72</sup>

---

<sup>72</sup>There are two problems with *brāhmaṇa* in Pali. (1) Oberlies (2001: 93) states that in word initial position in Pali only single consonants are allowed. This is certainly a strong tendency in Pali, but not a rule: the *br-* cluster is repeated in all derivatives of *brahman* and also in *brahant*, *brūti* and *brūheti* and their derivatives; initial clusters of, *ty-*, *tv-*, *dr-*, *dv-*, *by-*, *vy-* and *sv-* are also found in Pali dictionaries. (2) The long first vowel contravenes the Law of Morae, though Geiger, (1916 §7, 8) notes exceptions and the ancients were not always strict about this ‘law’ (§2.4). Oberlies (2001: 107), relying on Saksena, states “*brāhmaṇa-* is a Sanskritism and does not comply with any Pali sound law; its ‘etymologies’ (e.g. *bāhitapāpo ti brāhmaṇo*, Dh. 388) show that it was pronounced as *b(r)āhaṇa-*.” The Dhammapada abounds in polysemy and wordplay, so its ‘etymology’ may say more about a word play than its actual pronunciation; I cannot find an attested example of *b(r)āhaṇa*. Pace Saksena, one would not necessarily expect a borrowing from another variety to reflect the host language’s sound conventions. In English, *rendezvous* is frequently pronounced as in French with a nasalised first vowel although this feature is not part of English phonology. The French *hors d’oeuvres*, ‘starters’, is pronounced /ɔː də:v(ɾə)/ in Britain, approximating the French, although in America it is naturalised to /ɔːr də:vz/. Similarly, *brāhmaṇa* could be considered either as an original anomalous feature or as a loanword which



The MOTT claims that Pali represented a translation from Māgadhī, as exemplified in the Eastern Aśokan inscriptions, into a western literary variety. If so, it would be most unlikely to have features pre-dating Classical Sanskrit, Ardha-Māgadhī, the Aśokan inscriptions or Epigraphic Prakrit. If, however, Pali is a natural language of the 5<sup>th</sup> century BCE, one would expect a few of its features to evolve in its reflex, Epigraphic Prakrit, and in other varieties. This is what is actually found.<sup>73</sup>

### 5.3 Linguistic conditions supporting a standard language

In §2.3, there were arguments for the mutual intelligibility of Indo-Aryan in the Buddha’s day and beyond. I now argue that the standard vocabulary recommended by the Buddha (§5.1) would have strengthened mutual intelligibility among the dialects. For, with standard vocabulary, their differences were overwhelmingly of accent, with very few of syntax and of morphology.

First, we look at a controversy. Regarding the Aśokan Rock Edicts, Cousins (2013: 97) seems to agree on mutual intelligibility: “...the North-Western forms of Middle Indian in the early Mauryan period were certainly relatively close to the dialects spoken on the Gangetic plain.” However, Levman (2016: 2), speaking first of the Buddha’s time, disagrees:

---

was naturalised in later varieties.

<sup>73</sup>Lamotte (1958: 627-8) came to the opposite conclusion: that Pali post-dates the Aśokan inscriptions on the basis of archaisms found in the north-west, e.g. three sibilants compared to one in Pali, *dhrama* (Pali *dhmma*) and *draśana* (Pali *dassana*), showing *r* retention; he also found archaisms in the west: e.g. long vowels before consonant clusters, e.g. *ātpa* (Pali *atta*), and unassimilated clusters such as gerundives in *-tavya*. His reasoning is fallacious because: (1) he would not have found these archaisms had he made the comparisons with the dominant eastern Aśokan variety; (2) language varieties do not change at the same rate in the same features, Thus typically any variety compared to another will have both more archaic and fewer different archaic features. Lamotte (1957: 626) acknowledges that Pali has some archaic features closer to Vedic than Sanskrit, but these are apparently ignored when coming to his final conclusion that Pali is less archaic; (3) the retention of *r* in *dhrama* and *draśana* continued into the Common Era in the north-western variety’s reflex, Gāndhārī. On Lamotte’s reasoning we could wrongly deduce that Pali post-dates the beginning of the Common Era; arguably, these are not archaisms, but later developments by metathesis from Sanskrit *dharma* and *darśana*; (4) to make a valid relative dating, the **same** features need to be compared across **all** available varieties to identify trends in the majority of varieties, as above in the bullet points of §5.2.3, but Lamotte selected different features in each variety and compared only two varieties at a time. Lamotte also refers to a six-member compound *aniccucchādanaparimaddanabhedanavidamsanadhammo* (D I.76) ‘an impermanent thing, subject to erosion, abrasion, dissolution and disintegration’ and correctly points out such long compounds are not characteristic of early Indo-Aryan texts; however, six member compounds are untypical of Pali and suggest later editing.

We know, for example, that the north-western dialect (Gāndhārī) was quite different from the eastern dialects, and it was unlikely that they were mutually comprehensible. Certainly by Aśoka's time the dialect differences between, for example, the dialect of Shābāzgarhī in the north and Kālsī in the east were considerable.

Levman is not alone, as Gombrich (1988: 128) opines:

The Mauryan Empire was a political unit of a new order of magnitude in India, the first, for example, in which there were speakers of Indo-Aryan languages ... so far apart that their dialects must have been mutually incomprehensible.

Levman and Gombrich offer no evidence and, to argue the contrary, an excursion into linguistics is offered, exemplifying what Bechert (1991: 6) advocated: "... we should make use of the results of research into related developments outside India." We start with this comparison from Fennell (2001: 92) who is demonstrating the influence of Old Norse on English:

*The children are playing in the street* (standard Present Day English)

*The bairns are lakin out on't street* (Modern Humberside dialect)

*Barnen leker ute på gatan* (Modern Swedish).

To an English native speaker, the last sentence is unintelligible and the middle sentence is understandable only if the local vocabulary (*bairn, lak out*) of the Humberside dialect is understood. The different phonology (final /n/ instead of /ŋ/ in *lakin* and reducing *the* to /t/) presents no problem to a native English speaker.

Therefore, the criteria suggested for mutual intelligibility are: a difference of phonology alone, i.e. of accent, does not normally create a barrier to comprehension; a difference of vocabulary will be a problem until it is learnt. However, differences of grammar, will pose a barrier that mere exposure to the variety cannot resolve, as in the Swedish example.<sup>74</sup>

---

<sup>74</sup>Even a pidgin based on English vocabulary, such as Bislama from Vanuatu in the West Pacific, is unintelligible to a native English speaker, so crucial is the role of grammar. Freeborn (2006: 420ff) offers parallel versions of the Bible, Matthew 26:73:

**Bislama** (Gud Nyus Bilong Jisas Krai): "Gogo smol taem nomo, ol man ia we oli stap stanap long ples ia, oli kam long Pita, oli talem long em, oli se, 'Be i tru ia, yu yu wan long olgeta. Yu

### 5.3.1 Underlying similarities in Aśokan inscriptions

The opening of the first Aśokan (Aś) rock edict (RE1) illustrates Levman’s comparison:

|   |  |
|---|--|
| <p><b>Kālsī</b> (AśK), India, near Dehra Dun. (Hultzsch 1925: 27): (A) <i>Iyaṃ dhammalipi Devānaṃpiyenā Piyadasinā lekhitā</i>. (B) <i>hidā nā kichhi jive ālabhitu pajohitaviye</i>.</p> | <p><b>Shābāzgarhī</b> (AśSh), Pakistan, near Peshawar. (Hultzsch 1925: 51): (A) <i>Aya dhramadipi Devanapriasa raño likhapitu</i>. (B) <i>hida no kichi jive arabhitu prayuhotave</i>.</p> |
| <p>‘(A) This law edict is inscribed by Piyadasi, Beloved of Gods. (B) Here no life whatever is to be taken and sacrificed.’</p>   | <p>‘(A) This law edict of the King, Beloved of Gods, is inscribed by order. (B) Here no life whatever is to be taken and sacrificed.’</p>  |

Although the wording of the two sets is not identical in all cases, it is striking that there are arguably no differences in syntax in either set of inscriptions.<sup>75</sup> In terms of morphology, the differences that could potentially affect comprehension are similarly very small, 1-2%<sup>76</sup>, and so we can rule out the unintelligibility

---

luk, tok bilong yu itok bilong man Galili ia.”

**English** (New English Bible): “Shortly afterwards the bystanders came up and said to Peter, ‘Surely you are another of them; your accent gives you away?’”

<sup>75</sup>Twelve differences are noted using Hultzsch’s edition: RE1(A) instrumental with non-causative v genitive with causative; RE1(G), (H) *Tiṃni/ tini pānāni v trayo praṇa* - neuter v masculine; RE111(C) *etāye v etisa* – genitive v dative; REIII(C) *aṭhāye v karaṇa* - dative v ablative; REIII(E) *hetuvatā v hetuto* – instrumental v ablative; REIV(B) *dasanā v draṣanaṃ* – plural v singular; REVI(H) *uṭhānasā v uṭhānasi* genitive v locative; REVIII(C) *nikhamithā v nikrami* - imperfect v aorist; REIX(G) etc. *dāsa-bhaṭakasi v dāsa-bhaṭakasa* – locative v genitive; REXI(A) *hā v hahati* – aorist v perfect; REXI(C) *dāśa-bhaṭakaśi v dasa-bhaṭakanam* – locative singular v genitive plural. Ten of these differences are merely stylistic and the two instances from REI(G) and (H), masculine v neuter, *tini pānāni v trayo praṇa*, may not be a mistake in the other dialect, as Pali has both *pāṇāni* and *pāṇā*. These are a mere twelve out of 134 sentences in AśK and 137 in AśSh, 133 of which are comparable. Thus, the differences in syntax are minimal and probably non-existent.

<sup>76</sup>Whether a variation is phonological or morphological is sometimes moot: Beal (2010: 10) states that article reduction from *the* to *t’* could be seen as either (cf. the Humberside example from Fennell); Bloomfield and Edgerton (1932: 15) acknowledge that what they treat as phonetic variation is often accompanied by or results in lexical or morphological change; at AśK and AśSh, two instances, REVI(E) *hakaṃ v ahaṃ* and REX(B) *dhatāye v etāye*, appear to defy classification into morphological or phonological or lexical. Following Hultzsch (1925: lxxxvii), in this paper variant inflections involving a single phonetic change are taken as phonetic variants; for example,

of unfamiliar grammar paralleled for English speakers by Swedish. Crucially, there is hardly any lexical difference, only four lexemes in AśSh representing about 1% of vocabulary used.<sup>77</sup>

### 5.3.2 Mutual intelligibility despite different accents

Despite great uniformity in other features, there is considerable variation in phonology, i.e. in accent. While differences of accent are often initially difficult to

---

with nominative masculine singular *-a* declension endings in *-a*, *-e* or *-o*, my reasoning is: (1) There are many obvious phonetic variations in the inscriptions and a native Aryan speaker would have been more likely to process variant inflection as yet another variant pronunciation. (2) At both sites, all these variants are found (Hultzsich 1925: lxxvi, xc), which suggests that they were processed by native speakers as morphophonemes, i.e. interchangeable pronunciations of the same nominative masculine singular morpheme.

Some morphological variants in both varieties are found in a single language, Pali: at REII(A) etc. *-asi* (AśK) and *-e* (AśSh) are both locative singular endings. The same goes for the causative variants at REIII(E) etc. *anapayisaṃti v anapeśaṃti*; the standard AśK causative stem in *-aya-* is also found in AśSh at REVI(F) etc. in *anapayami* instead of the usual AśSh *-e-* stem; Pali also shows both forms in many verbs. Moreover, *-asi*, the standard AśK ending, is also found in the other variety, AśSh, at REI(F), etc. (The AśSh variant of *-aspi* for *-asi* is regarded here as a phonological variation.) Therefore, these variants are likely to be mutually comprehensible.

One objective measure of morphological change is a difference in the number of syllables. On this basis, the remaining morphological variants which may cause initial difficulty in comprehension are: REI(G), XIV(A) *yevā v vo*; REII(A), (E), (F), II(A) (twice), 4(B), (F), VIII(F), XIII(A) *lājine v raṇo*; REIII(B), (K) *duvādasa v badaśa*; REIV(K) *lājina v raṇa*; REV(D) *mamayā v maya*; REVIII(B) *husu v abhuvasu*; REIX(H) *suvāmikena v spamikena*; REXII(G) *kalata v karamino*; REXII(J) *kayānāgā v kalaṇaḡama*; REXIII(H) *evā v vo*. This amounts to 20 changes that might be temporarily unintelligible in an estimated 10 lexemes per sentence and 133 comparable sentences, which equals 1330 tokens, between one and two percent of vocabulary.

(An exception to the number of syllables criterion is *-iy-*, as in REI(B) *pajohitaviye* versus *-v-* as in *prayuhotave*; this is *-vy-* in Girnar, as in *prajūhitavyam*; AśK inserts an epenthetic vowel to produce *-viy-*, whereas AśSh reduces to *-v-*. This change occurs in nouns as well as gerundives; at RE13(H) there is the noun *vyasanam* in Girnar, *viyaśanam* in AśK and *vasanam* in AśSh. Therefore the missing syllable *-iy-* in AśSh is a phonological not morphological change.)

<sup>77</sup>Hultzsich (1925: xlii) observes that AśSh *dipi* instead of *lipi* is found in Persian inscriptions and considers AśSh *nipesapita* (RE14(A)) as an ancient Persian loan word synonymous with *likhapita*. Norman (1970: 123-4) includes them in his list of independently verified dialect words, all of which come from AśSh, and adds *spasu* for *bhaginī* and *kupa* for *udupāna*. Thus the number of identified dialect words at AśSh is a mere 4. *Dipi* occurs 6 times, *nipista* 4, *nipesita* 1, *nipesapita* 1, *spasu* 1, *kupa* 1, a total of 14 occurrences. At an estimated 10 lexemes per sentence and 133 sentences, they approximate to 1% of the vocabulary used. (There are also what I infer are transmission errors, which are not counted, e.g. REIV(B) AśK *agi-kamdhāni*, ‘masses of fire’ v AśSh *joti-kamdhani*, ‘masses of light’.)

comprehend (and this may be all that Levman and Gombrich meant), they become comprehensible to a native speaker with sufficient exposure.<sup>78</sup> Native speakers would realise in the case of Sanskrit *dharma* (Pali *dhamma*) that in AśK *dhamma* *r* was assimilated, but in AśSh *dhrama* *r* was moved before the vowel. (In Pali there is a similar metathesis of *r* which realises Sanskrit *arya* as *ariya* or *ayira*). There is a similar pattern of *r* retention, dropping and metathesis in English:

| Standard spelling | General American   | Standard British                                       | Non-standard   |
|-------------------|--|--|--|
|                   | (Sanskrit analogy with <i>r</i> retention, <i>dharma</i> ) | (Kālsī analogy with <i>r</i> dropping, <i>dhamma</i> ) | (Shābāzgarhī analogy with <i>r</i> metathesis, <i>dhrama</i> ) |
| Perhaps           | /pərhaps/  | /pəhaps/   | /prəhaps/  |
| Performance       | /pərfɔːrməns/  | /pəfɔːməns/  | /prəfɔː(r)məns/  |

Until the 17th century, the *r* was pronounced throughout England as it is in modern General American English (Sanskrit analogy), but afterwards the *r* was dropped in southern British English. Non-standard pronunciations on both sides of the Atlantic of *\*prehaps* for the standard *perhaps* and *\*preformance*<sup>79</sup> for *performance* are analogous to Shābāzgarhī.<sup>80</sup>

It is cumbersome to describe this sound change in English, but a native speaker of any variety of English can process these differences within fractions of a second quite unconsciously and without loss of intelligibility. For the non-native speaker, however, such changes could be glaringly disconcerting. Similarly, the differences in the Aśokan inscriptions may have been unnoticed or irrelevant to native speakers with good exposure to different accents, while appearing highly significant to a modern scholar. Another asymmetry

<sup>78</sup>By sufficient exposure, I mean staying with a family which speaks the unfamiliar variety for a week. It would not be possible for an English speaker to learn Swedish in this time.

<sup>79</sup>For example, Paul Gambaccini, an authority on popular music, can be heard on the BBC using this pronunciation consistently.

<sup>80</sup>The comparison is not perfect, as the AśK accent compensates for *r* dropping by inserting *m*, whereas the British accent compensates by lengthening the preceding vowel if stressed; thus *car*, *bird*, *deer*, *care* and *poor* are (American first): /kɑr>kɑː/, bɜrd>bɜːd, dɪr>dɪə, keɪ>keə, pʊr>pʊə/, while the unstressed /pə/ in /pə'haps/and /pə'fɔːməns/remains unchanged.

between native Indo-Aryan speakers and modern readers, is the optical illusion of difference caused by broad and narrow transcriptions. English uses broad transcription where a single spelling of *exit* will stand for several pronunciations: /'eksɪt/, /'ɛgzɪt/, /'ɛgzɪʔ/ etc.; the position is reversed in the Aśokan inscriptions, and a single word, *dharma*, is spelt *dhrama* or *dhaṃma* to reflect its local pronunciation. Thus, Aśokan inscriptions can deceptively appear to have greater variation than modern languages with standardised spellings.<sup>81</sup>

An explanation for different accents in the Rock Edicts runs as follows: messengers memorised the edict in court language and recited it to the recipients who in turn memorised the edict, but inevitably coloured it with their own accent before it was written down or inscribed. Reliance on memory was the norm: Megasthenes, the Greek ambassador to the court of Aśoka's grandfather, remarked on the absence of written laws, and Salomon (1998: 7-8) argues that writing had a lesser status compared to the spoken word in much of Indian discourse. Reliance on memory explains the different wording and even omission of sentences in some versions of the Rock Edicts. (The uniform wording and language of the Pillar Edicts suggest lessons were learnt and written messages were used.) Therefore, the linguistic variation evident in the Rock Edicts would not significantly obstruct mutual intelligibility.

### 5.3.3 Implications for the Buddha's teaching language

The difference between the AśK and AśSh varieties is therefore overwhelmingly one of accent, though some embryonic dialect features are also evident in lexis and morphology. Different accents of the same language are generally mutually comprehensible.<sup>82</sup> We are now able to answer Norman (1980: 70), who objects:

---

<sup>81</sup> To illustrate the claim that narrow transcriptions made accents look more divergent than they sound to a native speaker, here are transcriptions from Trudgill (2000: 68) of *very few cars made it up the long hill* in regional accents:

NortheastVeree: few cahs mehd it oop the long hill, /veri: fju: ka:z me:d it ʊp ðə lɒŋ hɪl/

NW Midlands: Veri few cahs mayd it oop the longg ill, /veri fju: ka:z meɪd ɪt ʊp ðə lɒŋg ɪl/

Central Southwest: Veree few carrs mayd it up the long iooll, /veri: fju: ka:rz meɪd ɪt ʌp ðə lɒŋ ɪoʊl/

None of these differences interfere with intelligibility to a native speaker.

<sup>82</sup> A native English speaker challenged this claim from personal experience of being able to understand Krio, an English-based creole of Sierra Leone, only in written form, and not when spoken. I believe the different syntax, morphology and lexis of Krio make it difficult to separate out the beginnings and endings of words when spoken; on the page, the work of separation is done. However, with Indo-Aryan dialects, the differences of syntax, morphology and lexis have

Geiger suggested that the Buddha spoke a lingua franca, free from the most obtrusive dialect features, but it is hard to imagine what this could be, for while *ś* and *l* might well be obtrusive in an area where *s* and *r* were normal, the opposite would be the case in a region where a dialect with *ś* and *l* were spoken. It seems much more likely that the Buddha varied his language to suit his audience.

Firstly, these changes seem trivial and unobtrusive. It is plausible that these variations were widely acceptable and unobtrusive because they are merely differences of accent. Secondly, if these differences affected the meaning, the context would make matters clear. In English, a Chinese waiter referring to “fled lice” will be understood from the context as referring to ‘rice’ and not to a meal of insects.

Thus, the linguistic situation was not comparable to that of English and Swedish, but closer to standard English and Humberside. The Buddhist Sangha conversing with each other would merely have toned down their dialects, removing dialect words as required by the Buddha (§5.1). This would produce varieties which were essentially different accents with some minor morphological variation. The equivalent for Humbersiders would be to say *the children are playing on 't street* instead of *the bairns are lakin out on 't street*, eliminating the dialect words but retaining much of their local pronunciation and morphology. As explained in the following three sections, for recitation of texts this degree of variation would be further reduced to the Indo-Aryan equivalent of standard English, *the children are playing on the street*.

#### **5.4 Social conditions supporting a standard language and accent**

Rhys Davids (1903: 147), an early SOTT advocate, described how the Wanderers, *paribbājakā*, were able to converse with each other:

...the Wanderers talked in a language common among the cultured laity (officials, nobles, merchants and others), which bore to the local dialects much the same relation as the English of London, in Shakespeare’s time, bore to the various dialects spoken in Somersetshire, Yorkshire, and Essex.

---

been shown to be insignificant (§5.3.2).

The London and Home counties dialect of Shakespeare's time was also used by the educated nationwide.<sup>83</sup> Originally an East Midlands, i.e. central, dialect that had spread south and had reached London some time after the Black Death of 1348, it also became a national sociolect of the educated two centuries later. A sociolect can be spoken in a standard accent or a local accent. Trudgill (2000: 2-3) states that 'BBC English' is Standard English spoken in an accent called Received Pronunciation and is spoken by 3-5% of the population of England; however, another 7-12% speak Standard English in a regional accent. Standard English is understood everywhere in the country and is used in some homes. Similarly, a standard Indo-Aryan could have originated as a prestigious geographical dialect in the west, but also have become within several generations a sociolect of the educated, and not be restricted to formal situations. The Buddha's accent apparently differed from *Soṇa Kuṭṭikaṇṇa*'s, but they could both be speaking standard Indo-Aryan.

However, there was no single pre-eminent political and commercial centre in the Buddha's day equivalent to medieval London to promulgate a standard language. The drivers towards a standard language in 5th century BCE North India would have been at least three. Firstly, accommodation to other dialects over centuries would accomplish dialect-levelling and the emergence of regional standard languages, one of which could become a supra-regional sociolect. Crystal (2005: 243-8) gives a parallel in the history of English: dialect levelling in the 14<sup>th</sup> and 15<sup>th</sup> centuries and the emergence of regional standard dialects, with that of the East Midlands, a central and prosperous region, becoming a national standard.

Secondly, centres of learning existed in Taxila and Benares and perhaps elsewhere. Such centres could have performed a similar role in the standardisation of Indo-Aryan as the British boarding schools did for Received Pronunciation in English.<sup>84</sup> Not that this standard was a Taxila or Benares dialect;

---

<sup>83</sup>Cf. George Puttenham's *Of the Arte of English Poesie* 1589 Book 3 Chapter 3, quoted by Freeborn (2006: 321):

"...ye fhall therfore take the vfuall fpeach of the Court, and that of London and the fhires lying about London within lx. myles, and not much aboue. I say not this but that in euery fhyre of England there be gentlemen and others that fpeake but fppecially write as good Southerne as we of Middlefex or Surrey do, but not the common people of euery fhire, to whom the gentlemen, and alfo their learned clarkes do for the moft part condefcend..."

<sup>84</sup>Crystal (2005: 468-9) describes the development in the 19<sup>th</sup> century of Received Pronunciation, originally called *Public School Pronunciation*:



the probability is that dialects from across India accommodated to the regional standard language of an economically active and reasonably central area of Aryan India, Avanti for example. Such an area would produce more students, and their central speech would be more readily understood and influential in forming a new variety, based on the western dialect. Such a variety could then filter down through government, education, the religious and the army and by intermarriage across dialect areas. We have examples of these processes. The Buddha had contact with five men who had studied at Taxila: King Pasenadi of Kosala; Mahāli, a Licchavi chief and educator; Jīvaka, physician to King Bimbisāra of Magadha; Aṅgulimāla, the mass murderer, possibly a Kāli devotee<sup>85</sup>, turned Buddhist monk and Yasadatta, a Mallan convert to the Buddhist order.

King Pasenadi could have influenced the spread of this standard language through his connections across dialect boundaries: two Mallans, Bandhula, a classmate of King Pasenadi at Taxila, and Dīghakārāyāṇa, Bandhula's nephew, served as generals in the Kosalan army; Kosaladevī, Pasenadi's sister, was married to King Bimbisāra of Magadha; Vajirā, Pasenadi's daughter, was married to Ajātasattu, the following King of Magadha. This educated sociolect could move through such networks across dialect boundaries. Once used at home, the children would become native sociolect speakers and the sociolect would spread beyond the formally educated and the elite through networks of government officials, merchants and religious wanderers to become a standard language accepted throughout India.

A third driver would be the recent rise of urban settlements. Chambers and Trudgill (2004: 166) argue that linguistic innovations move, not as a ripple from the innovation source, but as a pebble skipping across water in discontinuous movements, going from population centre to centre and then rippling out from the many centres. Urban settlements would have facilitated the spread of the educated sociolect in this more efficient pattern.

---

“There was never total uniformity, therefore, but this new accent was certainly one which was more supra-regional than any previous English accent had ever been. The regional neutrality, Ellis believed, had come from a natural process of levelling, with educated people from different regional backgrounds increasingly coming into contact and accommodating to each other's speech... University education had brought people from many regional backgrounds together. Schoolteachers were exercising an increased influence on their charges, and a momentum was building up within the schools themselves ...”

<sup>85</sup>Gombrich (1996: 135ff.)

Sanskrit did not perform the function of a standard language until the Common Era, judging from epigraphical evidence. Deshpande (1986: 317) writes of Sanskrit in the time of Patañjali “It was mainly restricted to the sacrificial and academic activity of learned Brahmins of the Āryāvartta. Women and non-Brahmins were not the normative speakers of Sanskrit.” Even though Sanskrit was available, probably as a taught second language among Brahmins, the Buddha would not have adopted it, for ideological reasons. As Deshpande (1978: 41) states:

On the higher philosophical plane, Buddha totally rejected hereditary caste rank . . . However, on the lower plane, there is a clear assertion that Ksatriyas are superior to Brahmins. Thus, from his point of view, far from being an inferior dialect, Buddha must have considered his own dialect superior to that of the Brahmins, as he considered his own Ksatriya rank superior to theirs . . . only on this interpretation can we explain why the Pāli Buddhist tradition came to view Pāli to be the supreme original language of all beings including gods.

The standard language of inscriptions before Sanskrit was Epigraphic Prakrit. A language with the geographical scope, longevity and non-sectarian character of Epigraphic Prakrit could not have appeared out of nothing. It must have had a long and broad hinterland of several centuries of development across India. From the evidence of §5.2.3, the precursor of Epigraphic Prakrit, an India-wide standard language based on a western dialect, was Pali. We have now seen the processes whereby Pali, also a western dialect, could have developed into the standard Prakrit by the time of the Buddha.

### **5.5 Pali was not originally an artificial language**

Geiger (1916: 2) states “There is now on the whole a consensus of opinion that Pāli bears the stamp of a “Kunstsprache,” *i.e.* it is a compromise of various dialects.” Geiger (1916: 5-6) proposed that an artificial language, *Kunstsprache*, must have been created after the death of the Buddha based on his language, supposedly a lingua franca based on Māgadhī. Geiger’s evidence for *Kunstsprache* was the lack of homogeneity in Pali, and indeed Lamotte called it a composite language (§5.2.1). What *compromise* and *composite* mean in this context is not clear. English, for example, commonly uses words of Danish origin like ‘they’, ‘them’, ‘window’ and words from French, Latin, Dutch etc., but this does

not make English an artificially constructed language. Lack of homogeneity is normal,<sup>86</sup> and even in the exceptionally homogeneous Sanskrit, there are borrowings from non-Aryan languages. Geiger argues that numerous double forms prove Pali to be a mixed dialect and therefore an artificial compromise, but a natural language like English has these features. Crystal (2005: 75) gives the following dialect words derived from Old Norse followed by the standard English: almsow/ alms, ewer/ udder, garth/ yard, kirk/ church, laup/ leap, nay/ no, scrive/ write, trigg/ true. Furthermore, in the past tense we see the following doublets: bet/ betted, burnt/ burned, cost/ costed, crept/ creped, crew/ crowed, dove/ dived etc. English is not an artificial language like the completely regular Esperanto; instead it is as natural a language as any, despite double forms.

*Kunstsprache* means literally ‘art-speech’ and has connotations of an artificial, constructed, literary language. This was perhaps a natural influence on a German scholar 45 years after the unification of Germany in 1871, when there was a need for a state-wide standard language and committees, sometimes at the behest of government, designed compromises amongst the varieties to establish standard vocabulary and also standard pronunciation for the stage (Russ 1994: 12-16).<sup>87</sup> However, for the Buddha’s day, a better model to use would be that of 15<sup>th</sup> century England, which was not a newly-formed state, which lacked state education, which had little literacy and in which printing was a novelty. According to Fennell (2001: 123-5), English had evolved three standard varieties, one of which, the East Midlands dialect, was adopted by a government department from 1430 onwards and by Caxton as a printing standard in 1476; this went on to become standard throughout the educated classes by Shakespeare’s day.

In acknowledgement of its prominence, the East Midlands dialect had help from government and printing which accelerated its progress to becoming a national standard. However, prior to this Fennell (2001: 125) believes there existed unattested standard languages to meet the needs of a mobile population involved in pilgrimages, crusades, universities, inns of court and royal households. These examples suggest that evolution and natural or deliberate

---

<sup>86</sup>Lodge (1993: 35) writes“... Latin, like any natural language, was an amalgam of varieties, not the homogeneous monolith depicted in the ‘Latin myth’.”

<sup>87</sup>Oberlies (2003: 166) suggests that Māgadhī was based on an artistic MIA *Dichtersprache*. *Dichtersprache* was a supra-regional written literary language of the 12<sup>th</sup> century Hohenstaufen court. However, it is doubtful that this language, based on Swabian, a southern German dialect, was ever a spoken supra-regional standard, even in royal courts. As he did not refer to *Dichtersprache*, I assume Geiger did not have this in mind for *Kunstsprache*.

selection can also produce standard languages. Natural selection in medieval England is a better model for Pali than the intelligent design of 19<sup>th</sup> century Germany. The earlier social situation matches ancient India better, and no-one would design a language like Pali with so many irregularities and double forms. By contrast, the precision of Sanskrit suggests intelligent design to eliminate irregularities from an existing language.<sup>88</sup>

A religious *Kunstsprache* of a new order of celibate monks could not have produced the scope, longevity and non-sectarian character of Epigraphic Prakrit. It would not have become the standard inscriptional language ahead of the Sanskrit of centuries-old Brahmanism. The social need for communication across dialects would not have originated with the Buddha and must have been met already by an existing standard language.<sup>89</sup> *Kunstsprache* is implicit in the notion of *Hochdeutsch*, standard German, which originated as an amalgam of book German dialects; however, it is superfluous and anachronistic in oral Pali and puts an unnecessary wedge between the Buddha's language and Pali.

### **5.6 Conclusion: the Buddha adopted western Prakrit for his order**

We are now able to infer the language of the Buddhist order on formal and informal occasions. By requiring his monks to use standard vocabulary (§5.1), the Buddha ensured mutual intelligibility within the Sangha because the remaining dialect differences were essentially different accents (§5.3). The standard vocabulary would have been from western Prakrit (§5.2) and this would have ensured harmonious informal communication within the Sangha.

However, to meet the needs of a sufficiently fixed transmission for group recitation, standard morphology and pronunciation were also required and these too would have been adopted from western Prakrit. The result was Pali, which provided the medium for the ideal of fixed transmission and was lexically, morphologically and syntactically fixed, but phonetically somewhat fluid, although identifiably western. Effort would have been required for those unfamiliar with this sociolect to understand it, but preaching would often be in the local dialect by disciples native to that area, who could explain recited Pali texts; in any case, exposure to Pali recitation would make it more easily comprehensible. If he was not a native

---

<sup>88</sup>Kulikov (2013: 68) tracks the debate over whether Pāṇini's Sanskrit accurately reflected language in actual use.

<sup>89</sup>This argument against Geiger's *Kunstsprache* applies also to Gombrich's hypothesis (2018: 78) that the Buddha created a lingua franca.

speaker of Pali, the Buddha himself would presumably have set a good example by preaching in Pali, at least on formal occasions or to a mixed-dialect audience, often the case with the accompanying Sangha. Furthermore, this sociolect of the educated would have offered the gravitas of a formal register to occasions such as the Buddha preaching, uposatha observances and recitation of texts.

## 6. The MOTT refuted

Oberlies (2003: 166) provides a recent and clear explanation of the MOTT. I summarise him first and then reject the MOTT point by point:

1. The Buddha is reported to have said that his teachings should be given to the people not in Sanskrit, but in their own language;
2. This report explains the varied languages used for early Buddhist texts;
3. The Theravada commentaries state that the Buddha spoke Māgadhī, a language of the Ganges basin of the fifth century BCE akin to that of the eastern Aśokan edicts;
4. Pali is the result of recasting proto-canonical Māgadhī into a western dialect, although some eastern features remain that represent ‘frozen’ phonetics;
5. As Buddhism spread westward, the transformation of Māgadhī into the more archaic language of Pali over-reached itself with hyper-forms such as *Isipatana* for *R̥ṣyavr̥jana*;
6. In that way Pali originated as a mixture of dialects, as a kind of lingua franca;
7. The transformation of the eastern proto-canonical language into Pali was done orally before the writing down of the Theravadin canon.

### 6.1 MOTT Point 1. *sakāya niruttiyā* does not mean ‘each in his own dialect’

§4 has shown there is no solid evidence for understanding *sakāya niruttiyā* in terms of the MOTT. Of the seven possible interpretations, Edgerton’s MOTT reading is clearly the weakest.

## **6.2 MOTT Point 2. The varied languages of early Buddhist texts**

§2 has argued against oral translation and for a single phonetically fluid oral transmission. §5 argues that a standard language, Pali, was available at the outset of Buddhism and was the only solution to the ideal of a fixed oral transmission. There is not a single Sutta to be found in these alleged varied languages.

## **6.3 MOTT Point 3. The ‘Māgadhī’ of the Theravadin commentaries is an eastern dialect**

§3 explains this is a misreading of the commentaries’ *Māgadhābhāsā* and *Māgadhiko vohāro*.

## **6.4 MOTT Point 4. Pali is the result of recasting Māgadhī into a western dialect**

Intentional recasting, transposition or translation could not have happened during oral transmission of Buddhist scripture (§2). Apart from the Kathāvatthu, the very few pre-eastern features found in Pali are better explained by imprecise oral transmissions of a pre-western variety among speakers with different accents (§3.6-8).

## **6.5 MOTT Point 5. Hyper-forms were created as Māgadhī transformed into Pali**

§3.7 explains that Māgadhī could not have changed into Pali by any natural process, therefore *transformed* is taken as and meaning ‘translated’ in this context. Norman (1989: 375) defines hyper-forms as “... forms which are unlikely to have had a genuine existence in any dialect, but which arose as a result of bad or misunderstood translation techniques.” However, the existence of hyper-forms does not prove translation as they exist in any natural language and are created by native speakers. In the case of English, Trudgill (2000: 76) writes:

Bristol speech is famous for the presence in this accent of a phenomenon known as the ‘Bristol l’. In the Bristol area, words such as *America*, *India*, *Diana*, *Gloria* are pronounced with a final ‘l’ - ‘Americal’, ‘Indial’, ‘Dianal’, ‘Glorial’ - with the result that it was once said that there was a family of three sisters there called ‘Evil, Idle and Normal’... Nobody can be sure why this is, but it may be an example of what dialectologists call hypercorrection. This happens when speakers try to acquire a pronunciation which

they perceive as having higher status than their own, but overdo it. An example is when northerners trying to speak with a southern accent change not only ‘oop’ /ʊp/ to ‘up’ /ʌp/ and ‘bootter’ /bʊtə/ to ‘butter’ /bʌtə/ but also ‘hook’ /hʊk/ to ‘huck’ \*/hʌk/ and ‘good’ /gʊd/ to ‘\*/gʌd/. It may be that ‘l’ at the ends of words disappeared in Bristol, as a natural sound change, and was then restored even where it did not belong by speakers trying to talk ‘correctly’.

English is not unique in having native speakers who produce hyper-forms. Vedic and German also offer examples,<sup>90</sup> and Pali should be assumed to have behaved similarly. Both Norman (1989: 375) and Oberlies (2003: 166) give the example of Pali *Isipātana* for Prakrit *\*isivayana*, Sanskrit *R̥ṣyavr̥jana*, as proof of translation. Yet native speaker hyper-corrections based on a confusion over whether the place name *\*isivayana* meant ‘gathering of the seers’ or ‘wild-animal enclosure’ are an alternative explanation. Oberlies (2001: 79) gives as examples of Pali hyper-forms using unvoiced for voiced consonants, including *ajakara* ‘python’ for Sanskrit *ajagara* and *chakala* ‘he-goat’ for Sanskrit *chagala*; he tendentiously calls them ‘hyper-translations’. Yet, in the first case, this could be a native speaker hypercorrection or simple mistake, as Pali normally has voiced *ajagara* and only one instance of unvoiced *ajakara* (J III.484). In the second case, Pali has both unvoiced *chakala* and voiced *chagala* and one could consider the latter to be the result of a process of lenition common across the Indian languages (Bubenik 1996: 54-8; Pischel §192, §198, §200, §203, §204); we should not assume, as Oberlies appears to, that the Sanskrit form is earliest.

## 6.6 MOTT Point 6. Pali originated as a mixture of dialects, as a kind of lingua franca

This opinion is not universally held even amongst MOTT advocates. Edgerton (1953: 11) regarded each of Pali and Buddhist Hybrid Sanskrit (BHS) as essentially a single dialect:

As is well known, Pāli also shows linguistic differences between the gāthās, canonical prose, later prose etc. . . . and dialect mixture in all of

---

<sup>90</sup>Bloomfield & Edgerton (1932: 20) refer to hyper-Sanskritisms in Vedic. Wells (1987: 310) provides the following examples of hypercorrection by a native German speaker, King Frederick Wilhelm I, father of Frederick the Great: *pfeldt*, *pfahren*, *pfertig*, *anpfandt*, *hundespfott*. I am informed by my translator that these refer to *Feld*, *fahren*, *fertig*, *Anfang* and *Hundespfote*.

them. I should add that, as in the case of Pāli, I find no reason to question the essential dialectical unity of the BHS Prakrit. Such differences as occur are minor compared to the great mass of resemblances.

Arguments against Pali being a lingua franca were presented in §5.4 and §5.5.

### **6.7 MOTT Point 7. The transformation into Pali was done orally**

Transformation into Pali, both intentionally (§2) and unintentionally (§3.7), has been rejected.

## **7. Evaluation of the MOTT and SOTT**

I believe I have shown how the commentaries could be correct in their claim that the Buddha spoke Pali (§3.2). Yet, under the philosophical system of Karl Popper, to which I subscribe, nothing in this paper definitively proves the Buddha did so, as definitive proof is never possible, only definitive disproof. However, I do claim that the MOTT has been disproved, whereas the SOTT has not.

For those who remain unconvinced, I now offer the argument that the SOTT represents the stronger hypothesis. There are two theoretical weaknesses of the MOTT: redundancy and lack of testability. Taking a very poor theory as an example, we could propose that the Buddha received his teachings from aliens from another planet. This is a poor theory because, first, there is no necessity to consider extra-terrestrials to explain the Buddha's teachings and because, second, there is no way to disprove their presence in India in the 5th century BCE; absence of evidence of aliens is not evidence of absence of aliens. Similarly, the MOTT is a poor theory because, first, there is no necessity to hypothesise multiple oral transmissions when a single transmission will explain the data more economically and, second, because the myth of lost oral transmissions in other dialects cannot be disproved. Thus, the MOTT does not meet the criteria of Occam's Razor and of testability.

On the other hand, the SOTT does not have these problems: there must have been at least one oral transmission for the written texts to exist, so there is no redundancy. It is also disprovable by the following evidence which would entail modification or abandonment of the SOTT:

1. Evidence from an oral culture, preferably of South Asia, of a large volume of communally recited material which changes, by accident or design, into another dialect.



2. Early Buddhist Texts in Māgadhī, Old Māgadhī or Old Ardha-Māgadhī.
3. Early texts describing a process of homogenising several dialects into one ‘literary’ language.
4. Texts describing translation from one dialect to another.

Moreover, the commentaries of Buddhaghosa are not just one body of work amongst that of many other scholars; they are evidence for the SOTT. The MOTT on the other hand has not a single sample to offer of the multiple dialects allegedly spoken by the Buddha, let alone a whole sutta or Nikāya. Modern scholarship has not disproved the Theravadin claim of the Buddha teaching only in Pali, it has merely misread it with a lack of sympathy. According to Carl Sagan’s dictum, an extraordinary claim, such as the MOTT’s implied oral translation of over 5,000 pages of text, requires extraordinary evidence. Until the Theravadin tradition is disproved by unambiguous evidence, the only intellectually rigorous position to adopt is to take the Buddha teaching in Pali as its working hypothesis. As Gombrich (1990: 8) states, “I also think it sound method to accept tradition until we are shown sufficient reason to reject it.” We have not been shown sufficient reason so far and must reject assertions such as von Hinüber’s (2006: 209) “the Buddha did not speak Pāli.”

## 8 Implications of the SOTT

One might think that adopting the SOTT would produce a major shift in Buddhist scholarship. Far from it. While the SOTT may increase subjective interest in Pali significantly, it might have only a minor incremental impact on objective scholarship. In fact, this potentially minor impact would explain why the MOTT has been unchallenged for so long.

### 8.1 A non-sectarian approach to Early Buddhist Texts

I suspect in the MOTT a kind of “political correctness”, the premise that all sects of Buddhism are to be treated impartially and, if some sects have had the language of their texts altered, then all sects must have had them altered.<sup>91</sup> If this is the motivation behind the misinterpretations<sup>92</sup> of the MOTT, then it is

---

<sup>91</sup>This may be the thinking behind Law (1933: xi-xxv) and Lamotte (1958: 607).

<sup>92</sup>Viz. the misreadings of *sakāya niruttīyā* (§4), *samaññaṃ nātidhāveyyāti* (§5.1) and

political correctness gone mad. The MOTT is entirely unnecessary to achieve the laudable aim of impartiality. Acknowledging that the Pali canon is in the original language of the Buddha does not mean that the Pali canon is the original transmission; that would be naive in the extreme. In the Pali transmission, there may have been mistakes, later additions and reorganisations peculiar to the Theravadins. Thus this paper in no way confers priority on the Pali version of the EBTs, which in fact needs other recensions to amend corruptions of its texts.<sup>93</sup> The SOTT in no way implies Theravada fundamentalism, and the trend of modern scholarship to treat each recension of scripture on an equal basis is completely unaffected.

## 8.2 A historical approach to Early Pali texts

The findings of this paper reject what Sujato & Brahmali (2015: 145) call Denialist Buddhism, “a rhetoric of scepticism”. For example, Gregory Schopen is quoted in Wynne (2004: 116) as saying:

Even the most artless formal narrative text has a purpose, and that in “scriptural” texts, especially in India, that purpose is almost never “historical” in our sense of the term... Scholars of Indian Buddhism have taken canonical monastic rules and formal literary descriptions of the monastic ideal preserved in very late manuscripts and treated them as if they were accurate reflections of the religious life and career of actual practising Buddhist monks in early India.

Wynne counters that such narrative can be taken as circumstantial evidence and therefore valuable in a historical approach. This paper follows Wynne’s approach and infers the linguistic reality in which the Buddha operated. It rejects Schopen’s approach as potentially corrosive and leading to a slippery

---

*magadhabhāsā* (§3.2).

<sup>93</sup>For example, Levman (2014: 34-8) gives the example of Dhṛ 414-b *saṃsāraṃ moham*, ‘samsara, delusion’ as compared to Udānavarga 33.41-b *saṃsāraugham*, ‘the flood of samsara’. He believes the original Pali was *saṃsāramo(g)ham*, ‘the flood of samsara’ and *ogha*, ‘flood’, was reduced to *oha*, leaving the ambiguous *saṃsāramoham* meaning either ‘the delusion of samsara’ or ‘the flood of samsara’; the Pali commentary refers to both these meanings, but the PTS and Be texts support only the meaning ‘delusion’. As the BHS, Sanskrit, Chinese and Tibetan versions support only the meaning of ‘flood’, it appears that the Pali is corrupt, and other recensions are needed to correct its text.

slope: if we do not have the language of the Buddha, then we do not have the original teachings, and, if we do not have the original teachings, then perhaps the Buddha is a mythical figure.

Furthermore, this paper reinforces Wynne's conclusion (2004: 124) that Pali literature can in principle be stratified, for Pali was a natural, not artificial, language.

### **8.3 The continuity of Pali and Epigraphic Prakrit**

The claim of a Pali/ Epigraphic Prakrit continuity is explicit here, but has been hinted at by several scholars (§5.2.2). Yet it has not so far been examined in detail by the academic community, perhaps because of the prevailing MOTT ideology that Pali is an artificial language. A database of Epigraphic Prakrit is a desideratum to investigate this claim further. Such a database might also assist the stratification of Pali texts.

### **8.4 Reworking Pali historical (socio)linguistics**

The SOTT requires adjustments to Indian historical (socio)linguistics. Principally, the search for the Ur-language of Buddhism can be dropped as the evidence points to its being Pali. This means that the extent of Sanskritisation and Māgadhisms in Pali needs to be revised downwards, though the data collection on which the claims are based remains useful. That Pali was the Buddha's teaching language could also be a partial explanation why the Aśokan inscriptions did not use Pali or Epigraphic Prakrit: Aśoka for reasons of state did not wish to show partiality towards Buddhism just as he equally avoided the Sanskrit of the Brahmins and the Ardha-Māgadhī of the Jains. The narrative of translation from Māgadhī needs to be abandoned together with the claim of Pāli being an artificial language. For Pali remains overwhelmingly a western dialect with minimal intrusions from other dialects and it is a testimony to the efforts of perhaps hundreds of thousands of personnel across twenty-five centuries to preserve, not only the Buddha's teachings, but also his language.

## Abbreviations

Abbreviations are as in the *Dictionary of Pāli, Volume 1* with the following additions:

|      |   |
|------|---|
| AMg  | Ardha-Māgadhī                               |
| AN   | Aṅguttara Nikāya                            |
| BBC  | British Broadcasting Corporation            |
| Be   | Burmese/ Myanmar edition                    |
| BHS  | Buddhist Hybrid Sanskrit                    |
| Cv   | Cūḷavagga                                   |
| DN   | Dīgha Nikāya                                |
| DOP  | Dictionary of Pāli                          |
| DPPN | Dictionary of Pāli Proper Names             |
| DPR  | Digital Pāli Reader                         |
| EBT  | Early Buddhist Text                         |
| fn   | footnote                                    |
| H    | Hāthīgumphā inscription                     |
| IPA  | International Phonetic Alphabet             |
| Mg   | Māgadhī                                     |
| MIA  | Middle Indo-Aryan                           |
| MN   | Majjhima Nikāya                             |
| MOTT | Multiple Oral Transmission Theory           |
| Mv   | Mahāvagga                                   |
| MW   | Monier-Williams Sanskrit English Dictionary |
| PE   | Pillar Edict                                |
| PED  | Pāli-English Dictionary                     |
| PTS  | Pāli Text Society edition                   |
| RE   | Rock Edict                                  |
| Skt  | Sanskrit                                    |
| SN   | Samyutta Nikāya                             |
| SOTT | Single Oral Transmission Theory             |

## References

Brackets () around a second year of publication indicate reference to the later edition's pagination.

- Agrawal, Ashvini 1989: *Rise and Fall of the Imperial Guptas*. Delhi: Motilal Banarsidass.
- Allon, Mark 1997: *Style and Function: A study of the dominant stylistic features of the prose portions of the Pāli canonical sutta texts and their mnemonic function*. Tokyo: The International Institute for Buddhist Studies.
- Anālayo, Bhikkhu 2009: "The Vicissitudes of Memory and Early Buddhist Oral Transmission" pp. 5-19 *Canadian Journal of Buddhist Studies* Number Five, 2009.
- 2011: *A Comparative Study of the Majjhima-nikāya*. Taipei: Dharma Drum Publishing Corporation.
- 2016: "Selected Madhyama-āgama Discourse Passages and their Pāli Parallels" in *Dharma Drum Journal of Buddhist Studies*, no. 19 pp. 1-61 (2016).
- Barnes, Michael 2010: "The Study of Norm" pp. 26-47 in *Northern Lights, Northern Words. Selected Papers from the FRLSU Conference, Kirkwall 2009* edited by Millar, Robert McColl. Aberdeen: Forum for Research on the Languages of Scotland and Ireland.
- Barua, Benimadhab 1926: *Bharhut Inscriptions*. Calcutta: University of Calcutta.
- 1929: *Old Brāhmī Inscriptions*. Calcutta: Calcutta University Press.
- Beal, Joan C. 2010: *An Introduction to Regional Englishes*. Edinburgh: Edinburgh University Press.
- Bechert, Heinz 1980: "Introduction." pp. 11-18 in *Die Sprache der ältesten buddhistischen Überlieferung = The Language of the Earliest Buddhist Tradition* edited by Bechert, Heinz. Göttingen: Vandenhoeck & Ruprecht.
- 1991: "Methodological Considerations Concerning the Language of the Earliest Buddhist Tradition." pp. 3-19 *Buddhist Studies Review* 8, 1-2.
- Bilalić, Merim & McLeod, Peter 2014: "Why Good Thoughts Block Better Ones." pp. 75-79 *Scientific American* March 2014, Volume 310, No 3.
- Bloomfield, Maurice & Edgerton, Franklin 1932: *Vedic Variants: Volume II Phonetics*. Philadelphia: The Linguistic Society of America.
- Bodhi, Bhikkhu 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*. Bristol: The Pali Text Society.
- 2017: *The Suttanipāta: An Ancient Collection of the Buddha's Discourses*. Somerville: Wisdom Publications.
- Bronkhorst, Johannes 2007: *Greater Magadha: Studies in the Culture of Early India*. Leiden: Brill.
- Brough, John 1962: *The Gāndhārī Dharmapada*. Delhi: Motilal Banarsidass.
- 1980: “Sakāya Niruttīyā: Cauld kale het.” pp. 35-42 in *Die Sprache der ältesten buddhistischen Überlieferung = The Language of the Earliest Buddhist Tradition* edited by Bechert, Heinz. Göttingen: Vandenhoeck & Ruprecht.
- Bubenik, Vit 1996: *The Structure and Development of Middle Indo-Aryan Dialects*. Delhi: Motilal Banarsidass.
- Buswell, R.E. Jr. & Lopez, S. Jr. 2014: *The Princeton Dictionary of Buddhism*. Princeton & Oxford: Princeton University Press.
- Cardona, George & Jain, Dhanesh (Editors) 2003: *The Indo-Aryan Languages*. London: Routledge.
- Chambers, J.K. & Trudgill, Peter 2004: *Dialectology* 2<sup>nd</sup> Edition. Cambridge, Cambridge University Press.
- Childers, Robert Caesar 1875: *A Dictionary of the Pāli Language*. London: Trübner & Co.
- Cone, Margaret 2010: *A Dictionary of Pāli: Part II g-n*. Bristol: The Pali Text Society.
- Cousins, L.S. 2013: “The Early Development of Buddhist Literature and Language in India.” *Journal of the Oxford Centre for Buddhist Studies*. (5): 89-135.
- Crosby, Kate 2004: “The Origin of Pāli as a Language Name in Medieval Theravāda Literature.” *Journal of the Centre for Buddhist Studies, Sri Lanka*, 2. pp. 70-116.
- Crystal, David 2005: *The Stories of English*. London: Penguin Books Limited.
- D’Alwis, James 1863: *An Introduction to Kachchayana’s Grammar of the Pali Language*. London & Edinburgh: Williams and Norgate.
- Deshpande, Madhav M. 1979: *Sociolinguistic Attitudes in India: An Historical Reconstruction*. Ann Arbor: Karoma Publishers, Inc.
- 1986: “Sanskrit Grammarians on Diglossia.” pp. 312-321 of *Structure, Convergence and Diglossia*. Edited by Krishnamurti, B.H. Delhi: Motilal Banarsidass.

- Edgerton, Franklin 1953: *Buddhist Hybrid Sanskrit Grammar and Dictionary Volume I: Grammar*. Delhi: Motilal Banarsidass Publishers Private Limited.
- Fennell, Barbara A. 2001: *A History of English: A Sociolinguistic Approach*. Oxford: Blackwell.
- Fraser, Bruce 1996: "Pragmatic Markers." *Pragmatics* 6:2.167-190.
- Freeborn, Dennis 2006: *From Old English to Standard English*. Basingstoke: Palgrave Macmillan.
- Geiger, Wilhelm 1916, (1943): *Pāli Literature & Language*. Translated by Batakrişna Ghosh. Delhi: Oriental Books Reprint Corporation. [First published as *Pāli Literatur und Sprache*, Strassburg: Karl J. Trübner].
- Gokhale, Balakrishna Govind 1982: "Early Buddhism and the Urban Revolution." pp. 7-22 in *Journal of the International Association of Buddhist Studies* Vol. 5 No.2.
- Gombrich, Richard F. 1988: *Theravada Buddhism: A social history from Ancient Benares to Modern Colombo*. London: Routledge & Kegan Paul.
- 1990, (2012): "Recovering the Buddha's Message." pp. 5-20 in *The Buddhist Forum: Vol I Seminar Papers 1987-88*. Edited by Skorupski, Tadeusz. Tring, UK & Berkeley, USA: The Institute of Buddhist Studies. [First published by the School of Oriental and African Studies (University of London), 1990]
- 1994: "Introduction: What is Pāli?" pp. xxiii-xxix in *A Pāli Grammar by Wilhelm Geiger translated into English by Batakrişna Ghosh revised and edited by K.R. Norman*. Oxford: Pali Text Society.
- 1996: *How Buddhism Began: The Conditioned Genesis of the Early Teachings*. London: The Athlone Press.
- 2009: *What the Buddha Thought*. London: Equinox Publishing.
- 2018: *Buddhism and Pali*. Oxford: Mud Pie Books.
- Gornall, Alistair 2014: How many sounds are in Pāli? *Journal of Indian Philosophy* (2014) 42:511– 550. DOI 10.1007/s10781-014-9221-z.
- Hazen, Kirk 2011: "Labov: Language Variation and Change." pp. 24-39 in Wodak, Ruth, Johnstone, Barbara & Kerswill, Paul (Editors) *The Sage Handbook of Sociolinguistics*. London: Sage Publications Ltd.
- Hock, Hans Heinrich & Pandharipande, Rajeshwari 1978: "Sanskrit in the Pre-Islamic Sociolinguistic Context of South Asia." pp. 11-25 in *International Journal of the Sociology of Language* Issue 16. The Hague: Mouton de Gruyter
- Hock, Hans Heinrich 1991: *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.

- Horner, I.B. 1959: *The Middle Length Sayings (Majjhima-Nikāya) Vol III* Oxford: Pali Text Society.
- 1962: *The Book of the Discipline (Vinaya Pitaka) Vol IV (Mahāvagga)*. London: Luzac & Co.
- 1963: *The Book of the Discipline (Vinaya Pitaka) Vol V (Cullavagga)*. London: Luzac & Co.
- Hultzsch, E 1925, (1991): *Corpus Inscriptionum Indicarum Vol I: The Inscriptions of Aśoka*. NewDelhi: Archaeological Survey of India. [First published Oxford: Clarendon Press.]
- Ireland, John D. 1976: “The Kosambi Suttas.” pp. 105-121 in *Pali Buddhist Review*. Vol 1, No 2 1976.
- Knipe, David M. 2015: *Vedic Voices: Intimate Narratives of a Living Andhra Tradition*. Oxford: Oxford University Press.
- Kulikov, Leonid 2013: “Language vs. grammatical tradition in Ancient India: How real was Pāṇinian Sanskrit? Evidence from the history of late Sanskrit passives and pseudo-passives.” *Folia Linguistica Historica* 34 (2013), 59–91.
- Lamotte, Étienne 1958, 1988: *History of Indian Buddhism*. Translated from the French by Sara Webb-Boin. Louvain & Paris: Peeters Press. [First published as *Histoire du Bouddhisme Indien*, Louvain: Bibliothèque du Muséon.]
- Law, B.C. 1933: *A History of Pāli Literature*. London: Kegan Paul, Trench, Trubner & Co., Ltd.
- Labov, William 1972: “Some principles of linguistic methodology.” *Language in Society* 1, 97-120.
- Lévi, Sylvain 1915: “Sur la Récitation Primitive des Textes Bouddhiques.” pp. 401-447 in *Journal Asiatique Mai-Juin 1915*.
- Levman, Bryan 2008/2009: “Sakāya niruttīyā Revisited.” *Bulletin d’Études Indiennes*, 26/27, 33-51.
- 2013: “Cultural Remnants of the Indigenous Peoples in the Buddhist Scriptures.” pp. 145-180 in *Buddhist Studies Review* 30.2 (2013).
- 2014: *Linguistic Ambiguities, the Transmissional Process, and the Earliest Recoverable Language of Buddhism*. A thesis submitted for the degree of Doctor of Philosophy. [https://tspace.library.utoronto.ca/bitstream/1807/68342/1/Levman\\_Bryan\\_G\\_201406\\_PhD\\_thesis.pdf](https://tspace.library.utoronto.ca/bitstream/1807/68342/1/Levman_Bryan_G_201406_PhD_thesis.pdf)
- 2016: “The Language of Early Buddhism.” pp. 1-41 in *Journal of South Asian Languages and Linguistics* JSALL 2016; 3(1): 1–41.



- Lodge, R., Anthony 1993: *French: From Dialect to Standard*. London and New York: Routledge.
- Lüders, Heinrich 1911 “The Lingual La in the Northern Brahmi Script.” *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 1911, pp. 1081–1089. [www.jstor.org/stable/25189956](http://www.jstor.org/stable/25189956).
- 1954: *Beobachtungen über die Sprache des Buddhistischen Urkanons*. Berlin: Akademie-Verlag.
- , Waldschmidt, E. & Mehendale, M.A.1963: *Bharhut Inscriptions*. Ootacamund: Government Epigraphist for India.
- Manné, Joy 1990: “Categories of Sutta in the Pāli Nikāyas.” pp. 29-87 in *Journal of the Pali Text Society, Vol XV*.
- Masefield, Peter 1994: *The Udāna*. Oxford: The Pali Text Society.
- Ñāṇamoli, Bhikkhu & Bodhi, Bhikkhu 1995: *The Middle Length Discourses of the Buddha: A New Translation of the Majjhima Nikāya*. Boston: Wisdom Publications.
- Norman, K.R. 1970, 1999: “Lexical variation in the Aśokan inscriptions.” pp. 130-43 in *Collected Papers Vol I K.R. Norman*. Oxford: Pali Text Society. [First published in *Transactions of the Philological Society* 1970 pp. 121-36]
- 1971, 1999: “Middle Indo-Aryan Studies VIII.” pp. 122-129 in *Collected Papers Vol I K.R. Norman*. Oxford: Pali Text Society [First published in *Journal of the Oriental Institute (Baroda)* XX, 1971, pp. 329-36]
- 1976, 1999: “Pāli and the language of the heretics.” pp. 238-246 in *Collected Papers Vol I K.R. Norman*. Oxford: Pali Text Society [First published in *Acta Orientalia* 37 (1976) pp. 117-26]
- 1979, 2003: “Māgadhisms in the Kathāvatthu.” pp. 59-70 in *Collected Papers Vol II K.R. Norman*. Oxford: Pali Text Society. [First published as pp. 279-87 in *Studies in Pali and Buddhism (in Honor of Bhikkhu Jagdish Kashyap)*. Delhi: B.R. Publishing Corporation.]
- 1980: “The Dialects in which the Buddha preached.” pp. 61-77 in *Die Sprache der ältesten buddhistischen Überlieferung = The Language of the Earliest Buddhist Tradition* edited by Bechert, Heinz. Göttingen: Vandenhoeck & Ruprecht.
- 1983: *Pāli Literature*. Weisbaden: Otto Harrassowitz.
- 1988, 2008: “The Origins of Pāli and its Position among the Indo-European Languages.” pp. 225-243 in *Collected Papers Vol III K.R. Norman*. Oxford: Pali Text Society. [First published in *Journal of Pali and Buddhist Studies* I (1988) pp. 1-27.]

- 1989, 2008: “Dialect Forms in Pali.” pp. 46-71 in *Collected Papers Vol IV K.R. Norman*. Oxford: Pali Text Society. [First published as pp. 369-92 in Caillat, Colette (Ed.) *Dialectes dans les littératures Indo-Aryennes*. Paris: College de France, Institut de Civilisation Indienne.]
- 1992: “Pāli Lexicographical Studies IX.” pp. 71-79 in *Collected Papers Volume V K.R. Norman*. Oxford: Pali Text Society. [First published in *Journal of the Pali Text Society XVI*, 1992, pp. 77-85.]
- 1993, 1994: “The Languages of Early Buddhism.” pp.146-168 in *Collected Papers Volume V K.R. Norman*. Oxford: Pali Text Society. [First published as pp. 83-99 in *Premier Colloque Étienne Lamotte*. Louvain: Publications de l’Institut Orientaliste de Louvain]
- 1997: *A Philological Approach to Buddhism*. London: School of Oriental & African Studies.
- 2002, 2007: “Pāli and the languages of early Buddhism.” pp. 96-119 in *Collected Papers Volume VIII K.R. Norman*. Lancaster: Pali Text Society. [First published as pp. 135-50 in *Indo-Iranian Languages and Peoples*, Sir Harold Bailey Centennial Volume, Proceedings of the British Academy 116, London.]
- Oberlies, Thomas 2001, (2012): *Pāli: A Grammar of the Language of the Theravāda Tipiṭaka*. New Delhi: Munshirma Manoharlal Publishers. [First published in Berlin: Walter de Gruyter.]
- 2003: “Aśokan Prakrit and Pāli.” pp. 161-203 in Cardona, George & Jain, Dhanesh (Editors) *The Indo-Aryan Languages*. London & New York: Routledge.
- Pischel, R. 1900, (1957): *Comparative Grammar of the Prākṛit Languages*. Translated from the German by Subhadra Jhā. Delhi: Motilal Banarsidass. [First published as *Grammatik der Prakrit-Sprache*. Strassburg: Verlag von Karl J. Trübner]
- Piyasīlo, Ven 1996: “What language did the Buddha speak?” pp. 143-148 in *Buddhist Studies Review* 13.2 (1996)
- Pollock, Sheldon 2016: *The Language of the Gods in the World of Men: Sanskrit, Culture, and Power in Premodern India*. Berkeley: University of California Press.
- Rhys Davids, T.W. & Oldenberg, Hermann 1882: *Vinaya Texts: Part II*. Oxford: Clarendon Press.
- , T.W.1903: *Buddhist India*. London: T. Fisher Unwin.
- , T.W & C.A.F 1910: *Dialogues of the Buddha, Translated from the Dīgha Nikāya Part II*. London: Henry Frowde, Oxford University Press.

- Ruegg, D. S. 2000: “On the Expressions *chandaso āropema*, *āyataka gītassara*, *sarabhañña* and *ārṣa* as applied to the ‘Word of the Buddha’ (*buddhavacana*).” pp. 283-306 in *Harānandalaharī: Volume in Honour of Professor Minoru Hara on his 70th Birthday* edited by Tsuchida, Ryutaro & Wezler, Albrecht. Reinbeck: Verlag für Orientalistische Fachpublikationen.
- Russ, Charles V J. 1994 (2005): *The German Language Today: A Linguistic Introduction*. London: Routledge.
- Salomon, Richard 1998: *Indian Epigraphy: A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the other Indo-Aryan Languages*. New Delhi: Munshiram Manoharlal.
- Scharfe, Hartmut 2002: *Education in Ancient India*. Leiden: Brill.
- Schopen, Gregory 1997: “If You Can’t Remember, How to Make It Up: Some Monastic Rules for Redacting Canonical Texts.” pp. 571-82 in *Bauddhavidyāsudhākaraḥ: Studies in Honour of Heinz Bechert on the Occasion of his 65<sup>th</sup> Birthday* edited by Kieffer-Pülz, Petra and Hartmann, Jens-Uwe. Swittal-Odendorf: Indica et Tibetica Verlag.
- Southworth, Franklin C. 2005: *Linguistic Archaeology of South Asia*. Abingdon: Routledge Curzon.
- Sujato, Bhikkhu & Brahmali, Bhikkhu 2015: *The Authenticity of the Early Buddhist Texts*. Toronto: Chroniker Press.
- Tatelman, Joel 2005: *The Heavenly Exploits: Buddhist Biographies from the Divyāvadāna*. New York University Press & JJC Foundation.
- Thomas, E. J. 1927: *The Life of Buddha as Legend and History*. London: Routledge & Kegan Paul.
- Trudgill, Peter 2000: *The Dialects of England*. 2<sup>nd</sup> Edition. Oxford: Blackwell Publishing.
- 2010: *Investigations in Sociohistorical Linguistics: Stories of Colonisation and Contact*. Cambridge: Cambridge University Press.
- Von Hinüber, Oskar 1977, 2005: “On the History of the Name of the Pāli Language.” pp. 76-90 in *Selected Papers on Pāli Studies: O. von Hinüber*. Oxford: The Pali Text Society. [First published as: *Beiträge zur Indienforschung Ernst Waldschmidt zum 80. Geburtstag gewidmet* (Berlin, 1977), pp. 237-46; also, *Revista de Estudios Budistas* 7, 1994, pp. 55-72.]
- 1982: “Pāli as an Artificial Language.” pp. 451-458 in *Oskar von Hinüber: Kleine Schriften. Teil I*. Edited by Falk, Harry and Slaje, Walter. Wiesbaden: Harrassowitz Verlag. [First published in *Indologica Taurinensia* 10: 133-140.]

- 1983, 2005: “The Oldest Literary Language of Buddhism.” pp. 177-197 in *Selected Papers on Pāli Studies: O. von Hinüber*. Oxford: The Pali Text Society. [First published in *Saeculum* 34 (1983), pp. 1-9.]
- 1985a, 2009: “Epigraphical Varieties of Continental Pāli from Devnimori and Ratnagiri.” pp. 489-504 in *Oskar von Hinüber: Kleine Schriften. Teil I*. Edited by Falk, Harry and Slaje, Walter. Wiesbaden: Harrassowitz Verlag. [First published in *Buddhism and its relation to other religions*. Kyoto:1985 185-200]
- 1985b, 2009: “Pāli and Paīśicī as variants of Buddhist Middle Indic.” pp. 505-521 in *Oskar von Hinüber: Kleine Schriften. Teil I*. Edited by Falk, Harry and Slaje, Walter. Wiesbaden: Harrassowitz Verlag. [First published in *Bulletin d’Études Indiennes* BEI 3 (1985): 61-77.]
- 1987, 2005: “Buddhist Law and the Phonetics of Pāli.” pp. 198-232 in *Selected Papers on Pāli Studies: O. von Hinüber*. Oxford: The Pali Text Society. [First published in *Studien zur Indologie und Iranistik* 13/14 (*Festschrift für Wilhelm Rau*, 1987), pp. 101-27.]
- 1995, 2009: “Linguistic Considerations on the date of the Buddha.” in *Oskar von Hinüber: Kleine Schriften. Teil I*. Edited by Falk, Harry and Slaje, Walter. Wiesbaden: Harrassowitz Verlag. [First published in *When did the Buddha Live? The Controversy on the Dating of the Historical Buddha* Edited by Bechert, Heinz. Delhi: Sri Satguru Publications pp. 185-194.]
- 2006: “Hoary Past and Hazy Memory: On the History of Early Buddhist Texts.” *Journal of the International Association of Buddhist Studies* Vol. 29, No. 2 2006 (2008).
- Walshe, Maurice 1987, (1995): *The Long Discourses of the Buddha: A Translation of the Dīgha Nikāya*. Boston: Wisdom Publications. [First published as *Thus Have I Heard: the Long Discourses of the Buddha*.]
- Warder, A. K. 1970, (1980): *Indian Buddhism* 2nd Revised Edition. Delhi: Motilal Banarsidass.
- Wells, C.J. 1987: *German: A Linguistic History to 1945*. Oxford: Clarendon Press.
- Witzel, Michael 1989: “Tracing the Vedic Dialects.” pp. 97-264 in Caillat, Colette (Ed.) *Dialectes dans les littératures Indo-Aryennes*. Paris: College de France, Institut de Civilisation Indienne.
- 1997: “The Development of the Vedic Canon and its Schools: The Social and Political Milieu.” pp. 257-348 in Witzel, M. (Ed.) *Inside the Texts, Beyond the Texts*. Cambridge: Harvard Oriental Series, Opera Minora 2.
- Wimalawamsa, P. & Perera, L.P.N. 1976: “What is Pali?” pp. 3-7 in *Pali Buddhist Review* Vol 1.1.

Woodward, F.L. 1979: *Gradual Sayings (Anguttara-Nikāya) Vol. I*. London: Routledge, Kegan, Paul.

Wynne, Alexander 2004: "The oral transmission of early Buddhist literature." *Journal of the International Association of Buddhist Studies*. Vol. 27, No. 1. 2004: 97-127.