Pāli and Buddhism

Pāli and Buddhism:

Language and Lineage

By

Bryan G. Levman

Cambridge
Scholars
Publishing

TABLE OF CONTENTS

Section Three: Pāli, The Transmission *Sankamana*

PREFACE

The central theme of this work is the teaching of the Buddha as embodied in the Pāli language. To some, Pāli was the language the Buddha spoke, while I argue that, although it was indeed close to the demotic speech of his day, the Pāli that has come down to us has changed quite a bit from its fifth century BCE usage by the Teacher. This is only natural, as change is the fundamental law of Buddhism and the world, and language is no exception.

Pāli is a great treasure to mine. It contains the earliest record we possess of the Buddha's teachings about life and within it his "technology" for liberation, a series of procedures which, followed as he instructed, are the *ekāyana*, the one way path to the end of suffering.

Pāli is a rich tapestry with many strands. Originating, as I argue, from a *koiné* (ἡ κοινὴ διάλεκτος, "the common dialect"), a simplified common language of north India used for social, cultural, political and business purposes at the time of the Buddha and prior, it contains elements of all the Indo-Aryan dialects of the time. Equally important is the imprint of the indigenous languages, preserved in Pāli just under the surface like a palimpsest. Because no record of the Buddha's teachings has been preserved in an indigenous language (Dravidian, Munda or Tibeto-Burman for example), scholars tend to view the Buddha's insights as the exclusive product of an Indo-Aryan culture. But in north-east India, the Indo-Aryan speaking peoples were a minority, even during the time of the Buddha. The majority were the indigenes whose first language was not Indo-Aryan. In order to participate in the new order of things, these peoples had to learn the Indo-Aryan languages which were the linguistic strata of what was fast becoming the dominant culture; and this bilingualism left a lasting imprint of borrowings from the local culture on the Indo-Aryan languages and Pāli in particular: hundreds, perhaps thousands of words, dozens of phonological constraints and idiosyncrasies, and several key and syntactical structures were adopted from the local peoples into the Indo-Aryan languages. More important than this is what these adaptations and adoptions represented, that is cultural appropriation, which is one of the main themes of this book.

Buddhism is an eclectic teaching, drawing from many sources, especially from the established culture of the autochthons. This same polymorphism may be ascribed to Pāli whose complex fabric contains many diverse features to be examined here: 1) Pāli is a mixed language, containing elements of the various vernacular dialects of north India which evolved over many centuries from the late second millennium BCE. 2) The composite nature of Pāli reflects its origin from a "common" (Gk κοινός) language to facilitate communication between diverse dialects (koiné) and diverse language groups (lingua franca), like Indo-Aryan, Dravidian, Munda, etc. 3) The phonological structure of local dialects and the indigenous languages constrained and entrained various phonetic idiosyncrasies in Pāli by linguistic diffusion. 4) In Pāli one finds extensive word borrowing from the local languages—especially of local toponyms and biota—imperfectly adapted to a foreign phonetic structure. 5) These linguistic borrowings also reflect a significant amount of cultural borrowing from the local population which are also manifested in 6) The adoption of structural and syntactical features from the local languages superimposed on the normal Indo-Aryan sentence pattern. 7) Pāli shows the normal diachronic change as the language evolved over time as well as 8) Extensive Sanskritization, that is, back formation to make it more like the sacred Vedic language from which it (in part) evolved. 9) Pāli has also been affected by transmissional errors due to the oral system of propagating the teachings for the first few centuries after the Buddha's passing, and, 10) Harmonization of the language by tradents at various stages in the language's history, bringing divergent readings into agreement, making corrections, etc., and 11) Standardization by the Pāli grammarians in medieval times.

These are some of the principal themes of this book which traces Pāli's lineage from the time of the Buddha down to modern times, with a penultimate chapter on the meaning of the important doctrinal term *sati* (usually translated "mindfulness") and a final chapter on Pāli's "correct" pronunciation. The first approximately two-thirds of the book examine the effect of the local languages on Pāli itself and the Buddhist teachings it represents, a very under-studied and important new area of research which adds a whole new dimension to Buddhist studies.

This book stands on the shoulders of many great Buddhist scholars and linguists, with a lineage that extends back over a century: Richard Pischel, Sylain Lévi, Wilhelm Geiger, Heinrich Lüders, Ernst Waldschmidt, Thomas Burrow, Murray Emeneau, Manfred Mayrhofer, Franciscus Kuiper, John Brough, Franklin Southworth, Kenneth Norman, Oskar von Hinüber, David Stampe, Michael Witzel, to name some of the more prominent ones; without

their work and such essential data bases as the Dravidian (Burrow and Emeneau), Munda (Stampe) and Sanskrit etymological dictionaries (Mayrhofer), it would be impossible to even make a start. The same must be said for the Pāli Lexicographers starting with Robert Childers (Dictionary of the Pāli Language), Thomas William Rhys Davids and William Stede (Pāli-English Dictionary), Vilhelm Trenckner, Dines Andersen, Helmer Smith, Hans Hendriksen and later editors (Critical Pāli Dictionary), Franklin Edgerton (Buddhist Hybrid Sanskrit Dictionary), and Margaret Cone (Dictionary of Pāli). Hundreds of other important sources are listed in the Bibliography.

Closer to home I would like to thank the many colleagues and associates who have taken the time to read and comment on different aspects of the book and whose suggestions have helped to improve the final manuscript. Dr. Christoph Emmrich of the University of Toronto; Ven. Ṭhānuttamo and Ven. Ariyadhammika of the Sāsanārakkha Buddhist Sanctuary in Malaysia; Ven. Bhikkhu Bodhi of the Chuang Yen Monastery, New York; Ven. Ashin Saraṇa, Ven. Dhammasami and Ven. Paññabhoga from Myanmar; Dr. Madhev Deshpande and Dr. George Cardona, professors emeriti. A special thanks as well to my two Tamil teachers Vaidehi Herbert and Mohan Thiruvengadam whose instruction in classical Tamil has deepened my understanding of the language; to Liwen Liu, doctoral candidate at the University of Toronto who has been an invaluable help with the indexing, proofreading and final editing of the book; to my jubilee wife Rosemarie for her proofreading and ever constant support and love; and to my son Jacob who created the computer program to sort the very complex indices of this book according to the syllabaries of the various non-Western languages. Lastly I honour the Buddha who made all this possible with his brilliant discoveries into the nature of life *yathābhūtaṃ* ("the way things are"); in the words of Ānanda, *lābhā vata me suladdhaṃ vata me yassa me satthā evaṃ mahiddhiko evaṃ mahānubhāvo ti* (AN 1, 228[20–21], Ānanda, speaking to the Buddha). "It is a gain to me, it is a great gain to me indeed that my teacher is possessed of such great power, of such majesty!"

His teachings are as relevant and essential to us today as they were twenty-five hundred years ago. Homage (*namo*) to that Venerable one (*tassa Bhagavato*), to that noble one (*Arahato*), the fully enlightened one (*Sammāsambuddhassa*)!

This book is dedicated to the memory of my first love, my mother, Belle ("Beautiful") Sachs-Levman 1922–2020, ever beautiful in body and soul.

CHAPTER ONE

INTRODUCTION

Abstract

Chapter One provides a brief introduction to the major themes in this book. Its first section deals with the history of early Buddhism which was largely written by Brahmans centuries after the Buddha lived. Although the Buddha came from a mixed ethnic and linguistic background, much of his background has been obscured by a Brahmanical overprint which obfuscates his connection with the indigenous tribes and presents the Buddha as an exclusive product of an Indo-Aryan Brahmanical culture. Yet parts of the older, earlier culture can still be found in the *suttas* and especially in the loan-words from the Dravidian and Munda languages adopted into Pāli. Section one of the book will be examining these words and interpreting their significance in terms of the cultural continuity of the autochthonous population in early Buddhism.

Section Two discusses the nature of the Pāli language and its evolution and Sanskritization.

Section Three gives two case examples of the transmission of Pāli over time. The first looks at the word *sati* and its semantic evolution from "memory" to "present moment awareness," arguing that memory of the Buddha's teachings is still a core meaning of the word; and the second looks at nasalization in Pāli and how the popular refuge formula *Buddhaṃ saraṇaṃ gacchāmi* is to be pronounced.

—

That history is written by the winners is an old adage of historiography. In the case of the Buddha and the community he founded, only a very lopsided account of its history has survived, one written largely by Brahman converts primarily for a Brahman audience. Although the Buddha came from an eastern tribe—the Sakyas—who were looked down upon by the Indo-Aryan immigrants as the "other" in language, culture and race, in Aśvaghoṣa's Sanskrit biography written many centuries after his death (the *Buddhacarita*,

ca. second century CE), he is historicized as "the crowning and consummation of the Brahmanical religion" (Olivelle 2008, xix). He is portrayed as an Aryan prince, son of an Aryan king and a prominent *khattiya* (warrior class) member of the Aryan caste system; his biography is thoroughly Brahmanized:

> The young Buddha is represented as the fulfillment of a long line of famous Brahmanical and Vedic ancestors; he is given a Brahmanical *gotta* (family or clan name), Gotama; recognized as a *Mahāpuruṣa* (P. *Mahāpurisa*) by the court *purohita*s (priests) with all the marks of a great man, "handed down in our Vedic mantras" (*āgatāni…amhākam mantesu*); likened to the Vedic gods; and administered the *saṃskāra*s (sacred Vedic rites) starting with the naming ceremony (Rhys Davids 1878, 160; Cowell 1895, 8–9; Olivelle 2008, 15–17, 23)…yet Suddhodana as ruler of the Sakyas was not a king, but an elected leader of a *gaṇa-saṅgha*, a tribal republic (Thapar 2002, 46; Malalasekera 1938/ 2003 = DPPN, s.v. Sakiyā) and Gotama never called himself a prince. His teachings are sometimes presented as a heterodox or reformist development of an earlier Upaniṣadic tradition; yet they are often radically different from orthodox Brahmanical beliefs. Buddhism itself is viewed, not as a separate teaching that may or may not have cultural affinities with the indigenous peoples, but as a set of antitheses to Brahmanical doctrines, or even a schismatic reform movement from within Brahmanism. (Levman 2013: 153–54).[1]

His mixed, cultural background, originating in the indigenous tribes of the sub-Himalayan foothills, is thoroughly camouflaged. By the time these

---

[1] There are four major biographies of the Buddha, all very late: the earliest of these is the *Mahāvastu*, parts of which date perhaps from the middle of the second century BCE. The *Lalitavistara* dates perhaps from the first century CE and is thought to be a work of the Sarvāstivādin school. The *Nidānakathā*, which has a Sri Lankan provenance, was written as a preface to the *Jātaka* commentary and is probably the latest of the four ancient bios. All treat the Buddha as the son of a king (which he wasn't) and the descendant of a long line of *cakkavattin* (*cakravartin*), world-rulers, an ancient Brahmanical legend preserved in the Vedas (it isn't), a legend apparently invented by the Buddhists to curry favour with the Brahmanical establishment (see below page 57 and Levman 2013: 162–65, hereinafter CR). Johannes Bronkhorst talks of this bias in his 2011 book and 2014 article where he describes the Brahmanization of the Buddha's background and the obfuscation of his own culture which he calls "Greater Magadha." He distinguishes the presentation of the Buddha's story in the *Buddhacarita*, which is wholly Brahmanized, from its presentations in the *Lalitavistara* and the *Mahāvastu* which were "quite different", that is, not as overtly Brahmanized (page 318). I take issue with this latter statement (see for example the Brahmanization of Asita in the *Mahāvastu* and *Jātaka Nidāna*, page 57 below), however, it is not particularly germane to this discussion, so for another time.

biographies were written (the earliest parts of the *Mahāvastu*, perhaps 2nd–1st century BCE) the tribes had been thoroughly assimilated into the Indo-Aryan political and cultural hegemony, their past independent history forgotten or recalled only with embarrassment, their customs lost or assimilated, the Buddha's teachings in their language(s) lost. For the Sakyas were historically a Dravidian and/or Munda speaking group—or perhaps more accurately described, based on the proportions of words borrowed into Pāli, as a Dravidian speaking group with a Munda substrate—and the Buddha almost certainly taught them in their own language, but none of his teachings have survived except the loan-words borrowed into Middle Indic. The "intensive interrelations between Dravidian, Munda and Aryan dating from pre-Vedic times" (Kuiper 1948a: 9, hereinafter Kuiper) make it difficult, sometimes impossible to settle etymological and diffusionary questions amongst these languages.[2] This is further complicated by the increasing political and economic dominance of the Indo-Aryan (IA) immigrants (beginning well before the birth of the Buddha), which forced the Sakyas and other tribes of north-eastern India to learn their language as a second tongue; they were Middle Indic as a second language speakers (MISL)—like the twentieth century immigrants to North America, who had their own language but had to learn English to interact with society at large, with the major difference that the Sakyas were the native population and the Indo-Aryans the immigrants. At the time of the Buddha it was a complex and very mixed linguistic culture.

There are hundreds, perhaps thousands, of non-Indo-Aryan words preserved in Pāli, the earliest and only relatively complete record of the Buddha's teachings we possess in an Indic language, a linguistic form very close to the vernacular that the Buddha actually spoke. Other Middle Indic idioms have also survived, but nothing has survived from the indigenous languages of the time. Yet there is no reason to believe that the Buddha only spoke in Middle Indic. There is too much evidence to the contrary.

---

[2] For a summary of some of the substrate and adstrate strata in the India linguistic area at the time of the Vedas and following see Emeneau 1980 (a compendium of his earlier essays by A. S. Dil), Masica 1976, Witzel 1999a: 1–5 (hereinafter Witzel), and Southworth 2005: Chapter Two (pp. 39–61, hereinafter Southworth). The word "Dravidian" is itself a native (*desi*) word, derived from proto-Dravidian **tammiẓ*, "own speech" and *tamiẓ*, "Tamil language or people" with the characteristic change of *-m-* > *-v-* (Southworth 77). For derivation of the word Munda/*muṇḍa*, see Chapter Four. In this book I use the word *desi* to refer to the autochthonous languages of Dravidian, Munda and other tribes. It is also sometimes employed to signify IA Prakrits.

Most of the place names of north-east India are derived from Munda or Dravidian etymons, as are the names for the flora and fauna of the region which were not known to the IA immigrants (Thomas 1931: 23). This is to be expected. More important are the many religious, cultural and political terms of the north-eastern clans assimilated into Middle Indic. This shows more than just word borrowing, but reveals a heretofore untold history of pervasive, indigenous infusion of autochthonous values into the Buddhist belief system. One cannot just view Buddhism as a reaction to or against Brahmanical teachings; this is only part of the story, and perhaps the less important part.

Linguistic scholars have long noted the influence of the indigenous peoples on the incoming IA peoples. Emeneau called this process "Indianization" and attributed bilingualism as the cause (1956: 7). For Southworth the process was facilitated by the need for economic cooperation which he called "village coexistence" (1979: 207–08), where features of the majority language (Dravidian) diffused into that of the minority. According to Krishnamurti Middle Indic was "built on a Dravidian substratum," where the structure of the Dravidian language was imported into and imposed on the IA language by autochthonous speakers who were constrained to accept the dominant IA language as their *lingua franca*; they adopted their vocabulary but kept a lot of the structural features of their mother tongue (2003: 41). The influence of this substrate is apparent in lexical, phonological and structural borrowing from Dravidian into the IA languages (for a summary, see Krishnamurti 2003, 38–42).

How does one take this influence into account when none of the teachings have survived? One method is through the words and what they signify; examining them can prove very revelatory. They uncover, for example, a highly developed indigenous ascetic tradition which the Buddha inherited. The matted hair ascetics (Jaṭilas, a Dravidian word) are looked down upon in the Buddhist *suttas*, but Kaṇha, a Jaṭila, was present at the Buddha's birth and prophesied his future enlightenment. Descended from King Okkāka himself,[3] the legendary founder of the Sakya tribe, Kaṇha is a native seer,

---

[3] *Ambaṭṭhasutta*, DN 1, 93[4–5]: *rañño kho pana Ambaṭṭha Okkākassa Disā nāma dāsī ahosi. sā kaṇhaṃ nāma janesi*, "Disā was a slave of King Okkāka, Ambaṭṭha. She gave birth to Kaṇha." Okkāka is a non-IA name per Kuiper (1991: 7) and see Mayrhofer 1956–76: vol 1: 84, s.v. *ikṣu*, "sugar cane" who says the derivation is "unklar" (unclear, hereinafter M1). Witzel (2009: 90) derives it from Dravidian **iṭ-cu* "sweet juice", after Southworth (218). Mayrhofer 1992–96, vol. 1: 185 says it is "wohl Fremdwort" (probably a foreign word; hereinafter M2), mentioning Berger who takes it from a prefixed Austro-Asiatic (AA) form with the meaning of "bitter squash." The Berger article was not available to me.

despite the fact that Buddhaghosa later tries to Brahmanize him as a court *purohita* (Brahmanical priest).[4]

The *kaṭhina* practice is inherited from the indigenous recluse tradition. The word *kaṭhina* refers to both the wooden frame on which cloth is fashioned to make robes (*cīvara*) for the monks, and, by synecdochic extension, the ceremony itself in which cloth is presented to the Sangha by the laity and robes are made. Both *kaṭhina* and *cīvara* are Dravidian words, as are many of the technical terms used to describe the practice (see below pages 65).

The likening of the Buddha to a snake (*nāga*) or tree spirit (*yakkha*) as a mark of respect, alludes to snake- and tree-worshipping practices amongst the indigenous tribes, a cultural tradition which was assimilated into Buddhism at a very early time. Indeed, right after the Buddha's enlightenment, the serpent king Mucalinda[5] protects the Buddha by coiling around him seven times and covering his head with his hood (*phaṇa*).[6]

---

[4] Pj II, 483[30]: *abhisitta-kāle purohitoyeva ahosi.* "At the time of his (Suddhodana's) consecration, he (Kaṇha) was just the *purohita*." In this book I will divide all compounds with a hyphen (-) to make it easier for the reader to parse them, regardless of whether the PTS edition (which I use wherever possible) employs the hyphen or not. Some volumes do and some do not; there is no standard practice. Proper names and names of *suttas* will not be hyphenated; capitalization of proper names will follow the PTS practice where they are usually capitalized (but not always). Negative compounds beginning with *a-* or *an-* will not be hyphenated. When the parsing of the compound is unclear, I will leave it as in the source text.

[5] Mucalinda is also the name of the tree spirit who inhabits the tree; according to M1 vol. 2: 649 (s.v. *mucukunda*), it is a *Pterospermum suberifolium* tree, which is "wohl unarische Planzennamen" (probably non-Aryan plant names). M2, vol. 2: 360, "nicht klar" (not clear). Cp Santali *mackundạ*, *Pterospermum acerifolium.*

[6] Connected with the word *phaṭā*. Kuiper 1948a: 163 (hereinafter Kuiper) gives the word a Munda etymology (< Santali *puṭi*, "to swell"), while Burrow (1948: 386) a Dravidian one (cp. Kannada *peḍe*, Tamil *paṭam*, "expanded hood of a cobra"). See M1 vol. 2: 389–91, s.v. *phāṭā*, "expanded hood of a serpent." The Munda root has a more extensive distribution than the Dravidian (Sora, *pēl*, "to swell"; Bodo-Gadaba *pulei*, "to swell"; Juang *puli*, "to swell"; Ho *puti*, "to swell"; Sora *puŋ*, "to be swollen"; *puŋ-ʾpuŋ*, "to swell"; Juray *puŋ-dɔɹ-*, "to swell"), which, barring borrowing, tends to suggest its chronological priority; see discussion below on methodology (on page 31f). Dravidian etymologies are taken from Burrow and Emeneau's *Dravidian Etymological Dictionary* (DED) unless otherwise noted. For Munda derivational material see footnote 17.

The *gaṇa-saṅgha* political organization, rule by a council of equals, in contradistinction to the monarchies of the Indo-Aryans, was also a home-born mode of governance, as are the linguistic terms associated with it.

The funeral practices of the native people were also adopted by the Buddhists and had little or nothing to do with Brahmanism.

The Indo-Aryan immigrants were—at least initially—numerically inferior to the indigenous population; presumably this was also the case with the distribution of their linguistic terms. Deshpande notes that non-Sanskrit linguistic usages (that is, Prakrit or *apaśabda*, "vulgar speech") were numerically more than Sanskrit usages (1979b: 9). But this is only part of the story; autochthonous designations for common things often outnumbered the OI and MI terms. Take the famous statement in the *Araṇavibhangasutta* (MN 139) where the Buddha says that one should not insist on local language (*jana-pada-nirutti*) and should not override normal usage (*samaññaṃ*). He lists six synonyms for a receptacle, only one of which might be considered "standard" Indic, *patta*, "bowl" (OI *pātra* < OI *pā*, "to drink" + affix *-tra*, "that by which something is drunk"), and its fem. form OI *pātrī*, P *pāti*, "vessel, plate, dish, pot" (*pātra* > *pātrī* > *patti* > *pāti*); presumably this is "normal usage", the others are "local language", all *desi* terms, that is loan-words from the local languages, either Dravidian or Munda.[7] These outnumber the standard Sanskrit term by five to one. The Buddha's injunction is quite ambiguous and has been interpreted in various ways, but if one goes with Lamotte's explanation (1958/1988: 553), the Buddha is saying that no one single term (e.g. the Sanskrit one) is correct,

---

[7] MN 3, 234[30–31]: *jana-pada-niruttiṃ nābhiniveseyya*, *samaññaṃ nātidhāveyya*; the other words are 1) *vittha/vitta*, "bowl" (OI ?) < etymology unknown; 2) *serāva/sarāva* (OI *śarāva*), etymology not understood per M1 1956–76, vol. 2: 307 ("nicht befriedigend erklärt," not satisfactorily explained). Possibly from Dravidian Tuḷu *teriya* "circular pad of wicker or straw placed under a vessel to make it steady" by metonymy and change of *t- > s-* (no *s-* in proto-Dravidian). Or from Munda, cp Santali *soṛwa* or *soṛha*, "leaf cup"; 3) *dhāropa/harosa* (var.) "bowl, dish, pan" < etymology unknown ? < *dhṛ*, "to hold"? Cone: a dialect word for a bowl, a dish; *hapax legomenon*, cp Kannada *ḍoppe, doppe*, "cup or dish of leaves"; 4) *poṇa/hana* (var.), "pot," *hapax legomenon* < unknown etymology; cp Gta' (Munda language) *boṛna*, "small pot"; 5) *pisīla/sīla/pipila* (var.) Skt *piśīlam*, "wooden vessel"; Mayrhofer: "nicht genügend erklärt" (not sufficiently explained), perhaps from root *piś*, "hew out, carve out, cut into shape" *hapax legomenon*; more likely non-Aryan, cp Dravidian Tamil *patalai*, "large-mouthed pot."

but one may adopt the term in use in the region one is in (see also Levman 2014: 110 for a different interpretation).

This book is divided into three sections. Section One consists of Chapters Two to Five, all of which deal with certain aspects of the linguistic scene and what they reveal about the political, cultural and religious roots of Buddhism in the autochthonous population, one of the most unexplored, opaque areas of Buddhist study.

In Chapter Two I will provide a general introduction to the linguistic scene in the fifth and fourth centuries BCE, both the indigenous languages and the Middle Indic dialects which flourished at the time, and how they interacted and affected each other. I will provide examples (from the Vinaya) of linguistic mixing of Pāli and indigenous languages to show the reader some of the general theoretical and methodological principles of this section of the book. A detailed methodology on how to determine whether a word has been borrowed into MI is also introduced and discussed.

Some of this material I have already discussed in my 2013 article "Cultural Remnants of the Indigenous Peoples in the Pāli Scriptures" (hereinafter abbreviated as CR) which provides a great deal more detailed information on what Emeneau calls the "India Linguistic Area" (1954), the complex linguistic scene of local languages and Middle Indic dialects which interacted at the time of the Buddha.

Chapter Three expands on this article with many more details from the historiographic and linguistic side, looking at what we can piece together of the indigenous *samaṇa* (recluse) tradition, including the above-mentioned Jaṭilas, some details on the *kaṭhina* practice, and more information on the burial customs. I will also look at the names of all the villages the Buddha visited during the last few months of his life as they provide a lot of information on the cultural and linguistic mixing at the time the Buddha lived (see, for example, Appendix Three on page 149 which discusses the town Bhaṇḍagāma and the meaning of *bhaṇḍa*). I also try to deal with some potential objections to my overall thesis–that linguistic mixing is isomorphic with cultural mixing–on page 133.

Chapter Four continues the discussion on Dravidian influence on OI and MI in terms of syntactical structure. Here I look at the "Tamilization" of Pāli, an expression coined by Koenraad de Vreese in 1980 to describe certain grammatical correspondences between Old Tamil and Pāli which he attributed to the influence of the former language. Here I also look at the

similar structure of complex sentences between the two languages, composed of strings of non-finite verbs.

Chapter Five is an article on the meaning of the non-Aryan word *muṇḍa* ("bald") in the Pāli scriptures, reprinted from the *Canadian Journal of Buddhist Studies*, 2011. Scholars are generally agreed that the word has a non-IA source, but differ as to its various meanings. I argue that it was used as a pejorative name for a non-Aryan ethnic or tribal group—the Buddha was called a *muṇḍaka* as a form of insult—and demonstrates the disapprobation of the IA immigrants for the indigenous tribes.

Section Two is about the languages the Buddha spoke, that is, the languages we know of that have been handed down in the oral tradition. Ever since Buddhaghosa announced that the Buddha spoke the language of Magadha (Māgadhī), which he considered identical to Pāli, this has been a controversial subject. Most Theravādin monks take Buddhaghosa at his word and believe that Pāli or Māgadhī was the actual language that the Buddha spoke. But starting with Sylvain Lévi's important article in 1912, in which he identified "*une langue précanonique*" underlying Pāli, a scholarly consensus has emerged that has uncovered a pre-Pāli layer in the canon inferable by comparing cognate parallel passages in the different Indic transmissions. This earlier layer is closer to the language the Buddha spoke and has been characterized as a *lingua franca* or *koiné*, that is, a common inter-language of trade and commerce, what is called a *vohāra* in Middle Indic (< *vy-ava-hṛ*, "to carry on commerce, trade, deal in"). As a result of its *vohāra* heritage, Pāli shows a mixture of many Indic dialects from the north, east and west, as well as a large number of borrowings from the indigenous Munda and Dravidian languages; and to further complicate matters Pāli contains a lot of Sanskritizations, where backward-looking Old Indic elements are introduced into the idiom, dovetailing with a general trend towards the desacralization, secularization and politicization of OI from around the time of Asoka. Although the vernacular, colloquial character of the pre-Pāli idiom was never lost, many artificial OI elements were added in, in addition to the mixed, synthetic character of the *koiné* itself.

Although Pāli was not the language the Buddha spoke, it was close; and it is the oldest record we possess, closest to the time of the Buddha. Von Hinüber, for example, dates the core parts of the *Mahāparinibbānasutta*, which tells the story of the Buddha's death, to within sixty years of his death (2009b: 64). And he has this to say about those who doubt the historicity of the Pāli tradition:

During the past decades the age of the Theravada tradition has been thrown into doubt occasionally, not rarely in a rather general and sweeping way. However, wherever it is possible to use comparatively hard arguments that is to say linguistics, it becomes soon more than obvious that it is possible to dig considerably deeper into the past here than in any other tradition. Even though the old Pāli texts are created out of a Buddhist Middle lndic, and, consequently, nowhere preserve, but at best reflect the language of the earliest Buddhism, they contain the earliest redaction of Buddhist texts, linguistically near to the Asokan inscriptions at Girnar, followed by the Mahāsāṃghika (-lokottaravāda) and of course the Dharmaguptaka texts in Gāndhārī. This concerns first of all the age of the redaction, which also protects the content. On the other hand, revisions such as a change of language, e.g., from Middle Indic to Sanskrit, always opens the opportunity to introduce new concepts. In this context it is remarkable that new concepts sometimes found their way only into the Theravada commentaries, while they still could be included in canonical scriptures of other traditions, which points to a rather early closure of the Theravada canon…Therefore, T. W. Rhys Davids and H. Oldenberg were not at all wrong in their judgement when pointing out the comparatively high age of the Theravada tradition, which, however, does not mean that all old material is preserved only there, and that all other tradition are necessarily recent in each and every respect, but only that the roots of the Theravada tradition reach much deeper into the soil here and there than elsewhere (ibid, page 48–49, footnote 43).

While I would disagree with Prof. von Hinüber that the Pāli texts "nowhere preserve, but at best reflect the language of the earliest Buddhism" (see Levman 2019a, and Chapters Five and Six of this volume), it is undeniable that Pāli has undergone a lot of changes from the pre-Pāli idiom which the Buddha and his immediate disciples used.

The change of the language of the Buddha and its evolution into Pāli until it was at least partially fixed in the first century BCE by writing it down was caused by a number of factors, including normal diachronic language change over time, synchronic change from natural dialect variation, artificial Sanskritizations, and borrowings from the indigenous languages, to name the four principal factors. Some still claim that Pāli was the actual language the Buddha spoke—despite so much evidence against it—and maintain that all variation found in Pāli is due to "natural variation and transmission errors" (Chapter Seven), denying Sanskritization and diachronic change completely. I also have friends in the monkhood who are doubtful that Pāli wasn't the language the Buddha spoke, as stated by Buddhaghosa, for, with so many monks reciting the Dhamma, they don't understand how change would have been introduced. This would make Pāli the only *dhamma*, linguistic or otherwise, where the law of *anicca* did not

apply. Linguistic change, especially in an oral environment where there is no written guide to adhere to, happens naturally over time, according to regular linguistic laws to which Pāli was not immune. To argue that this is all "natural variation" is also to miss the point; this is like the Ājīvaka heretics, contemporaries of the Buddha whose leader Makkhali Gosāla argued that everything in the universe was random, and there were no laws, including no law of karma that led to either liberation or affliction. Certainly there is such a thing as random change, but the priority of *paṭicca samuppāda*, that is, change according to specific causes and conditions, is not only a linguistic, but a universal law. Language too must follow the *ti-lakkhaṇa*, being impermanent, changing according to causes and conditions; inaccurate and unsatisfactory, for it posits things (like the "I") which do not truly exist and are without a permanent essence. An unchanging Pāli as the language the Buddha spoke is another form of essentialism, which the Buddha eschewed.

In my 2016 paper "The Language of Early Buddhism" I outlined the methodology and argument for identifying a *koiné* or inter-language as the earliest recoverable language of Buddhism, and probably the language the Buddha spoke—or at least one of the languages he taught in—that later, through the processes noted above, evolved into Pāli. It is reprinted here as Chapter Six.

An article entitled "Sanskritization in Pāli" was originally planned for this chapter, but in the interim it was accepted for publication by the *Journal of South Asian Languages and Linguistics* (2020b). The article examines cognate correspondence sets of the *Dhammapada*, comparing Pāli verses with parallel Gāndhārī, Sanskritized Prakrit and Sanskrit verses, confirming the existence of the *koiné*, the interpretation of which is in many cases the source of the phonological differences in the transmission.

Chapter Seven is a new discussion on the "Evolution of Pāli" which picks up on an article entitled "The Language the Buddha spoke" published in 2019 in the *Oxford Journal of Buddhist Studies*. That article answered Stefan Karpik and Richard Gombrich's assertion that the Buddha spoke Pāli.

Section Three examines the transmission and evolution of the Buddha's teachings over time. This is a very complex process involving multiple vectors and tens of thousands of protagonists over many centuries. Outside of Pāli, the earliest renditions we have preserved of Buddhist teachings is in Gāndhārī, the language of the Gandhāra kingdom in the north-west (parts of present Afghanistan and Pakistan); fragments of Gāndhārī manuscripts

have been dated to as early as the first century BCE, making them the oldest Buddhist manuscripts in existence (for comparison, the oldest Pāli manuscript is dated to approx. the ninth century CE). Evidently the Buddha's teachings were transmitted along the silk routes, both northwest and northeast, starting from the lifetime of the Buddha, while at the same time the teachings diffused within India along established trade routes to the west and south. The Asokan edicts of the mid third century BCE, inscribed 100–150 years after the Buddha's *parinibbāna*, bear witness to at least three principal dialect areas in India at the time, in the northwest, west and east, and we find a mixture of all these dialect features in Pāli, especially from Girnār in the west and Gandhāra in the northwest. In the middle of the third century BCE Asoka's son Mahinda brought the teachings and commentary to Sri Lanka and by the late first century BCE the canon was written down there and at least partially fixed. While the island was being converted to the Buddha's teachings, they continued to spread on the continent through central and southern India, monumentalized in art and votive inscriptions at such famous sites as Sanchi, Bharhut, and Amāravatī in the last three centuries BCE.

The MI lineage as it manifested in China has been an area of continued interest to scholars. In 2018 I published a study on the transmission of the *Buddhadhamma* from India to China, studying the *dhāraṇī*s (sacred formulae for the retention of the Buddha's teachings) of the *Lotus Sutra*; examining the Chinese transliterations of the *dhāraṇī*s showed that the earlier work was composed in a north-western Prakrit similar to Gāndhārī. A north-western source is also probably true for the earliest translations of the *Buddhasāsana* into China which date from the mid-second century CE and likely earlier; the teachings were no doubt transmitted to the country along the eastbound trade route to Chang'an well before the common era, at least in inchoate form. A new study of the *Milindapañha* ("Questions of King Milinda") examines the nature of what was transmitted before the first full-fledged translations of the Buddhist *suttas* into Chinese took place: a short summary of the basic tenets of Buddhist doctrine in a question and answer form—a Buddhist "catechism" that was used as a basis for improvisation and expansion by the Chinese translators, rather than word-for-word translation. This study was originally planned to form the first chapter of section three of this monograph, but has now been published in the *Journal Asiatique* ("Revisiting *Milindapañha*," 2021b). Comparison of the Chinese version of the Pāli *Milindapañha* shows that the former is one of the earliest of Chinese renditions, showing a lack of knowledge of basic Buddhist theory, and reflecting an early time before the Chinese standardization of a common Buddhist terminology. So the Chinese

translation, although nominally dating from the third century CE, is based on a work which is much earlier than this and may represent one of the first and earliest attempts at rendering the Buddhist doctrine into Chinese.

That language changes over time is an incontrovertible fact. Usually this statement refers to the evolution of languages' phonologies, i.e. sound-systems, which evolved according to regular rules discovered by the Neogrammarians in Germany in the latter part of the nineteenth century. It is Darwin's law of descent with variation in the field of language; however, whereas biologic evolution changes due to mutation and adaptation to changing environmental and biotic conditions, language evolves based on other principles like lenition, assimilation, vowel loss and merger, borrowing etc., which we see in Middle Indic as it evolved from Old Indic. Not only does phonology change, but meaning as well. A very common phenomenon is the broadening of a word's meaning over time through metaphorical extension. So, for example, a word like "metaphor" originally only meant "to carry from one place to another, to transfer" (from the Greek *μεταφέρω*, *metaphéro*) but gradually developed an extended meaning in the field of rhetoric as a "transferring to one word the sense of another." So it is used in English today, and while it has lost the meaning of a physical transference, as can be seen, the core meaning of "carrying over" is still there.

The word *sati* in Pāli has undergone a similar phenomenon except that it has kept its core meaning (< OI *smṛti*, "remembering, recollection"), while developing additional meanings in the Buddhist context ("present awareness, mindfulness, alertness, etc."). Some scholars have argued that *sati* no longer has the meaning of "memory" in Buddhism, but a careful study of the texts shows that its core meaning of "memory" is ever present, always in the sense of remembering the Buddha's teachings. Chapter Eight is about the varying usages of the word *sati* in the Buddhist scriptures as understood in the Burmese tradition. I choose the Burmese tradition for two reasons: firstly, because they have the largest and most comprehensive Pāli dictionary in the world, and secondly because the Burmese school of mindfulness training is the lineage source for much of the present mindfulness movement in North America; that movement's understanding of the word *sati* or "mindfulness" is often quite narrow in relation to the Burmese understanding of the term.

Pāli is a phonetic language, so theoretically what you see is what you get. Generally that is true, but there are some aspects of Pāli which are problematic in pronunciation as they are in Old Indic. I am referring to

nasalization, called *anusvāra* ("after sound") in OI and *niggahīta* ("checked") in Pāli.[8] No one is quite sure how this element is pronounced in OI or MI (represented by *-aṃ*). This is important for several reasons. Although Pāli is no longer a spoken vernacular, it is still recited by monks every day, carrying the message of the Buddha's teachings; one would like to ensure that it is chanted the way the Teacher himself intoned it. In addition, Pāli is used for official declarations and resolutions within the Buddhist Sangha (i.e. for *kammavācā*s); these are supposed to be pronounced properly so as not to invalidate the proceedings, like for example, the ordination formula of a novice monk or layperson, *Buddhaṃ saraṇaṃ gacchāmi, dhammaṃ saraṇaṃ gacchāmi, saṅghaṃ saraṇaṃ gacchāmi.* There are several possibilities for pronouncing the nasal and various people do it in various ways. Is there a standard? This is also important because in the change from OI > MI, nasalization increased quite markedly as a form of lenition, that is, various inflections in OI ending in *-am* changed to a nasalized vowel ending (*-aṃ*) in MI, so *Buddham paśyāmi* ("I see the Buddha") in OI became *Buddhaṃ passāmi* in MI with the *-am* changing to *-aṃ*; that is the nasal consonant *-am*, pronounced as written in OI—*Buddham*—changes to a nasalized vowel *-aṃ* pronounced *Buddhã* (as in French *enfant*) in MI, or pronounced as a velar nasal *Buddhaŋ* ("Buddhang") depending on what consonant follows in the next word, or sometimes pronounced simply as a normal nasal consonant (as in OI, as *Buddham*). Hardly a sentence in Pāli goes by without encountering this *-aṃ niggahīta*, so one would like to ensure that it is pronounced correctly.

Chapter Nine discusses the different possibilities and offers some suggestions as to the "correct" pronunciation, although the article makes the point that this is quite a sticky wicket, as we have no live recordings of the pronunciation from two thousand five hundred years ago and the grammarians do not describe the phenomena in enough detail to extrapolate a definite answer.

Chapter Ten is a short summary of all these different themes. One inescapable conclusion is the importance of language, and especially Pāli, in understanding the roots of Buddhism and how it evolves. This echoes Norman's observation in the 1994 *Bukkyō Dendō Kyōkai Lectures*, that any work on Buddhism must take into account philology, and must be based on a full understanding of the linguistic sources (1997/2006: 229).

---

[8] Because the articulatory organs are not completely closed as they are in a normal consonant, and air is forced through the nose by a partial lowering of the soft palate.

In Pāli itself lie buried several hundred or perhaps thousand autochthonous words and technical terms, and in those one finds a rich tapestry of indigenous customs, culture and religion; this field, what one might call "linguistic archaeology," is still very much in its infancy and the account I have given of it here in Section A is merely a beginning.

Comparing Pāli to other Middle Indic dialects and Old Indic, while taking into account the effect of word borrowing and the different phonemic structures of native languages, provides insight into the nature of diachronic language change over time and the synchronic influence of the local population. The process allows us to make some meaningful statements about the pre-Pāli idiom the Buddha taught in, or at least the earliest recoverable language of Buddhism, its evolution, Sanskritization and standardization.

Finally the diachronic study of Pāli and the Prakrits over time opens up a possible new paradigm on the origin of the early Chinese translations; clarifies the modern day interpretation of the word *sati* by examining its multi-valent usage by the Buddha and his followers; and helps one to understand the pronunciation of the nasalized phonemes in both ancient and modern times.

SECTION ONE

THE LINEAGE

PARAMPARĀ

CHAPTER TWO

THE LINGUISTIC SCENE
AT THE TIME OF THE BUDDHA

Abstract

Chapter Two provides an overview of the linguistic scene at the time of the Buddha where, initially at least, the autochthonous peoples far outnumbered the newer Aryan immigrants. This resulted in significant structural influence on the Indo-Aryan languages and extensive word borrowing which is found in the names of local biota, toponyms, proper names and most importantly, in the names of certain cultural, technical and religious phenomena which were carried over from the local culture into Buddhism. The entire robe practice of the local *samaṇa* (mendicant) culture, for example, was adopted by the Buddhists. The chapter discusses the implications of this borrowing and provides a comprehensive discussion on methodology, that is, how to determine whether a word is borrowed from the indigenous languages into IA or from IA into the indigenous languages; it gives an example of the complexities of the process with a discussion of the etymology of the word *amba* meaning "mango."

—

The early history of Buddhism is biased from the very start. Not just because its history was written centuries after the Buddha had passed, by a small group of Brahman converts who had their own story to tell, but also because so much of the early history is missing.

Linguistically we have only one of the many language groups that were current at the time the Buddha lived: Middle Indic and its various dialects, the most important of which is Pāli, as it is the only dialect which preserves a relatively complete record of the Buddha's teachings. But we know that the languages spoken in fifth and fourth century BCE were much more varied than just MI and its various dialects; in fact MI speakers at that time were certainly the minority, with the indigenous languages forming the plurality. Two of the principle language families at that time were Dravidian

and Austro-Asiatic (AA); Dravidian still forms the dominant language family of south India to this day (Tamil, Kannada, Telugu, Malayalam, to name the four most populous ones), comprising over 200 million speakers, while Munda speakers (Santali, Mundari, Kharia, etc., part of the Austro-Asiatic language family), although reduced to less than ten million speakers today, were at one time along with Dravidian a dominant linguistic group in north-eastern India. The Tibeto-Burman language family was also represented as was proto-Burushaski, a language isolate, plus an unknown substrate which Masica calls "Language X" (1979).

Judging from the etymology of hundreds of key words—toponyms, personal names, names for local biota and cultural terminology—the Sakya tribe in which the Buddha was born was Dravidian and/or Munda speaking, or perhaps more accurately described as Dravidian with a Munda substrate, based on the relative etymological proportions of the native words that are preserved. The reader will have to forgive the vagueness of the terms "Dravidian and/or Munda" for a description of the language of the Sakyas (and other north-eastern tribes) which can often not be resolved at a finer scale; the only record we have of this language are those loan-words preserved in OI and MI which are indeed very mixed; some can be demonstrated to be AA in origin, others Dravidian, and others from language isolates like proto-Burushaski, and still others from unknown sub- or adstrate layers, sometimes called languages of the Indus Valley (Witzel §6; Southworth, Chapter Three). The language is extremely complex, layered and commingled as will be demonstrated in detail in the following chapters. Southworth's pioneering study on linguistic archaeology hypothesizes that the eastern, sub-Himalayan India where the Buddha lived and taught was primarily Munda and Tibeto-Burman—speaking of the middle of the first millennium BCE (2005: 329, Figure 10.2). However, this study shows Dravidian to be the dominant language, based on words preserved in the Pāli *suttas* and in the local place names. Munda words are also represented, but less so than Dravidian, and Tibeto-Burman influence is very limited. In addition, the tribes spoke a form of Middle Indic; in order to interact with the increasingly hegemonic culture of the IA immigrants, they had to learn their tongue, as they certainly did as a second language. But they were outside the Indo-Aryan fold, of "mixed" heritage" (*saṃkīrṇa-yonayaḥ*), living well beyond the confluence of the Ganges and Yamuna rivers, which was the eastern limit of the Aryan influence at that time (Bronkhorst, 2007: 1–9).

In fact there was a Middle Indic language of trade, commerce and government (a "*Verkehrssprache*" and/or "*Kanzleisprache*") which appears

to have been common to all the varied linguistic groups; it has been variously termed a *koiné* (inter-dialectic) or a *lingua franca* (inter-linguistic) by scholars and it is from this idiom that Pāli evolved (see this story in Chapters five and six). The Buddha likely spoke in this idiom and the dialect that developed out of it, Pāli, is the earliest record we have of his teachings. But he certainly also spoke to his own people, the Sakyas, who initially at least comprised a large number of his converts, in their own language, and none of that has survived. What has survived however are many of the words from the Sakyan language which were imported into Middle Indic. Part of the reason is just practicality. They provided names for biota and places that the Indo-Aryans were unfamiliar with; more important is what they represent in terms of borrowings from the culture of the lender: a rich religious and socio-political culture that was adopted directly into Buddhism. The source of these ideas have been completely obfuscated by later tradents, often Brahmans writing for other Brahmans who were out to show that the Buddha was the best of all Brahmans—and not a *mlechha* (OI, "foreigner, barbarian"; P *milakkha*)—though the whole notion of the IA class system was foreign to his or the Sakyas' thinking.

It is well known that there were many *samaṇa* (ascetic) groups in north India at the time of the Buddha. Even when some of their practices, like *jhāna* meditation, were adopted, the Buddhist Sangha is represented as opposing the customs and beliefs of these groups, yet many of their customs and practices were adopted and adapted directly into Buddhism. The technical language used is not Middle Indic, but Dravidian or Munda, that is, indigenous. It is almost as if the Middle Indic exemplar, Pāli, has been translated from a local language with all the technical terms kept intact in the original language as a common, shared cultural heritage.

I will be giving hundreds of examples of cultural continuity in the next few chapters. Here I would like to provide a few examples of what I am saying with reference to the robe practice in Buddhism. This was one of the four requisites provided by the laity to the monks, i.e., food, lodging, robes and medicine.[9] The practice of donating cloth and making it into robes for the

---

[9] One must think of these as categories of requisites as both the *Samantapāsādikā* and the *Sumaṅgalavilāsinī* provide for eight possessions which a monk was allowed:
    ticīvarañca patto ca, *vāsi sūci ca bandhanaṃ.*
    parissāvanena aṭṭh-ete, *yutta-yogassa bhikkhuno.* (Sp 1, 240; Sv 206)

monks was not invented by the Buddha; it was taken over from the local *samaṇa* culture. The key words *kaṭhina* (the structure on which the cloth was stretched prior to being cut and sewn), *dussa* (the cloth itself) and the name of the robe (*cīvara*) are all of Dravidian origin (see page 62f). So are the instructions on how to make the robe; without divining the words' indigenous etymologies it is almost impossible to understand the instructions. Of course none of this takes away from the originality and brilliance of the Buddha's teachings about liberation; but understanding his roots opens up a whole new perspective on the Buddha's antecedents in the indigenous Indian culture, his noble lineages (*ariya-vaṃsā*), which are *aggaññā rattaññā vaṃsaññā porāṇā asaṃkiṇṇā asaṃkiṇṇa-pubbā* ("ancient, longstanding, traditional, primeval, pure and unadulterated now as then…" Woodward 1933/2001, vol. 2: 31, AN 2, 27[16–7]).

An Example

Once, when the Buddha was en route to Dakkhiṇāgiri ("Southern mountain") in Magadha he observed how the fields were laid out and that gave him an idea for how to make the monks' robes. He said to Ānanda:

Three robes and a bowl, a knife, needle and a waistband

With a water strainer, there are eight, which a monk who is properly equipped possesses.

Of these eight, five and possibly six are indigenous terms. For the *cīvara* etymology see discussion below on page 62f; a small hatchet or *vāsi* can be traced to a Dravidian source (cp Tamil *vai*,*"sharp"*; *vaci*, "to cut"; Kota *vac*, "skewer"; Kannada *basi*, "point, sharpen"; Telugu *vasi*, "nail, thorn"; Kui *vast*, "sharpen"; Burrow 1948: 393). M2 vol. 2: 548 wonders whether the word can be derived from late Avestan *vāsī*, "pointed knife"; the needle or *sūcī*, traditionally derived from the IA root *siv*, "to sew" and related to *śūka* ("awn of a grain; spike of an insect" M1 vol. 3: 363), originates from a Finno-ugarian source of the word for both OI and Dravidian (per Burrow 1946: 28–29); the Munda languages can also lay claim to the word (and also may have received it from the same source) as it is widespread there: in Santali *sui*, Mundari *sui*, Kharia *cuci*, Juang *cuṇci*, Korku *suwi/suwa*, all meaning "needle"; the *patta* or bowl is IA (< *pā* "to drink" + instr. suffix *-tra*, "that by which is drunk"), as is *parissāvana* ("water strainer" < *pari* + caus. of *sṛ*, "to flow", literally "to cause to flow around/through. Note the similarity to the Tamil word *vaṛi*, "to flow" and the Koṇda verb *vaṛ*, "to drip down as through a filter"; *bandhana* (= *kāya-bandhana* or "waistband") lit., "binding [round] the body") may be < OI root *bandh*, "to bind"), or more likely (as the meaning fits better), from the Dravidian word for "stomach, belly" (cp Tamil *paṇṭi*, *paṇṭam*, "belly, paunch, body") + suffix (*-anai*, "touching"?), misconstrued as an IA word because of the similar phonology. In proto-Dravidian voiced stops were allophones of their unvoiced counterparts.

passasi no tvaṃ Ānanda Magadha-khettaṃ acci-bandhaṃ pāli-bandhaṃ mariyāda-bandhaṃ siṅghāṭaka-bandhan ti?
evaṃ, bhante.
ussahasi tvaṃ Ānanda bhikkhūnaṃ eva-rūpāni cīvarāni saṃvidahitun ti ussahāmi, bhagavā' ti. Vin 1, 287[10–13].

"Ānanda, do you see this Magadha field which *is acci-, pāḷi- mariyāda-siṅghāṭaka-bandhaṃ*?"
"Yes, Bhante."
"You should try to prepare robes of such a kind for the monks."
"I will try Bhagavā."

Now various attempts have been made to decipher what the Buddha meant.

Pāli, Vin 1, 287[11]	Horner 1938–66/ 2001–07: 1832, hereinafter Horner	Rhys Davids & Oldenberg 1882: 207–208	von Hinüber 2006: 7
acci-bandhaṃ[10]	laid out in strips	divided into short pieces	laid out in squares
pāli-bandhaṃ	laid out in lines	divided in rows	laid out by dams
mariyāda-bandhaṃ	laid out in embankments	divided by outside boundaries (or ridges)	laid out by embankments
siṅghāṭaka-bandhaṃ	laid out in squares	divided by cross boundaries	laid out by cross roads

The commentary is as follows (Sp 5, 1127[4–9])

acchi-baddhan ti catur-assa-kedāraka-baddhaṃ

acchi-baddhaṃ = put together in a square field

pāli-baddhan ti āyāmato ca vitthārato ca dīgha-mariyāda-baddhaṃ.

pāli-baddhaṃ = laid out in long borders by width and length

mariyāda-baddhan ti antara-antarā rassa-mariyāda-baddhaṃ.

mariyāda-bhaddhaṃ = laid out by short boundaries at intervals

---

[10] With variants *acchi* and *baddhaṃ*.

siṅghāṭaka-baddhan ti mariyādāya mariyādaṃ vinivijjhitvā gata-ṭṭhānena siṅghāṭaka-baddhaṃ, catukka-saṇṭhānan ti attho	*siṅghāṭaka-baddhaṃ* = laid out in the form of a square, having crossed one boundary with another at the place where they meet; the configuration of a square is the meaning.

acci (acchi)-baddhaṃ

The word *acci* is neither derived from OI *arci* ("flame" < *arc* "to shine") nor from OI *akṣi* ("eye" P *akkhi*) as the meanings are obviously wrong. It is possible that it is derived from OI *akṣa* ("die for gambling, cube") the eastern MI form of which is *akkha*; however P is showing the western conjunct form here, *acch-*, in one of its variants, with an *-i*, not an *-a* ending. The meaning does seem to be "square." As is often the case, the clue comes in the commentary where the *acci/acchi* is described as a type of *kedāra*, "field" (OI and P idem). *Kedāra* is a Dravidian word.[11] Looking further we find that *acci/acchi* is almost identical to a Dravidian word *accu* (cp Tamil, "ridge in a field"; Old Telugu inscription, *accu kaṭṭu sēyu*, "to make a ridge along the boundary") with the expression *accu-kkaṭṭu* as a verb "to form a ridge around a field" and the noun "field with ridges, land divided into beds to admit and retain water for the cultivation of paddy" perfectly describing what the Buddha saw in his journey. Except that here we don't have the verb *-kkaṭṭu*, "to build" as the second member of the compound but what appears to be a perfectly good IA root *badh* ("arrange, put together") in past participle form. Or perhaps we are simply witnessing the common Dravidian verb *paṭu*, *paṭṭ* (e.g. Tamil, "do, make, establish, cause to grow") adopted to IA phonetics with the change of *p- > b-* which is lacking in proto-Dravidian and the weakening of the intervocalic retroflex voiceless stop to a voiced stop, *-ṭ- > -ḍ-*, which is also a common phenomenon in proto-Dravidian (Zvelebil 1990: §1.7.2) and in Middle Indic (Pischel §198; hereinafter Pischel). Since *baḍḍa/u* did not mean anything in MI, it was changed to the closest meaningful word, *baddha.*

What the Buddha was describing was an indigenous field, in indigenous language. This became the platen for the manufacture of the monks' robes.

---

[11] M1 vol 1: 265 "wohl dravidisch" (probably Dravidian); Burrow 1948: 375, compare Kannada *kesaṟu*, "mud, mire"; Tamil *kaitai*, "paddy-field"; Tuḷu *kēḍu̥*, *kesaṟu̥*, "mud, mire" and others.

The other four expressions are the same, they are Dravidian terms, descriptions of the fields that were cultivated at that time in Magadha.

pāli-baddhaṃ

Pāli is a Dravidian word meaning "row, line, margin, edge" (cp Malayalam *pāḷi*; Kannada *pāṛi*, "row, line, regularity, regular order or way, method rule", Telugu *pāḍi*, "justice, propriety, nature, quality"). In this particular case it simply refers to the boundaries between the fields, laid out in rectangles and squares.

In the sacred Buddhist texts, by metonymy the term is transferred to designate the language as a whole. Dozens of attempts have been made to derive the word from an IA source—since Pāli is an Indo-Aryan language—and none are convincing. I have discussed all this in detail in Levman 2020a: 1–3. The latest attempt is to derive Pāli from the verb *paṭh* ("to recite," Gombrich 2019), but the two words are phonologically quite distant from each other and in any case, *paṭh* is itself a Dravidian word (see page 76 below).

The word *pāli* in OI has several other meanings, in addition to "row, line, margin, etc.," that is, "tip or lobe of the ear; dam, dike." These are apparently adoptions and conflations of other similar sounding Dravidian words, viz., Tamil *pālam*, "bridge, jetty, dam"; Malayalam *pālam*, "bridge over rivers or to connect the walls of compounds; Kota *palm* "bridge"; Toda *polm*; Kannada *pāla*, idem; Koḍagu *pala*, "bridge made of trees." In the meaning of "tip of the ear," cp Kannada *bāvali/bāvuli*, "ornament worn on the tip of the ear"; *bāli*, "earring. "Note the change of *-āva- > -ā-* in Kannada, which presumably also happened in the other south Dravidian languages. This was also a peculiarity of IA historical phonology (von Hinüber 2001: §145). For syllable contraction in proto-Dravidian see Krishnamurti 2003: 96, §4.3.2.1.

Is it possible that Pāli was so-named because the teachings were translated from an original Dravidian tongue? No one has even asked the question or considered the possibility. Passages like these we are examining here (and we will be looking at dozens of others) give one pause to consider what their significance might be.

mariyāda-baddhaṃ

The OI word *maryāda* (P *mariyāda*, "frontier, limit, boundary" < *maryā*, idem) has no convincing IA etymology (M1 vol. 2: 597; M2 vol. 2: 331, "nicht zufriedenstellend erklärt," not satisfactorily explained). Przyluski (1931: 613) derives the word from the Austro-Asiatic word for "ocean" (*maru* and **mari*). It is closer to home in the proto-Dravidian (PD) word for "boundary, limit, shore," a descendant of which is found in modern Tamil, *varai*, ("small ridge as of paddy field, bank, shore, limit, boundary, measure, extent, place, time"); Malayalam *varu*, "boundary, border"; Kannada, *vari*, *vare*, "limit"; Telugu, *vaṟa*, "limit." The first definition fits perfectly with the present context. As is well known, in proto-Dravidian *m* and *v* interchange initially and intervocalically (Zvelebil 1990: §1.7.8) and (perhaps under Dravidian influence) this interchange is not uncommon in MI dialects (Pischel §251, 261 and Brough 1962: §36 for Gāndhārī). The suffix *-da* seems to be null as to meaning.

siṅghāṭaka-baddhaṃ

The OI word *śṛṅgāṭa* (P *siṅghāṭa*, "*Trapa bispinosa*") is the name for a local plant, but somehow got the meaning "a place where three or four roads meet" especially in the *-ka* suffixed form which we have here. Presumably this was because of the triangular shape of the leaf which changed from "dreieckiger Platz" (triangular square or place) to "Wegkreuzung" (crossroads, junction; M1 vol. 3: 370–71, "unklar; wohl Fremdwort" unclear, probably a foreign word). This seems to be a Dravidian derivation, cp Tamil names for *Trapa bispinosa*: *cimkhara*, *cinkara*, *cinkara-k-kottai*, *cinkarak kottai*, *cinkarakkottai*, *cinkari*, *inter alia.* [12]

These four compounds—*acci (acchi)-baddhaṃ*, *pāli-baddhaṃ*, *mariyāda-baddhaṃ* and *siṅghāṭaka-baddhaṃ*—all basically mean the same thing, that is, a field arranged in square or rectangular rows. The phrases follow the waxing-syllable principle of four-four-six-six for emphasis and mnemonic memorability, each looking at the field from a slightly different point-of-view—squares, boundaries, margins and crossroads—in keeping with the

---

[12] Found at
http://envis.frlht.org/plantdetails/49608330a08d30e9f1a01da7faf2ed79/5aad36299 e8ec84ad14a8e1dcfceffd2 accessed June 10, 2020. These seem to correspond to DED 2063 *ciṅka-vāṟai*, *ciṅkaṉ-vāṟai*, called "a variety of plantain." Proto-Dravidian had no *s-* consonant.

Buddha's de-centered approach to language and use of near-synonymic repetition (Levman 2020a: 20).

The Buddha uses this field-geometry as the model for robe-tailoring. Given his instructions by the Buddha, Ānanda then makes up the robes for which he is praised for "understanding the meaning in detail of that which the Buddha only spoke in an abridged fashion" (*mayā saṃkhittena bhāsitassa vitthārena atthaṃ ājānissati*, Vin 1, 287[21–2]):

> *kusimpi nāma karissati, aḍḍha-kusimpi nāma karissati, maṇḍalampi nāma karissati aḍḍha-maṇḍalampi nāma karissati, vivaṭṭampi nāma karissati, anuvivaṭṭampi nāma karissati, gīveyyakampi nāma karissati, jaṅgheyyakampi nāma karissati, bāhantampi nāma karissati*

> Ānanda will make what is called a *kusim*, also an *aḍḍha-kusim*, etc.

Once again, we are dealing with technical terms whose meaning is not immediately obvious. There are nine terms:

Pāli, Vin, 1 287[22–26]	Horner p.1833–34	Rhys Davids & Oldenberg, ibid p. 208–09	von Hinüber ibid: 7
kusim	cross-seam	cross seams	seam
aḍḍha-kusim	short cross-seam	intermediate cross seams	half a seam
maṇḍalam	circular seam	greater circles	(large square) piece
aḍḍha-maṇḍalam	short circular seam	lesser circles	half a (square) piece
vivaṭṭam	central piece	turning in	a piece in the middle
anuvivaṭṭam	side pieces	the lining of the turning in	a piece at the sides
gīveyyakam	neck piece	collar piece	a neck-piece
jaṅgheyyakam	knee piece	knee piece	a calves-piece
bāha-antam	elbow piece	elbow piece	an outer end

The commentary, which is attributed to the *Mahāaṭṭhakathā* (parts of which may go back to Mahinda's visit to Sri Lanka in the third century BCE),[13] is as follows (Sp 5, 1127[13–28])

kusī ti āyāmato ca vitthārato ca anuvātādīnaṃ dīgha-paṭāna metaṃ adhivacanaṃ	That (*kusi*) is a name for long pieces of cloth which are sown in length and width (along a seam)
aḍḍha-kusīti antara-antarā rassa-paṭānaṃ nāmaṃ.	The word *aḍḍha-kusi* is a name for short pieces of cloth at intervals.
maṇḍalan ti pañca-khaṇḍika-cīvarassa ekekasmiṃ khaṇḍe mahā-maṇḍalaṃ.	A *maṇḍala* is a big circle/square in each piece of the five-part robe.
aḍḍha-maṇḍalan ti khuddaka-maṇḍalaṃ	An *aḍḍha-maṇḍala* is a small circle/square.
vivaṭṭan ti maṇḍalañca aḍḍha-maṇḍalañca ekato katvā sibbitaṃ majjhima-khaṇḍaṃ.	The word *vivaṭṭa* means, having placed the *maṇḍala* and the small *maṇḍala* together, the middle section is sewn.
anuvivaṭṭan ti tassa ubhosu passesu dve khaṇḍāni atha vā anuvivaṭṭan ti vivaṭṭassa eka-passato dvinnaṃ eka-passato dvinnaṃ pi catunnaṃ pi khaṇḍānam etaṃ nāmaṃ	*anuvivaṭṭa* means two pieces on both sides of it (the middle piece), or *anuvivaṭṭa* is the name for the four parts, two on one side of the *vivaṭṭa* (middle section) and two on the other side.

---

[13] The Pāli Dictionary from the Vipassana Research Institute (dictionary.sutta.org) says that the *Mahā-aṭṭhakathā* "was the oldest and most important of the commentaries upon the Tipiṭaka: the tradition is that it was rehearsed at the first Council, and brought to Ceylon by Mahinda who translated it into Singhalese." However, Endo (2013: 19), quoting Mori, says that the *Mahā-aṭṭhakathā* (plural form) "contains many records of both Sri Lankan and Indian origins" from perhaps the 1st century BCE to the end of the Sīhala-aṭṭhakathā period; it is the *aṭṭhakathā* (sg.) used in the singular in the commentaries, that refers to the Mahindan material brought in the third century BCE. In this particular case we are dealing with the singular form (Vin 5, 1127[29], *Mahāaṭṭhakathāyaṃ vutto*, "as said in the Great Commentary").

gīveyyakan ti gīvā-veṭhana-ṭṭhāne daḷhī-karaṇa-ttthaṃ aññaṃ sutta-saṃsibbitaṃ āgantu-kapaṭaṃ

gīveyyaka means another piece of incidental cloth which is sewn with thread for the purpose of strengthening the place where it covers the neck.

jaṅgheyyakan ti jaṅgha-pāpuṇana-ṭṭhāne tatheva saṃsibbitaṃ paṭaṃ. gīva-ṭṭhāne ca jaṅgha-ṭṭhāne ca paṭānamevetaṃ nāmantipi vadanti

jaṅgheyyaka is a piece of cloth which is sewn in just this way that it reaches to the calf of the leg. They say, "This is the name for the pieces of cloth in the neck-place and calf-place."

bāha-antanti anuvivaṭṭānaṃ bahi ekekaṃ khaṇḍaṃ. iti pañca-khaṇḍika-cīvaren'etaṃ vicāritan ti. bāhantan ti su-ppamāṇa-cīvaraṃ pārupantena saṃharitvā bāhāya upari ṭhapitā ubho antā bahi-mukhā tiṭṭhanti tesaṃ nāmaṃ.

bāhanta means: outside of each side piece (*anuvivaṭṭa*) is (another) piece. It is considered thus due to the five-piece robe. (Or) *bāha-anta* means, putting on a robe of the correct size, gathering it up with the arms, having raised (the ends) up, both ends remain outside the face—it is the name for each of these (ends).

Kusi

Appendix One (page 140) has a diagram of the different parts of the robe which will help clarify their organization and function.[14] PED defines *kusi* as "one of the four cross seams of the robe of a *bhikkhu*" and *aḍḍha-kusi* as an "intermediate cross seam" which is how the translators have taken it above. Cone defines it as a "long piece of cloth (part of a *bhikkhu*'s robe)" which is based on the commentary in Sp above. It is a word with no OI counterpart; it appears to derive from the Dravidian root *kuṭ-* or *kuṭṭ-*, "to sew" (Kolami), cp Telugu *kuṭṭu*, "to prick, pierce, bore, sting, stitch, sew"; and the root is widespread in all the Dravidian languages, incl. Parji *kuṭṭ*, "sew"; and Gadba *kuṭ/kuṭṭ*; Gondi *kuṭ*, all "to sew." Kui has the verb *kuṭa*, *kuṭi*, "to incite, instigate, urge, rouse, stir to action" which is a metaphorical

---

[14] I am very grateful to Ven. Ariyadhammika of the Sāsanārakkha Buddhist Sanctuary in Taiping, Malaysia for sending me drawings showing how the *cīvara* is made.

extension of its basic meaning of "prick." There are also a parallel series of words with dental -*t*- or geminate -*tt*- which may be the source of *kusi*: Tamil *kuttu* or *kutti*, "to puncture, pierce, bore, perforate, stab, sew," etc; Kota, Toda, *kut*-, "to sew, pierce, prick" (DED 1719). Since proto-Dravidian weakened their cacuminal (retroflex) consonants to an -*ṛ*- or -*ḍ*- intervocalically and had no such thing as an -*s*- sound, and -*t*- weakened to a fricative (-*ð*-) intervocalically, it is quite likely that *kuṭa/i* or *kut*- were heard as *kusi*, or something close allophonically (Zvelebil 1990: §1.7.2).

Aḍḍha-kusi

For *aḍḍha-kusi*, CPD gives the definition "an intermediate cross-seam (on the robe of a *bhikkhu*)," quoting the commentary above (with the variant reading *pattānaṃ* instead of *paṭānaṃ*). But the commentary does not say that. The word *aḍḍha* is usually derived from OI *ardha*, "side, part, half," Pāli *aḍḍha* or *addha* ("half"), so it may simply mean "short" here, i.e. the long pieces of cloth cut in half. *Aḍḍha* could also be a Prakrit form for *aṣṭa* ("eight") where it appears in the Asokan seventh Pillar Edict in Toprā, (*aṣṭa* > *aṭṭha* > *aḍḍha* written *aḍha*) and it could also represent *aṭṭa*, as for example in the *Kevaḍḍhasutta*, with var. *Kevaṭṭasutta*. Here the most likely derivation is from the Dravidian *aṭṭam*, "cross direction" (Tamil), Malayalam *aṭṭam*, "what is across, transverse"; *aḍ*, "at right angles, crosswise" (Kolami); Telugu *aḍḍa*, "horizontal, crosswise," etc. It is a very common term shared by all the Dravidian dialects (DED 73). So, whereas a *kusi* is a long strip of cloth, an *aḍḍha-kusi* is a crosswise piece of cloth, and probably shorter, which is why the term *aḍḍha* in its meaning of "half" still fits the context (although not correct). Note this interpretation is confirmed in Appendix One, where the *aḍḍha-kusi* runs at right angles to the *kusi* strips. The same diagram shows that they are not half the length of a *kusi* strip, but perhaps one-third. The Dravidian derivation is the only one that fits the meaning.

Maṇḍala

This is probably an indigenous word. See footnote 209 for a discussion of the etymology of *maṇḍala*, which does generally seem to mean "circle" or "globe" in Pāli, OI and the putative Dravidian derivation (cp Tamil *muṭalai* < **moṇḍale*, Burrow 1948: 389). However, examining drawings of a *cīvara's* design (Appendix One), shows no circles in the robe, so von Hinüber seems to be right in his translation "(large square) piece" and "half a (square) piece" for *maṇḍala* and *aḍḍha-maṇḍala*. He refers to a sculpture

of a robe which shows square seams from Bharhut (2009: 8), a photo of which may be viewed in Cunningham 1879, plate 48, section 9).[15] It is likely then that *maṇḍala* has another meaning ("square" or something similar) that has been lost. In favour of this interpretation is the fact that such a temporary pavilion, a *maṇḍala-māḷaṃ* (described as a *maṇḍapa*, "temporary shed or pavilion, awning, tent" in the commentary) was erected for the Buddha's funeral celebration by the Mallas (see page 117 below). This structure, although the PED calls it a "circular hall with a peaked roof"—no doubt because of the word *maṇḍala*—must have been square in design as the Dravidian-source word *māḷa* suggests (~ Tamil *māṭam*, Kannada *māḍa*, "house, terrace, upper storey, platform," see Appendix Four, number 10, page 168). Interestingly Santali also has a similar word *maṇḍwa/maṇḍua* which is a (square) temporary shed erected for a marriage. Another related word is the Dravidian *maṇṭai* (Tamil, "mendicant's bowl, earthen vessel, standard of measure"), which again suggests a round shape, but the "earthen vessel" might also be square and perhaps used as a standard of measure. The commentary describes an *aḍḍha-maṇḍala* as a "small *maṇḍala*" which doesn't tell us anything new and is consistent with both interpretations of *aḍḍha* (as "half" or a cross-piece) mentioned above. Examining the Appendix One drawing (and other, similar ones), the *aḍḍha-maṇḍala* does not appear to be half of a *maṇḍala*, so "cross-piece" is the more accurate translation.

Vivaṭṭa

The word *vivaṭṭa* seems to represent the central piece created by sewing the *maṇḍala* and small *maṇḍala* together (number 1 and 2 taken together in Appendix One). But P *vivaṭṭa* (< OI *vi* + *vṛt*, "to roll, turn round, revolve, expand, develop, make or produce by turning, cast off"; Pāli *vivaṭṭati*,

---

[15] Von Hinüber refers to a work by Coomaraswamy which was not immediately available to me. Cunningham's work is freely available for viewing at https://ia802708.us.archive.org/14/items/stpabharhutabud00offigoog/stpabharhutabud00offigoog.pdf (accessed June 2020). See also Lüders 1963: 169–70 for an explanation of the strange picture of a mendicant's robe (*kanthā*) being "milked" and inscription, *V[e]duko katha dohati Nadode pavate*, "Veḍuka milks the tattered garment on Mount Nadoda." The garment is a *kanthā*, not a *cīvara*, so may not have been Buddhist (the word *kanthā* is never used in the canon); nevertheless, the patchwork composition of it was presumably similar. For a current photo of a patchwork robe laid out to dry (showing squares and rectangles, but no circles, see https://www.buddhistdoor.net/features/the-theravada-monastic-robe-the-design-and-meaning (accessed, June 2020). Note that this photo also shows that an *aḍḍha-maṇḍala* is not half of a *maṇḍala* piece.

"move back, revolve, turn away, turn over") has no meaning remotely close. In Pāli the nominal form *vivaṭṭa* is usually used to express the beginning or evolution of a new world cycle (sometimes the end); and it also refers to *nibbāna* or liberation in the sense of the absence of *saṃsāra* (*vaṭṭa*, "the round" or "the turning"). In OI *vṛtta* can mean anything from "round, circular, existing, past, elapsed" as an adjective to "circle, transformation, event, matter, conduct, behaviour, subsistence" as a noun.

Rhys Davids & Oldenberg's "turning in" translation is an attempt to capture that sense of "revolving" in the verb, although the meaning is not clear. The others use the word *majjhima* in the commentary to translate it as the central or middle piece of the robe. Perhaps Rhys Davids & Oldenberg meant "that central piece of cloth from which (or to which) the other pieces un(in)folded?"

What we have here then is a technical meaning of the word, probably adapted from the terminology of the robe practice of the autochthons, as the other terms are. That word is unlikely to come from Munda, as they have no letter *v-*. In Dravidian there are several possibilities the best of which is *maṭṭa* (with the typical change from *m-* > *v-* mentioned above on page 23), which is a common word meaning "measure, evenness, flatness, rule, line, gauging rod, limit, extent, bound, degree, guess, conjecture; equality in height, size, measure; moderation; whole quantity leaving no overplus" (Tamil). Its use here would be "standard of measurement" (Tamil, *maṭṭu*), that is, the central piece of standard size to which the other four parts—the side pieces, neck, calf and arms—are attached to make the final robe. For this derivation to work, one would have to assume that the word was adapted to IA morphology by the addition of the prefix *vi-* (perhaps for distinction or intensification, or perhaps with no meaning at all except to distinguish it from *vaṭṭa*, "round" or *maṭṭa* "polished" which already had established meanings), and *anu-* "alongside" to represent the side pieces.[16]

---

[16] There are many words in Pāli where the *vi-* prefix seems to be added gratuitously without a change of meaning. In the Vinaya for example, *naṭṭha* ("perished, destroyed, lost") and *vinaṭṭha* ("destroyed, ruined, perished") usually follow each other in the description of robe loss, but they both mean virtually the same thing. Presumably the *vi-* prefix was at one time added for intensification, but that usage has been lost. Other examples are *namati* and *vinamati*, "he bows down"; *budhyate* and *vibudhyate*, "he awakens, he perceives" etc.

Gīveyyaka

The word for "neck" is OI *grīva*, P *gīva* (*gīveyyaka* = "necklace, neck-piece")
is IA in lineage with a RV and Avestan pedigree. The Dravidian-derived
word is *kaṇṭha*, "neck" (see page 67 below). This is the piece of the robe
which is "at the neck" of the monk (Appendix one).

Jaṅgheyyaka

The word for "calf, shank" is *jaṅghā* (OI, P, *jaṅgheyyaka* = "knee-piece").
Mayrhofer treats the word as IA derived, with *zanga* as its Avestan reflex
("ankle"). However Chatterji (1929: xxiii) derives the word from AA, citing
Bengali *jāṅ*, Mon *jöng*, Khmer *jöng*, *cöng*, Santali *jaṅga*, "foot." To these
examples may be added phonetically similar words widespread among the
Munda languages: Sora *jeʔeŋ*, *jeŋ*, "leg" (*jeŋ* is the combining form); Juang
ijiŋ, "leg"; Korku *nãga*, *nʌŋga*, "foot, leg"; Mundari *jang*, "leg, calf".[17] The
jaṅgheyyaka is the robe-piece which is "at the shin" of the monk, when the
robe is worn.

Bāha-anta

Besides the word *gīva* ("neck"), *bāha-anta* are the only other words that
appear to be straightforwardly IA in derivation *bāha* ("arm") and *anta*
("end"). The word also exists in Kharia (*baha*, "upper arm") and Santali

---

[17] M1 vol. 1: 412 suggests that "Die austrischen Wörter…stammen vielleicht aus
dem Indoarischen und sind nicht dessen Quelle" (The AA words perhaps originate
from the IA and are not the source of this [word]). Unless otherwise noted, Munda
definitions are taken from the following dictionaries: for Santali, Anon (*English-
Santal dictionary*) which seems to have the same information as Campbell 1899, and
Bodding 1929–36 (hereinafter Bodding). For Kharia, Peterson 2009 and for Mundari,
Bhaduri 1931/1994; there are several other dictionaries at the Mundari Dictionary
Project at http://www.southasiabibliography.de/index.html and the Munda Etymological
Dictionary which may be found at
http://www.sealang.net/munda/dictionary/ (MED). Also see the University of
Hawaii's web page at http://www.ling.hawaii.edu/austroasiatic/. Unfortunately the
state of Munda dictionaries nowhere near approaches the completeness of the DED,
so the absence of words for some languages should not be looked upon as definitive.
The most complete dictionaries we have are for Santali, Kharia and Mundari; the
other languages are lacking, especially the south Munda languages (Bondo, Gutob,
Gta', Gorum, Sora).

bāĭh, *bāhĩ̃*, idem), but these are likely borrowed, as nothing similar occurs in the other Munda languages.

These are two highly technical passages on how to make a robe—the Buddha's instructions on making a robe based on his observation of the Magadhan fields and their execution by Ānanda—which for the most part incorporate the technical terms of the autochthons without making any attempt to translate them into MI.

In his discussion of inter-dialect "translation" practice, Norman notes how certain items in one dialect might remain unchanged when converted into another dialect:

1. Words which had a specific sanctity attached to them because they were regarded as technical or semi-technical terms and were therefore, despite the air of strangeness which they must have presented, too important to change;

2. Words which had no equivalent in the receiving dialect and which therefore had to be retained;

3. Words retained by oversight (1997/2006: 83).

What we have here appears to best correspond to number one above, except that the conversion is not inter-dialectic, but a true translation from a local language to MI (with the exception of the technical terms discussed which correspond to number two above). Since almost all the technical terms are preserved *in situ*, and their etymology is clearly correlated with the agricultural practice of the local farmers, it appears that the Buddha was directly adopting local practice to the inchoate Sangha's clothing needs. This of course still leaves open the question of what language the Buddha's "original" words were in, whether in some form of MI which incorporated technical terms from a local language, or whether he spoke first in the local language which was then translated into MI, while retaining technical local terms. These possibilities are also not mutually exclusive. He could have spoken in both languages at different times and probably did (Norman 1980a: 75; 1991: 144), although the temporal sequence is not determinable.

Methodology

Since the hypothesis of indigenous cultural influence on early Buddhism is to a large extent based on linguistic diffusion and word borrowing, the

determination of whether a word is taken into MI from a local language like Dravidian or Munda, or whether it is native to MI/OI, is critical.

How do we know which words were borrowed?

Although there is some uncertainty in the determination of word borrowings' directionality, the procedure involved does follow basic linguistic inductive principles. Burrow outlines these principles in his important 1946 paper, "Loan-words in Sanskrit", to determine the original provenance of a word common to Sanskrit and Dravidian (12–18; summarized in Emeneau 1954: 288). These rules are based on Caldwell's original observations (1875: 453–54), for the most part still valid today:

> (i.) When the word is an isolated one in Sanskrit, without a root and without derivatives, but is surrounded in the Dravidian languages with collateral, related, or derivative words; (ii.) when Sanskrit possesses other words expressing the same idea, whilst the Dravidian tongues have the one in question alone; (iii.) when the word is not found in any of the Indo-European tongues allied to Sanskrit, but is found in every Dravidian dialect, however rude; (iv.) when the derivation which the Sanskrit lexicographers have attributed to the word is evidently a fanciful one, whilst Dravidian lexicographers deduce it from some native Dravidian verbal theme of the same or a similar signification, from which a variety of words are found to be derived; (v.) when the signification of the word in the Dravidian languages is evidently radical and physiological, whilst the Sanskrit signification is metaphorical, or only collateral; (vi.) when native Tamil and Telugu scholars, notwithstanding their high estimation of Sanskrit, as the language of the gods and the mother of all literature, classify the word in question as a purely Dravidian one; - when any of these reasons is found to exist, and more especially when several or all of them coincide, I conceive we may safely conclude the word in question to be Dravidian, not a Sanskrit derivative.

The first step is to determine if the word has a valid IE etymology. Mayrhofer's works (M1 and M2) contain a useful summary of OI etymologies. Although the works, especially M1, contain a bias towards IA and IE derivations, Mayrhofer does list dissenting opinions (like Burrow, Emeneau, Kuiper and others) and provides enough information for the reader to form an opinion about the word's heritage.

If an IA/IE etymology is doubtful, analyze the currency of the word in the Dravidian languages. "If a word occurs widely in Dravidian and is of the nature of a basic element in the vocabulary, a corresponding word in

Sanskrit, if without IE etymology may reasonably be considered a borrowing from Dravidian" (Burrow 1946: 13).

Secondly, a word is shown to be Dravidian if it is clearly to be derived from a Dravidian root. Here Burrow gives the example of OI *candana* ("sandal wood") where there is no IE etymology and a Dravidian root is evident in all the Dravidian dialects. The word is widespread not only as a noun but as a verb with a meaning of "to rub into a paste;" and of course, the tree itself is native to Malabar, the southern part of the Indian sub-continent and far away from IA immigration routes. Its Indian provenance is thus another criterion to consider (Berger 1955: 25).

Thirdly, consider the antiquity of the word in Dravidian. If it goes back to the earliest texts, it is unlikely to be borrowed from Sanskrit, as the number of Sanskrit loan-words in early Dravidian text is very small. If the word appears late, it is more likely to be borrowed. Similarly, examining the timelines of the word in Sanskrit can corroborate this process; if it appears fairly late, it is more likely to be borrowed. Rare Sanskrit words not attested in the literature and found only in late lexicons are most likely loan-words.

A corollary to this last point (and the first point regarding currency) is to examine the word and its cognates' appearance in the various Dravidian language groups, south Dravidian, central and north. Its appearance in all three groups, or even two will be an indication of its age. Per Southworth "any etymon found in two of these groups can be presumed to represent a word in proto-Dravidian"; however, there are caveats (p. 230ff), the most important of which is the problem of linguistic diffusion and the narrowing and broadening of meaning over time. In a recent article Southworth provides some suggestions as to how to establish the most likely proto-meaning, viz., those which are common to all branches of the linguistic sub-groups and the most general and inclusive (2009: 105–08).

One may also confirm the Dravidian words' age by checking for their appearance in the oldest Dravidian literary works which go back to the middle to late centuries BCE. This of course has only limited value, because so many of the Dravidian languages are non-literary and therefore have no record;[18] plus the earliest Dravidian work preserved, the *Tolkāppiyam* (Tolk)

---

[18] The work to consult is a concordance published by the Institut Français d'Indologie in Pondicherry entitled *Index des mots de la litterature tamoule ancienne*, in three volumes, 1968–1970.

is no earlier than 400–500 BCE, so they do not compete in age with the earliest OI works which date from 1500 BCE (parts of the RV).

Fourthly, consider the phonetics/phonology of the word in question. Dravidian has, for example, a short -ŏ- and short -ĕ- sound which are lacking in Sanskrit where it is always long (in Pāli too they are usually long but can be short under certain circumstances; see Warder 1967: §34–35). So, imported words from Dravidian often change the -ŏ- > -a- (e.g. Tamil *mullai*; Kannada *molle*; Telugu *molla* ~ OI *mallikā* "jasmine") and the -ĕ- > -a- (e.g. Tamil *erukku* ~ OI *arka* "*Calotropis gigantea*"). Dravidian also has a retroflex voiced fricative -ṛ- (also written -ḻ- and sometimes by Burrow as -ḻ-), which is lacking in OI and tends to be represented by -l- (e.g. Tamil *kāṛ*, "blackness" ~ OI *kāla*, "black"). Other examples which Burrow cites to demonstrate OI as the borrower include loss of consonants (Tamil *puṛuka* ~ OI *puṅkha*, "feathered part of an arrow"); assimilation (Tamil *kuḷampu*, "hoof of an animal" ~ OI *kulpha*, "ankle"); and the alternation of voiced and unvoiced consonants which is rare in OI but common in Dravidian (which does not have a phonemic distinction between voiced and unvoiced stops). For example OI *jaṭā*, "matted locks" *śaṭā*, idem and *chaṭā*, "mass, lump" are all derivable from Tamil *caṭai*, "matted locks, thick bunch" or Kannada *jaḍe*, *jeḍe*; Malayalam *jaṭa*, *caṭa*, idem (Burrow 1948: 135 notes that the alternation *of j-*, *ś-* and *ch-* in OI is "an unusual feature in Sanskrit, but very familiar in Dravidian, and points strongly to the conclusion that the words are borrowed from that source").

Witzel (1999a: §0.4) makes the same point in a more systematic manner with a linguistic formula which shows what can or cannot belong to the IA language. So a RV word like *kīnāśa* ("cultivator of the soil"), which violates this formula, must be borrowed.[19] Berger simply says to look for suffixes not found in IA languages and "Nebenformen" (variant forms) that cannot be derived from Sanskrit sound-laws (1955: 25).

Fifthly, compare the meanings. When a meaning has developed a metaphorical extension from a simpler definition, the latter is probably the source. So, for example, *maḷḷaṉ* in Tamil means "strong, powerful, warrior" etc., while in OI it means "wrestler"—the wider and simpler Tamil meaning suggests that the Dravidian word is primary. But Tamil *mallaṉ*, "wrestler" with the identical OI meaning is a loan-word into Tamil.

---

[19] The formula is prefixes +/- {(s) (C) (R) (e) (R) (C/s)} +/- suffixes, where s = the letter s, C = consonant, R = semi-vowel and nasals, e = standard IE vowels.

The corollary of this meaning correspondence rule is that if the meanings have to be stretched to correspond (as for example Mayrhofer's complaint re: the IA word *nandā*, "joy" and the Dravidian *nantu*, "flourish"; see below), then the derivation is flawed. The same would apply to the (invalid) derivation of OI *āmra*, MI *amba*, "mango" < OI *amla*, "sour, acid" (see page 36 below). A corollary of this corollary is Emeneau's point that "comparative simplicity and avoidance of the assumption of tortuous phonological and semantic developments should also be aimed at, following the general practice of all disciplines ('Occam's razor')" (1954: 288).

Another point to consider, not mentioned by Burrow (because it is so obvious) is the provenance of the word. Fauna and flora not known to the immigrating Indo-Aryans are mostly loan-words. The same is true for toponyms, the names of local places and rivers which for the most part were retained by the Indo-Aryans, although often adapted to IA phonology/phonetics.

Many words with cerebrals in Sanskrit are also imported from Dravidian, where the cerebrals are part of the proto phonology. They are borrowed from Dravidian and possibly Munda into IA, introduced by Dravidian speakers adapting IA to their native phonology (MISL speakers; Deshpande 1979: 297).

A lot of variant forms (see, for example, *kaṭhina* below, or discussion on *jaḷogi* in Levman 2019a: 91–93) is another sure sign of a borrowed word. Many indigenous languages did not recognize IA phonemic distinctions (like aspirated stops, or voiced vs. unvoiced stops, or sibilants, etc.), and vice versa. So Dravidian allophones could easily be notated as phonemes, and many variant forms would result.

The overarching point that Burrow makes is that IA words can not be etymologized only from within IA, without considering the effects to the neighbouring languages. Although Dravidian has been pushed to the south of India in modern times, and Munda to a small area in the center-east (and Tibetan only exists natively in the Ladakhi dialect, in a tiny portion of the north-east in Ladakh), it is important to remember that from Vedic times the Dravidian and Munda groups were the dominant ones demographically and linguistically. Their influence on IA was therefore unavoidable. Similarly the Dravidian and Munda language families also influenced each other and it can often be the case that some Dravidian, Munda and OI words are traceable to a fourth language, an early form of Austro-Asiatic which Witzel calls "ParaMunda," an unknown, western AA language centered in the

Panjab associated with the language of the Indus valley culture (1999a: 25–26 for discussion of the word *lāṅgala*, "plough").

An Excursus on *Amba* ("Mango")

The problems of establishing a clear-cut derivation are sometimes insurmountable. Take, for example, the case of the common word "mango" (OI *āmra*, "mango tree, mango fruit" or *mākanda*, "mango tree"; MI/P *amba*). The English word is derived from the Tamil *māṅkāy* via the Portuguese *manga* (Yule and Burnell 1903/2000: 553; hereinafter HJ). Historically, it is a very old word in both Dravidian and Munda, pre-dating its IA appearance probably by several hundred years. We know this because of its occurrence in both Central and South Dravidian (per Southworth's groupings p. 50):

Central Dravidian (PCD):[20]
Kolami *māmḍī*, "mango"; *māmṛi, manḍi*, "mango fruit"
Naikṛi *māmṛi*, "mango"
South Dravidian (PSD2):
Telugu *māvi, māmiḍi, māviḍi*, "mango"; *mā̃gāyu*, "mango fruit"
South Dravidian (PSD1):
Tamil *mā, mānti, māti*, "mango"; *māṅkāy*, "unripe mango fruit"
Malayalam *mā, māvu*, "mango tree"; *maṅṅa*, "mango fruit"
Toda *mofin bum*, "mango fruit"; *mofin men*, "mango tree"
Kannada *mā, mām, māvu*, "mango"
Koḍagu *maŋge*, "mango"

There are no correspondences with Northern Dravidian (Brahui, Kuṛux and Malto), but this is probably an artifact of preservation (Southworth p. 230, §8.21) and lack of data. In any case any word found in at least two of the three groups can be presumed to be derived from proto-Dravidian (ibid §8.2), which Southworth reconstructs as **mām*. Proto-south Dravidian I (which includes Old Tamil) diverged from PD around approximately the middle second millennium BCE (ibid p. 253); so theoretically we can place the word for mango in PD long before it appeared in OI (in the *Bṛhadāraṇyaka Upaniṣad*, see below). Confirming its antiquity, the word for mango (*mā*) also appears in the oldest Dravidian literary text, the *Tolkāppiyam* (Chevillard 2010, chart one; Tolk. 1.7.29, v. 231), which may date from as early as the mid-first millennium BCE. Witzel however has

---

[20] For a breakdown of the different Dravidian languages by group, see Southworth p. 50, Figure 2.4 and Southworth 2009: 102, Figure 1.

shown that the earliest borrowings into the RV were first taken from the Munda languages (1999a: §0.5). The word's distribution here (all with the meaning "mango" unless otherwise noted) also shows an extensive chronological and geographical range:

North Munda:[21]	Santali, *ul*
	Mundari, Ho, *uli*
	Korku, *ambe/uli*
	Korwa, *uka*
	Asuri, *ul*
Central Munda:	Kharia, *kayar* ("green mango")
	Juang, *ole/ale*, *kayer* ("green mango")
South Munda	Sora, *uṛa/-ul*
	Gorum, *uṛaʔ/uṛa*, *aj-er* ("green mango")
	Gta', *ull*, *heʔwir/hiaʔwir* ("green mango")
	Remo, *ayer/ager*
	Gutob, *ili*
Burushaki (isolate)	*aam*

Zide and Zide (1976: 1297) reconstruct a proto-Munda (PM) root **uXli/*uXla*, "mango" and **kaj'-er*, **kag'-er*, "green mango".[22] The relationship of the Dravidian and Munda root to each other, to IIr and IE, and/or to a possible fourth language group is not immediately inferable.

Mayrhofer derives OI *āmra* "mango" from OI *amla* ("sour, acid; wood sorrel tree") because of the taste of its seeds (vol.1: 77, "Wohl (wegen des bitter schmeckenden Samens) zu *amlaḥ* zu stellen" (probably to be placed with *amla*, "bitter" because of its bitter-tasting seed). Not very convincingly he connects it with Armenian *amok'* "sweet," Albanian *ëmblë*, idem, Latin *amārus*, "bitter," Old High German *ampfaro*, "sorrel" and Lettish *amuols*, "clover." Presumably the derivation is *amla > amra* (change of eastern *-l-* > western *-r-*) > *āmra* (*guṇa* vowel gradation, "derived from/relating to *amla*") form > *āmbra* (addition of glide per Pischel §295 prior to assimilation or epenthetic vowel addition) > *amba*, "mango," or *ambira/ambila*, "sour, acid" in P.

---

[21] For a breakdown on the different Munda languages by group see Anderson 2008: 3–4, Figures 1.1, 1.2 and 1.3.

[22] "X is used here to indicate a vowel feature which must be reconstructed for Proto-Munda" (Zide & Zide 1976: 1330, note 4). The example words are taken from Table 1 of this article on page 1299.

Now, this derivation is quite complicated, but still regular, except for the change of meaning. How do we account for that? A mango is well-known as a sweet fruit, so why would the name be derived from the supposedly sour taste of its seeds? Who eats the seeds? The PW dictionary gives two definitions for *āmra*: "der Mangobaum, *Mangifera indica* und die Frucht des Baumes, bekannt wegen ihres Wohlgeschmacks" (the mango tree, *Mangifera indica* and the fruit of the tree, known for its pleasant taste, p. 674).

There are several other problems with this derivation. The word *amla* in OI (from which *āmra* is supposedly derived) represents the wood sorrel tree (*Oxalis corniculata*), the word *amlaka* the plant *Artocarpus lakuca*, *amlikā*, the tamarind tree and the wood sorrel tree. The word *āmra* itself also represents the "hog-plum" or *Spondia mangifera* tree (*āmrāta* or *āmra-vāṭika* = *āmra* "plantation").

The word *amba* also survives in Pāli with the meaning "bitter" in the compound *amba-kañjika*, "sour congee" as does the word *ambila*, "sour, acid" and it is common in the *sutta*s and the commentary as one of the six *rasa*s "tastes." Yet, as we have seen, *ambil(r)a* and *amba* are two different forms of the same word, with apparently opposite meanings.

Another problem is the provenance of the mango tree. They thrive in sub-tropical and tropical frost-free environments, so would probably be unknown to the north-west, incoming Indo-Aryan immigrants, until they had progressed further to the south. That would explain why the earliest mention of *āmra* (in the meaning "mango") is in the *Bṛhadāraṇyaka* (4.3.36) which is dated to perhaps the seventh or sixth century BCE (Olivelle 1996: xxxvi); the IA migrations of course started centuries before that and the Vedas are dated today to 1400–1000 BCE (Jamison & Brereton 2014: 5). The word *āmra* is unknown in the Rig Veda. The word *amla*, from which it supposedly derives is also unknown in the Vedas or the Upanisads. Its earliest occurrence is in *Manusmṛti* (first century BCE to second century CE, Olivelle 2004: xxiii), much later than the Pāli scriptures.

Being a native tree to the sub-continent, a much more likely derivation of the name is from the local languages, which, as noted above, pre-date its IA appearance by centuries. Three chronological lineages are discernible: 1) OI (*āmra/mākanda*) and its putative connection to IE; the latter word *mākanda* is clearly derived from the Dravidian **mān-kā*, but it is very late (Kāvya, early medieval), 2) Dravidian, which Southworth suggests is rooted in the proto-Dravidian etymons **mām-*, **mān-kā*, **maṯ-kāy*, "mango" (p. 215); and

3) Munda with its reconstructed roots *uXli/*uXla, "mango" (cp Santali *ŭl*; Mundari *uli*, etc.) and *kaj'-er*, *kag'-er*, "green mango" (Zide & Zide 1976: 1297; e.g. Kharia *kegɔr*). Munda also has some words similar to #1 (e.g. Korku *ambe*, Juang *amra*, Bodo-Gadaba *amɔ*), but these are a minority and probably loans from IA. Then there is the possibility or perhaps likelihood of a fourth unknown group (the Indus Valley language or one of them?) from which all of these may derive.

Apparently the names in question were used for a variety of different plants. Perhaps because of the similarity of their fruit and their popularity in cuisine, the words for the *Tamarind* and *Spondia*s plants were transferred to the mango. The *Spondia*s were once even called the "forest mango," although its fruit is bitter, not sweet. In any case, without any further data, determining the ultimate etymon for the word *amba* appears impossible (Southworth 79), although it certainly does not appear to be IE.

Derivations of this sort are not always this indefinite. Appendix Four of Chapter Three lists several derivations which show that it is often quite possible to determine the lineage of a word, with some degree of accuracy and confidence. Sometimes the facts seem evenly balanced between a Dravidian, Munda or IE derivation; in other cases the preponderance of non-IA data is indeed very convincing. Osada notes, quoting Masica and Zide, that a lot of Kuiper's Munda derivations were not well-founded (2009: 139) and points to Mayrhofer's scepticism in M2, a re-working of his IA etymological dictionary, published approx. twenty years after his first (ibid, p. 140). He notes the diminishing number of Austric entries in M2 but fails to note that a lot of the entries in M1 were omitted in M2, presumably because there was nothing further for Mayrhofer to add. Kuiper is indeed in some places hard to follow and makes some very tenuous connections, but Osada leaves out Witzel's work which in many cases confirms Kuiper's and does not mention Stampe's Munda Etymological Dictionary (which was not available to Mayrhofer), which contains data validating a lot of Kuiper's intuitions, as hopefully the reader of this section will discover.

—

Burrow ends the article with the admonition that "Indian linguistics must be based on the combined study of the three major linguistic families together with their external connections. This is a vast field for investigation of which only the surface has so far been touched" (1946: 30). Though 75 years have passed since these words were written, we are still a long way away from a full understanding of the complex currents and cross currents that have

shaped the evolution and mixing of both the Indo-Aryan and the autochthonous languages. Borrowing into IA from Dravidian and Munda and other languages was an important force as early as Vedic times, and as the Indo-Aryans moved further east and south, coming into increasing contact with the indigenous linguistic groups, more and more of the latter's vocabulary—representing everything from names of the local biota and places, to agricultural and artisan terminology, to religious, cultural and technical terms—became adopted into IA use; but although the IA culture and with it its language and dialects became increasingly dominant socially and politically, the cultural heritage of the indigenous peoples was at least in part preserved in the borrowed language assimilated into IA. It is those words which give us an important insight into the culture as it existed at the time of the Buddha, a culture which showed the tension of two different worlds meeting and the complex assimilations which resulted and divisions which remained. By itself, linguistic borrowing does not necessarily equate to cultural adoption; but there are a number of convincing arguments to adduce that, in the case of the IA-indigenous interaction, pervasive linguistic and structural borrowing do indeed mirror a strong cultural influence.

At the time of the Buddha, the importation of various structural forms and the adoption (better perhaps, "carry-over") of various cultural and religious terms from Munda and Dravidian into IA languages imply a certain symbiosis of these two ethnolinguistic groups, which was probably due, not so much to intermarriage, as to economic interdependence which required the native-speaking majority to learn the language of the immigrant (but politically and economically dominant) minority, thus providing the vehicle for continuous interaction and change.[23] There was also some degree of bilingualism on the part of IA speakers, but since the indigenous population was much larger, the diffusion of structural features was largely unidirectional, that is, from the local languages to IA (Southworth 118–23). To be sure, in this model, the IA languages also influenced the indigenous languages, but largely through word borrowing; there does not appear to be any innovations from IA incorporated into the Dravidian or Munda linguistic structures. This section is primarily concerned with the implications of the carry-over of various cultural, religious and socio-

---

[23] Initially at least the non-Indo-Aryan inhabitants of the sub-continent formed the plurality of the population. This has been noted by Burrow (1955/1973: 386), "a considerable element of pre-Aryan speakers"), Sjoberg (1992: 61), Emeneau (1974: 93; also in 1980: 198), Krishnamurti (2003: 15, 36), and is implicit in Southworth's models of linguistic diffusion (2005: 118–122)

political terms into IA from the native languages. It suggests a heretofore unacknowledged imprint of the local culture on early Buddhism, the details of which will be the subject of the next three chapters. The following Chapter Three expands on an article written in 2013 where this notion of cultural and linguistic diffusion was first examined (CR); it is all new research. At the end of the chapter I explore and answer some of the objections to the linguistic-cultural thesis (page 133f). Chapter Four studies the structural influence of proto-Dravidian on Old and Middle Indic, through the examination of Old Tamil writings. Chapter Five is a reprint of an article written in 2011 on what the word *muṇḍa* means in Pāli; here I try to demonstrate that, as well as being a pejorative appellation for the Buddha and his followers, it also is a name for the Muṇḍas and Puṇḍras, an indigenous non-IA, Munda language-speaking tribe.[24]

---

[24] The word *muṇḍa* in lower case and italics means, *inter alia*, "bald"; in upper case and non-italics, with retroflex dentals it refers to the Muṇḍa tribe; in upper case without retroflex letters it refers to the Munda = the AA language family. See Chapter Five.