

Pāli, the Language

Pāli, the Language:

The Medium and Message

Buddhavācānussati

(Remembering the Buddha's voice)

By

Bryan G. Levman

Cambridge
Scholars
Publishing



Pāli, the Language: The Medium and Message

By Bryan G. Levman

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2020 by Bryan G. Levman

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-4195-9

ISBN (13): 978-1-5275-4195-5

This work is dedicated to the memory of my father, William Levman, and my teacher, the Buddha:

*lābhā vata me suladdhaṃ vata me yassa me satthā evaṃ
mahiddhiko evaṃ mahānubhāvo ti*

“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!”

(Ānanda, speaking to the Buddha, AN 1, 228²⁰⁻²¹)

TABLE OF CONTENTS

Acknowledgments	ix
Preface	x
Chapter One.....	1
Introduction	
Chapter Two	36
Doctrine, <i>Anattā</i>	
Chapter Three	72
The <i>Bāhiyasutta</i>	
Chapter Four	77
The Buddha	
Chapter Five	92
The Middle Way	
Chapter Six	100
The Mind	
Chapter Seven.....	117
<i>Nibbāna</i>	
Chapter Eight.....	139
Karma	
Chapter Nine.....	151
The Vernacular	
Chapter Ten	169
Prolegomenon to a Buddhist Language	
Chapter Eleven	194
Dependent Origination	

Chapter Twelve	210
Meditation	
Chapter Thirteen	263
Conclusion	
Works Cited.....	267
Abbreviations	274
Index of Subjects	276
Index of Teachings	282
Index of Pāli Words Discussed.....	285

ACKNOWLEDGMENTS

This book takes its genesis from my doctoral thesis at the University of Toronto (2014) on “Linguistic Ambiguities, the Transmissional Process, and the Earliest Recoverable Language of Buddhism.” A central theme of that dissertation was the importance of the oral medium and how it affected the transmission of the *Buddhadhamma*. As time went on I began to realize that the oral nature of Pāli was an equal partner to the teachings; that is, that the message of the Buddha and the sonic medium in which it was transmitted were inextricably linked and mutually complementary – the medium was also the message. This book is my attempt to explore that insight in much greater depth.

I would like to thank my colleague and PhD supervisor Christoph Emmrich for the many productive discussions we have had on my dissertation and this book. Ven. Bhikkhu Bodhi read an early draft of the monograph and helped me with several doctrinal issues and I owe him a debt of gratitude. I also had numerous doctrinal discussions with Ven. Bhikkhu Piyadhammo which were very valuable, many of which have found their way into the book. I am further indebted to Prof. Madhav Deshpande for taking the time to review the monograph and for his numerous constructive suggestions. Numerous discussions with my students in the Advanced Pāli class at the University of Toronto have also helped me develop some of the content of this book, and I thank them for their questions and their suggestions.

I also thank my wife, Rosemarie, for her encouragement, patience and her proofreading expertise.

Bryan G. Levman
Toronto
Fall, 2019
bryan.levman@utoronto.ca

PREFACE

This monograph is a study of the Pāli language medium, which preserves *viva voce* (“with the living voice”) the Buddha’s many insights about the nature of life, suffering and liberation. It asks and answers the question, “What is it about Pāli that makes it a particularly appropriate communication medium for the Buddha’s message?” It is an extended study on Buddhaghosa’s observation that Pāli (which he called the Māgadhābhāsā, or language of Magadha) “carries the meaning easily” (*atthaṃ āharitum sukham hoti*, lit: “it is easy to carry the meaning”) and that “as soon as the sound impinged on the ear the meaning was present in a hundred ways, in a thousand ways” (*sote pana saṅghaṭṭitamatteyeva nayasatena nayasahassena attho upaṭṭhāti*, Vibh-a 388⁹, 11-13).

Buddhaghosa called Pāli a “natural language” (*sabhāva-nirutti*) by which he meant an original, or innate language. He maintained that a person without parents growing up in a forest isolated from human contact would speak Pāli of his/her own accord, because of his/her own inherent nature (*attano dhammatāya*, Vibh-a 387³⁵).

In the phrase *atthaṃ āharitum sukham hoti*, it is difficult to find the right word for “*sukham*.” *Sukham* is the opposite of *dukkham*, usually translated as “suffering” (see page 240 below), another notoriously difficult term the Buddha used. “Easily” is one possibility; others, equally germane, are “happily,” “comfortably,” “pleasantly,” “agreeably,” and “swiftly.” The point is that there is something in the structure and sound of the language which conveys meaning, parallel to and apart from the actual symbolic content of the words themselves. An analysis of what that something is, is the subject of this book.

While it is obvious that the symbolic content of a language often reflects the belief structure of a society,¹ it is perhaps not so obvious how the sound and

¹ In Central American Spanish, for example, the omnipresent interjection *ojala* (“if God wills it”), represents not only the Central American countries’ historical subjugation by Catholic Christian Spain, but also that latter country’s subjugation by Islam in historical times. *ojala* is indeed an Islamic expression (from Arabic, *‘in šā’a llāh*), and the sentiment expressed is much more consonant with the core beliefs of

structure of a language carry its meaning, both independent of and complementary to the symbolic denotation of the words. This observation is often expressed in vague generalities, like, for example, German being a rough, harsh and guttural language, Spanish a soft, romantic and mellifluous one and French aloof, artificial and condescending. We shall see that Pāli has some very special sonic and structural qualities which communicate affectively, on a stratum parallel to the symbolic.

One of the principal teachings of the Buddha was *anattā* or “not-self,” the emptiness of self and anything belonging to the self; the Buddha taught selflessness as an ultimate value and pathway to liberation. With *anattā*, the Buddha redefines the Vedic word *ātman* or “soul,” an ultimate value often associated with Brahma. *Nibbāna*, the state of final liberation was the extinction of the ego, and all personal desires and passions (< *nir* + *vā*, “to blow out). Although the verb *nir-vā* appears in the *Rig Veda*, the nominal form meaning liberative extinction was probably first used by the Buddha (Collins 1998, 137-38).²

The teaching of the Buddha is about renunciation, the gradual stripping away of all possessions, obsessions and views by which we define a self; it is the accumulation of these very attributes which cause conflict and suffering.

To communicate this vision of emptiness, in Pāli the Buddha cultivated a very sophisticated language of negation, *ἀποφασίς* (*apophasis*, < Gk, denial, negation), *ἐποχή* (*epoche*, < Gk, “suspension of judgement”) and *ἀταραξία* (*ataraxia*, “imperturbability”). Pāli is a language of the ineffable, the

that religion (“submission (to the word) of God”) than it is with Catholic Christian values which teach, not acceptance of an unknown fate, but self-determined, free will progress to a known spiritual goal. Yet that value was not enjoined upon the conquered peoples of Central America (nor of course did the Muslim invasions encourage Spanish individualism when under the Islamic hegemony), and their meek acceptance of the domination of the Spanish conquering and ruling class is a political fact of life to this day.

² The earliest usage of *nirvāṇa* as a noun or adjective is found in Pāṇini 8.2.50 where he writes *nirvāṇo* 'gniḥ = *upaśāntaḥ*, *nirvāṇaḥ pradīpaḥ*, *nirvāṇo bhikṣuḥ* = *uparataḥ* (“an extinct fire = ceased, an extinct lamp, an extinct mendicant = restrained”). It is not known whether Pāṇini lived before, after or contemporaneous with the Buddha. The earliest use of the verbal root (*nir* + *vā*) from which *nirvāṇa* is derived is in the *Rig Veda* (10.16.13) where it means to extinguish in a causative sense (*nirvāpayā*). The *Rig Veda* predates the Buddha by at least half a millennium, but there is no record of its usage there in the sense of “liberation.”

inexpressible, a privative language expressing “absence of” and “no-self” in a myriad different ways. While other religions described the ultimate, Buddhism eschewed all attempts to define the indefinable, recognizing that the linguistic medium itself was flawed. Conventional and arbitrary, language causes us to view an essentially counterfeit world as real and authentic. Starting with the omnipresent word “I” (which encapsulates the value of our Western culture), in its very structure Pāli expresses the “un-ness” and impersonality of both symbols and their referents.

Another core teaching of the Buddha is relativity, the observation of the interdependence of all phenomena, and the subsequent lack of intrinsic nature in anything, especially the self.

In Pāli the Buddha also developed a language to reflect and express this interconnectivity. Relativity means that nothing exists independently of other causes and conditions; everything is contingent and determined. The Buddha communicates this relativity in his descriptions of internal (within the self) and external phenomena (outside of the self), cultivating the art of synonymy and illustrating the interdependence of phenomena and the subtle interplay between them through an interlocking chaining structure which reflects relativity in the internal grammatical phrasing.

As well as examining the special usage of individual Pāli words, we will also examine how the overall structure of the language assisted the Buddha to communicate his teachings. Pāli is an inflected language, meaning that grammatical relationships are expressed in the terminations of nouns and verbs and not, for the most part, through prepositions (as in English) and postpositions, as in other non-inflected languages. The result is a communication medium of extreme conciseness and directness which is the perfect medium for clearly expressing the Buddha's simple view of reality. There is little of what one might call orthodox religion here; the Buddha examined the psychology of perceptions, and how we become attached to various afflictive phenomena, and reported what he saw in clear and simple language. So clear and simple that it is impossible for many of us to understand the simplicity of his vision of liberation.

The orality of the Pāli medium is always an important consideration, its presence and aliveness. The Pāli tradition recognized that analytic discrimination in regard to language (*nirutti-patisambhidā*) had more to do

with its utterance (*abhiḷāpe*) than its meaning.³ The sound comes to life in all the *suttas*, as the voice of compassion, the voice of sympathy for his fellows who are caught up in delusory perception and thinking about an “I” which leads to obsessive clinging and accumulation. The Buddha is teaching real people, with real, pressing problems. This is not the language of abstract philosophy, but the language of the *vulgus*, the language of the common people, a vernacular language in a popular tongue. Pāli has given up the rigid conjunct dissonances (joining of two dissimilar consonants which are often difficult to pronounce) of formal Vedic, the sounds of ritual, sacrifice and caste, of priestly invocation and teaching, evolving a euphonic naturalness and musicality characteristic of everyday speech. It is the opposite of authoritarian Vedic, a natural, reflective, come-see-for-yourself (*ehi-passiko*) conversational medium which is perfect for communicating the simple truths the Buddha taught.

Musicality and euphony are also important aspects of Pāli which we shall examine repeatedly in this book. For as in music, the rise and fall and interplay of the sounds of Pāli, their instantaneity and interdependency mirror the ephemerality (*anicca-tā*), formlessness (*animitta-tā*) and essential emptiness (*suñña-tā*) of life as taught by the Buddha. Pāli of course presents an additional semantic dimension, which, combined with the melodiousness and structure of the language results in a rich polyphony of sound and meaning unsurpassed in the history of human thought.

In Pāli, the language of relativity and impersonality is augmented by the Buddha's distinctive use of the passive voice to express the essential randomness and aimlessness of existence, where the actor is propelled mainly by external, determinate forces and is itself non-existent as a separate, distinct entity.

In Pāli we find a unique snapshot of the culture and values of mid first millennium BCE India, the social stratification expressed through various forms of address; a view of cyclical death-and-rebirth, a “going-nowhere” antithetical to the “somewhere” directional odyssey of the Western mythos.

³ There are four *paṭisambhidās* usually translated as “analytic insight” or “discriminating knowledge” which an Arahant possesses: knowledge of meaning, of *dhamma* (law), of language and of kinds of knowledge or perspicuity. Discrimination of language has sound as its object, not concepts (Ñānamoli 1991: 127; Vibh-a 387⁸⁻⁹: *Evam ayaṃ niruttipaṭisambhidā saddārammaṇā nāma jātā, na paññatti-ārammaṇā*).

These are some of the topics we will be taking up. Scholars do not believe (for reasons which we will examine) that Pāli was the actual language the Buddha spoke. It is geographically mixed in dialect and has some artificial elements and Sanskritization in the grammatical structure which betray interference with the natural form. Nevertheless in overall structure, word content, and in its vernacular, conversational character, it must have been very close to the (or a) language that the Buddha used to communicate his message. Pāli is a Prakrit (< Skt *prākṛta*, “original, natural, vulgar, unrefined, vernacular”) and all Prakrits go back to popular dialects as their source, evolving from the living speech of the people (Pischel 1900 [1981], 9; hereinafter PG); and the Buddha was of the people and spoke to the people and for the people. One might not dare to say that without Pāli the Buddha could not have communicated his message. But I would argue that the character of the language, fostered, confirmed and mirrored his teachings. And in perfect resonance, the Buddha's teachings fostered, confirmed and mirrored the essence of the language. In Pāli we have a euphonious polyphony (“many voices”) combining together to create a synergy unique in the history of psychological and religious thought.

The Plan of this Work

The Buddha's teachings represent a single theme with variations, like Bach's pinnacle violin Chaconne. It starts with the simplest of observations about the psychology of perception and proceeds from there to construct a magnificent pathway to liberation. His theme is the self and selfishness and he deconstructs it with illimitable variations and skill. Pāli is his clarion call to wake up to the truth of selflessness and impersonality.

As an examination of some of these variations, this book could start almost anywhere and similarly end, without diverging from the Buddha's core teaching of the emptiness of the self (*anattā*). All the teachings are closely interrelated and dovetail smoothly into a unified whole. So to a certain extent the plan of this work is arbitrary.

For the ease of those being introduced to Buddhist philosophy for the first time, the work is organized by doctrinal topics. It is impossible to appreciate the brilliance of the Buddha's didactic and linguistic virtuosity without understanding the overarching brilliance of his insights. As the reader becomes familiar with these I gradually introduce more and more illustrations about how the Pāli medium is used to sustain and augment the emotional force of his teachings. The message is the medium, and the message cannot be divorced from the sound of the language, its sonic energy.

From ancient times to this day the Buddhist community has recognized the importance of the *sound* of the teachings as a complement to their meaning, and daily recitations of the Buddha's words for both edificative and apotropaic purposes have always been an established part of the practice; indeed, several of the *suttas* have their genesis as group chantings (*saṅgīti*) and prophylactic recitations (*paritta*), and the regular *pāṭimokkha* recitation of the *Vinaya* rules have always been a central pillar of the *saṅgha*.

The Buddha's process of discovery was essentially heuristic, leading him from his initial observation of suffering to its cause (craving and the self) and the investigation into those factors which brought these afflictions into existence. This book follows a similar process in presenting, expounding and proving its central thesis of the complementarity of the Buddha's teachings and the medium in which they are taught.

Throughout the text I have inserted various “dioramas” on the Pāli language which reveal the inner structure of the language and concrete examples of etymological derivation, illuminating some aspect central or oblique to the text. These can be read independently, or even omitted, if the reader finds them too technical.

Chapter One introduces the Pāli language in the context of its historical evolution. Although Pāli contains some artificial elements, it is based on an oral, demotic dialect of the people which informs much of its natural, simple, spontaneous character. Opposing the formal, ritualistic language of the Vedas and the Vedic priests, Pāli was the perfect iconoclastic language for presenting radical new ideas, like those the Buddha taught.

Chapter Two presents the Buddha's unique teaching of selflessness through passages from the *suttas*. Here I present the Buddha's teachings on how the ego develops and how to eliminate it and I introduce some of the rhetorical devices and figures of speech the Buddha employs to maximize the impact of his teachings.

Chapter Three discusses the *Bāhiyasutta* as an example of the effectiveness of the oral teaching medium, focusing on its conciseness and directness characteristic of the Pāli medium.

Chapter Four focuses on the humanness of the Buddha. He was a real person who used the common language of the people. Vernacular language and vernacular song forms informed all aspects of his teachings. I show how many of the *suttas* are affected by popular music, both in terms of the rhythm of the words and the overall repetitive organization of the teaching.

In some cases (see page 86) popular songs are even incorporated into the *suttas*. Pāli's use of colloquial, earthy language, its sonic immediacy and visceral emotivity speaks to the here and now, to the urgency of human suffering and to the practicality of a philosophy which promises to end it.

Chapter Five presents the Buddha's teaching of the Middle Way and how he uses simile and metaphor from the natural world to universalize his teaching and impersonalize the subject.

Chapter Six introduces dependent origination, the Buddha's insight into how phenomena arise contingently and thus have no permanence or individual essence. It was this insight that led him to the realization of *anattā* and personal liberation.

Chapter Seven is about *nibbāna*, the ultimate state where the ego is completely extinguished. Here I introduce the chain structure of Pāli which mirrors the Buddha's teaching of relativity in the linguistic medium; I also discuss his use of apophatic language to describe the ineffable.

Chapter Eight deals with the conundrum of karma and how one can escape the recoil residue of our past lives to attain liberation. For if, as the Buddha maintains, we have lived limitless lives, then we are all responsible for limitless past misdeeds in previous lives.

Chapter Nine summarizes the vernacular elements of Pāli, forms of address and greeting, slang, and colloquial idiosyncrasies which reflect the social structure of the Buddha's day and add verisimilitude and immediacy to the teachings.

Chapter Ten, "A Prolegomenon to a Buddhist Language," examines systematically many of the linguistic features that have been introduced earlier in the book and discusses new observations. It argues, *inter alia*, that in Pāli the Buddha has created a language of impersonality, where the role of the actor and agent is minimized or omitted and passive linguistic structures represent a more accurate picture of reality where "things happen" rather than being made to happen by a free-willing agent. I discuss again the Buddha's extensive use of metaphors from the inanimate world as a linguistic proxy to express his insight into emptiness.

Because of its doctrinal centrality, in Chapter Eleven I return to a discussion of dependent origination, and a more detailed explanation on the use of grammatical and linguistic chaining in Pāli to complement the Buddha's teaching of relativity on the sonic plane.

The Buddha divides the practical aspect of his teaching into three parts, *sīla*, or morality, prohibitions against misdeeds; *samādhi* or meditation; and *paññā* or the realization of the ephemerality and impersonality of all phenomena. In Chapter Twelve, I discuss the central meditation practice of Buddhism, concentration meditation (*samādhi*) and mindfulness with recollection meditation (*satipaṭṭhāna*), and the various linguistic devices the Buddha uses to maximize the impact of his teachings.

CHAPTER ONE

INTRODUCTION

Pāli has always been something of a mystery language. Often called the language that the Buddha spoke, it is in fact probably not so. Buddhaghosa, the famous fifth century CE commentator on the early Buddhist *suttas*, equated it with Māgadhī, the language of Magadha, (*Māgadhikovahāro*), the ancient kingdom ruled by Bimbisāra and Ajātasattu where the Buddha spent a good part of his teaching career.⁴ Pāli is similar to Māgadhī, in that both are vernacular dialects derived from ancient Vedic Sanskrit, but it is not the same. Even the very name “Pāli” is a misnomer. In Vedic it means “row” or “line,” “margin,” “edge,” and by metaphor, a “dam,” “dike” or “bridge”. So it really just refers to a line in the sacred Buddhist texts recounting the words of the Buddha, in contradistinction to the commentary on it.⁵ As far as we can determine, it was seventeenth century French scholars, writing about Thailand’s religion and customs, who first used the name Pāli as a designation for a language as early as 1680 (Pruitt 1987, 124).

⁴ Sp 1214¹⁸⁻¹⁹: *ettha sakā nirutti nāma sammāsambuddhena vutta-ppakāro Māgadhika-vohāro*. “Here (the phrase) ‘His own terms’ (*sakā nirutti*) means the Māgadhian mode of speech with those qualities, according to the Fully Enlightened One.” For interpretation of *nirutti* as “terms” see Levman (2008/2009) and footnote 16. Those monks and scholars who assert that the Buddha spoke Māgadhī (the language of Magadha) immediately run into the problem that Māgadhī is quite different than Pāli (see the *Pāli-Myanmar Dictionary*, PMA, vol. 1, page 9 for the principal differences). The Burmese circumvent this issue by distinguishing between the Magadha language, which is the language of the Buddha (*Māgadhabhāsā*) and the Māgadhī language which is the vernacular demotic of the kingdom of Magadha, appearing in treatises like the *Rūpasiddhi*, in an early drama of Assaghosa, in Kālidāsa’s *Sākuntala* and various grammars (p. 8). They acknowledge that the two are different dialects, but do not try to explain their phonological relation.

⁵ Von Hinüber 1977, 243; Norman 1983, 1.

The name Buddha

The word “Buddha” is a title, not a proper name. It is the past participle of the verbal root *budh*, *budhyate* in Skt (*bujjhati* in Pāli), “to wake up, be awake, know, understand.” By adding the suffix *-ta* to the root, the past participle is formed, i.e. *budh* + *ta*, but phonologically it is impossible to pronounce an aspirated stop (*-dh-* in the middle of a word (*budh-ta*), so the aspiration moves to the final syllable (where it can be pronounced) and the unvoiced stop (*-ta*) becomes voiced (*budh-ta* > *budh-da* > *bud-dha*), assimilated to the previous syllable, a process known as progressive assimilation in linguistics. Although the root *budh* has the meaning of both awaken and understanding, it is the latter meaning, “know” and “understand” which is favoured in the Pāli suttas, while many of the later Sanskrit *sūtras* emphasize the fact that the Buddha has awakened himself and others from the sleep of ignorance.

For example, one trope which occurs over fifty times in the Pāli *suttas* is the Buddha's description of his enlightenment as *pubbe ananussutesu dhammesu cakkhum udapādi, ñāṇaṃ udapādi, paññā udapādi, vijjā udapādi, āloko udapādi*, “Understanding arose, insight arose, wisdom arose, knowledge arose, light arose in regards to phenomena previously unheard of.” It is in the later Sanskrit texts (like the *Nirvāṇa* and *Buddhabhūmi sūtras*), that the metaphor of the Buddha awakening himself and causing other beings to awaken from the sleep of ignorance appears (Malalasekera, 1971, 358).

Buddha is cognate with English “bid” (< Old English *bedoan*, to proclaim), English “beadle,” (< OE *bydei*, “herald, messenger”) and English “ombudsman” (< Old Norse *bodh*, “command”).

The etymology of the word is itself unclear and the present form that has come down to us may be a composite from several different sources (Mayrhofer 1956-1980, vol. 2, p. 263). Almost all Vedic words derive from primitive verbal roots, and some think that *pāli* derives from the root *pāl*, which means “to protect” “preserve” “defend,” the row or dam presumably being a line of defence. Others see a connection to the root *val* (“turn around”) from which the Indo-Aryan word *āvali* is derived (“row, range, continuous line”). Or it may have come from the Skt *paribhāṣā*, “rule” “maxim”, with the common *-r-* > *-l-* change one sees in the Prakrits. Others consider it a non Indo-Aryan word and derive it from the Dravidian language Kannada *pāri* or Malayalam *pāli*, both meaning “row, line.” A

recent book by Gombrich (2018, 11) connects it to the Vedic root *paṭh*, meaning “recite”; he suggests the word Pāli originally meant “text for recitation,” but the derivation of *pāli* from *paṭh* is problematic phonologically.⁶ It is a mystery, and there are other possibilities, but the closest, most direct derivation seems to be from Dravidian which would give the word Pāli an autochthonous pedigree.

The Burmese, who possess a very rich scholarly tradition in Buddhist matters, have their own etymological explanation of the word Pāli, contained in the introduction to their monumental twenty-four volume *Pāli-Myanmar Dictionary (Pāli-Myanama-Abhidhān)*. Some Burmese scholars trace the etymology of Pāli to the language of Pāṭaliputta, the capital of Magadha (*Pāṭali* > *Pāli*) (page 3). Others derive it from the Vedic word *pañkti* (“row, series, collection”) in the sense of continuity with the past (Burmese: *acañ*, “tradition, sequence, order succession”), as, according to the Burmese, Pāli has been spoken lifetime after lifetime as the natural language of mankind (page 3, footnote 2); or from the word *pakati* in its meaning of “natural,” “fundamental form” or “original.” Still others derive it from *dhamma-pariyāya*, “disquisition of the dhamma” in its dialect form of *paliyāya* (p. 3). Derivation from the root *pā* (“to protect,” related to the root *pāl* above) is also a possibility because it guards the teaching of the Buddha.

Most languages are named after the place and inhabitants of the people living there. But there is no such place as “Pāli” and attempts to localize it to any one area of India (like Pāṭaliputta) have proved fruitless, because it has characteristics of all the different dialects prevalent in India at the time of the Buddha, east, west and north. But though I say it is “probably not the language that the Buddha spoke” it is in fact very close to it, for historically we can trace it back to the third century BCE, within a century or two of when the Buddha lived, and it is probably the oldest of the Buddhist languages preserved. It is also the only Indic language which has preserved virtually the entire record of the Buddha’s teachings. The only other language that might compete with Pāli in terms of age is Gāndhārī, in which the oldest Buddhist manuscripts are found. But only a small number of

⁶ In a new article, published in 2019, Gombrich derives the word Pāli from Skt *pāṭhya* (“to be recited”) > *pāṭhiya* (insertion of epenthetic-*i*-) > *pāliya* (change of -*th*- > -*dh*- > -*ḍ*- > -*l*-) > *pāli* (with loss of -*ya* > Ø). So the actual language was called *pāliyaṃ* (neuter sing.) meaning “to be recited” which was mistaken as the locative form of the fem. noun *pāli* (which is identical, i.e. *pāliyaṃ*, “in Pāli”) and *pāliyaṃ* was back-formed to Pāli by removing the locative ending -*ya(ṃ)*.

canonical texts have been preserved in that dialect. So if Pāli is not the oldest Buddhist language it is certainly the best preserved and is very close to a language the Buddha spoke. I say “a language” as there is no reason to suppose that the Buddha only spoke one language.

Hundreds of years after the Buddha’s *parinibbāna*, in the so-called “Mahāyāna” writings (lit: “great vehicle” or perhaps “great knowing” per Karashima 2001, pp. 215-17)⁷ and in the later Theravādin commentarial tradition, it was common practice to attribute polyvocality to the Buddha as one of his supernatural qualities; that is, when addressing an audience, he spoke several languages at once, each directed to the corresponding language group in the audience and at the individuals’ particular reality level, skilfully crafted to waken their inner wisdom eye (for references see Levman 2008-2009, 44-45). In the *Itivuttaka* commentary, for example, it is claimed that, although the Buddha has just one voice, he “simultaneously communicated to beings of diverse languages according to their own language.”⁸ The Burmese dictionary claims that the language of Magadha can convey 128 languages at once.⁹ But by “polyvocality” I mean something more mundane. There were several different languages and dialects prevalent in northeast India at the time of the Buddha which one might organize in three broad groups. The ancestor languages of the indigenous peoples, that is the Dravidian speaking, Munda speaking and Tibeto-Burman speaking autochthonous groups; the various Indo-Aryan languages, starting with Vedic, the sacred, ritual language of the

⁷ The word *mahā* (Sanskrit/Pāli “great”) is cognate with English “mega” < Greek *μεγας*, idem. *yāna* means “vehicle” but Karashima argues that *yāna* is a Prakrit form of *jñāna* which means “knowing.” For polyvocality references see Levman 2014, p. 134.

⁸ It-a, 21³⁰⁻³⁴: *Eko eva Bhagavato dhamma-desanā-ghoso ekasmiṃ khane pavattamāno nānā-bhāsānaṃ sattānaṃ attano bhāsānāṃ vasena apubbaṃ acarimaṃ gahaṇ’ūpago hutvā atthādhigamāya hoti*. “The Bhagavan has just one voice for teaching the dhamma, and happening in one instant, after simultaneously reaching to each of the beings of diverse languages according to their own language, there is an understanding of the meaning.” Pāli quotes are generally from the PTS edition of the texts with variant readings from the Burmese recension sometimes used, when the latter seem more apt.

⁹ PMA vol. 1, page 7, footnote 1. “The Buddha is able to deduce and convey by word of mouth, preach and communicate quickly in many languages for the sake of those who hear the *dhamma*, so they who don’t know Māgadhi will understand, since this is a language which is quick and successful at being able to convey by word of mouth 128 languages to ordinary human beings who only speak one language.” (translation from the Burmese).

Brahmanical religion and its derivative colloquial dialects; and a *koine* or simplified inter-dialect language which was used for governmental and trade purposes across north India to eliminate interdialect phonological differences which impeded understanding, and harmonize the different dialects to a common intelligible language for everyday use. There were also foreign languages whose influence depended on who was in power at a particular time. Greek was important in the fourth century BCE because of Alexander's incursions into the sub-continent and in the sixth to fourth centuries BCE Aramaic was the *lingua franca* of the Persian hegemony.

By far the most important of these language appears to be the Indo-Aryan group, but that is largely, if not completely a function of preservation (i.e. luck), as the Buddhist teachings have not survived in any of the indigenous languages. That the Buddha was multi-lingual and taught in one or more of the indigenous languages seems quite plausible, as the phonetic constraints of these languages have had a significant phonological influence on Pāli (Levman 2016) and many of the local place names, designations for local fauna and flora and specialized vocabulary for local customs have diffused and been adopted into it as well (Levman 2013; 148-149, 169-173). Scholars are now adopting a more balanced view of the Buddha's genetic and linguistic heritage which is believed to have been a mixture of Indo Aryan (IA) and local ancestry straddling both cultures. The old view of the Buddha as an apex representative of brahmanical culture is losing ground; this now appears to be an historical overprint, that is propaganda fabricated to gain acceptance from the dominant brahman culture, or manufactured by the colonized, imitating the colonizers, or both (ibid, 153).

The Buddha lived and taught in a small sub-Himalayan area of north-east India, wandering between the Kingdoms of Magadha, Kosala and Kāśi and the tribal federations of the Mallas, Sakyas and Licchavis immediately to the east. So various scholars have opined that he spoke Māgadhī, Kosalan, Ardhamāgadhī (half Māgadhī, a language spoken to the west of Magadha), and the languages of the tribes. Indeed, he probably spoke all of these to one degree or another, and adjusted his teachings as necessary, moving between the dialects as his audience required (Levman 2016, 2). And as the Buddha's teachings spread rapidly to the north, center and west of India, they had to be adapted into even more dialects which were not necessarily mutually intelligible. Just how different they were we do not exactly know, but within two centuries of the Buddha's passing, in the mid 3rd century BC the Asokan edicts present a picture of significant dialect variation between different parts of the country. Problems of inter-communication and mutual intelligibility were resolved with the creation of a *koine*, that is, a common

or inter language which contained elements of all the dialects but was free from the most obtrusive dialect characteristics. This was probably pre-Buddha in origin as it was necessary for business, government and perhaps

Pāṭaliputta

All names have meaning, but sometimes it's difficult to piece together their original meaning.

Pāṭaliputta is the name of the capital of Magadha near present-day Patna in northeast India, but at the time of the Buddha it was only a small village known as Pāṭaligāma ("village of the Pāṭali trees"). It was named after the Pāṭali or trumpet flower tree which is known for its large flowers shaped like trumpets and its large, distinctive seed pods. Other names it had were Puppha-pura ("town of the blossoms") and Kusuma-pura ("town of the flowers"). The word *Pāṭali* is itself non-Aryan, stemming from the Munda or Dravidian language groups (Mayrhofer 1956-80, vol. 2, p. 245-46), two of the indigenous tribal groups inhabiting India before the coming of the Indo Aryans.

The second part of the name *-putta* is also a mystery. It appears to be derived from Skt *-putra* meaning "son", but that makes no sense in the context. More likely the earlier name was Pāṭalipuṭa where *-puṭa* ("seed pod of the Pāṭali flower" PED), referred to the long, distinctive seed pod. The meaning of *puṭa* then widened to include any container, box, bag or sack.

In the *Mahāparinibbānasutta* Pāṭaliputta is described as being in a "noble position" and "on the trade routes" and is called *puṭabhedanaṃ*, which is a Sanskrit word for "town, city". But it is clearly a descriptive compound with the meaning "opening of the *puṭa*." The Pāli commentary relates the compound to the opening of a container: *puṭabhedanaṃ* = "a place for opening containers of merchandise, a place for delivering bundles of goods" (Sv 2, 541), in other words a town or city which is a trade centre. But it is much more likely that the original meaning of *bhedanaṃ* (which is from the root *bhid* meaning "to split") actually refers to the splitting open of the seed pods of the Pāṭali plant, each of which has a large number of seeds, which profusion or seeds then became metaphorically associated with a profusion of merchandise.

Pāṭaliputta (cont'd)

Confirming this is the derivation of *puṭa* itself which appears to be non-Aryan word from the Dravidian language (CDIAL, entry 8253, Tamil *putar*, *putral*, Kannda *podar*), with the meaning of “bush” or “thicket.” So the original meaning may well have been “thicket of Pāṭali trees.” The Dravidian word was similar to the Middle Indic word for seed-pod (*puṭa*) and from there the word became generalized to mean any box or container, and the splitting of the seed pod became associated with the opening and packaging of merchandise in an urban centre

How then did *puṭa* become *putra*? Mayrhofer (ibid, 246) and Böthlingk Roth (SW 4, 633), two well known Indic philologists, both did not believe that *putra* was a mis-Sanskritization of Skt *pura* (“town” or “city”) and Mayrhofer proposes a hypothetical word **pūrta*, “ford” as the source. Pischel says it is simply a wrong Sankritization, but that seems unlikely with such a well-known word as Skt *putra* (PG, §238, note 2).

The story is complicated because two of the city’s earlier names (Puppha-pura and Kusuma-pura) ended in this word, *pura*. So there is the further possibility that the town, known as Pāṭalipura (“town of the Pāṭali trees”) with *puṭabhedanam* as an attributive compound associated with the town (“opening of the seed pods” evolving to a later meaning to do with a commercial urban centre) conflated *-pura* and *-puṭa* to produce *-putra* which was then simplified to *-putta* with the elimination of the conjunct consonants *-tr-*, a characteristic of all Prakrit languages.

artistic usage. By comparing cognate words in parallel passages that have survived from the different dialects, we can uncover some of the lexemic content of this *koine* which has been variously called a *langue précanonique* (“pre-canonical language”; Lévi 1912, 511); *Verkehrssprache* (inter language), *Hoch- und Gebildetensprache* (“high and cultured language”), that is a *lingua franca* (Geiger 1916 [2004], 3-5); *koine gangétique* (“Gangetic koine”; Smith 1952, 178); a *Kanzleisprache* (“chancery language”; Lüders, 1954, 8); and a *Dichtersprache* (“poetic language”; Bechert 1980, 34). Von Hinüber called it “Buddhist Middle Indic” (1983, 192-94), a “highly artificial,” “composite” language, “most likely close to the language of the Buddha himself,” which developed into Pāli. It was based on a vernacular dialect, traces of which are still preserved, but later standardized and “raised to a literary level” (1993, 107). Cousins called the

earliest language of Buddhism a “type of *koine* with vocabulary and syntax deriving from various dialects and without a standardized spelling” based on the language of Magadha, the administrative language of the Mauryan empire. He calls it “old Pali” which did not become “standard Pali” until the third or fourth century C. E. (2013, 128).

I have called it a *koine*, which is one of the media the Buddha and/or his immediate disciples used to communicate, especially when they had a varied linguistic audience, and it was from this *koine* that Pāli developed.

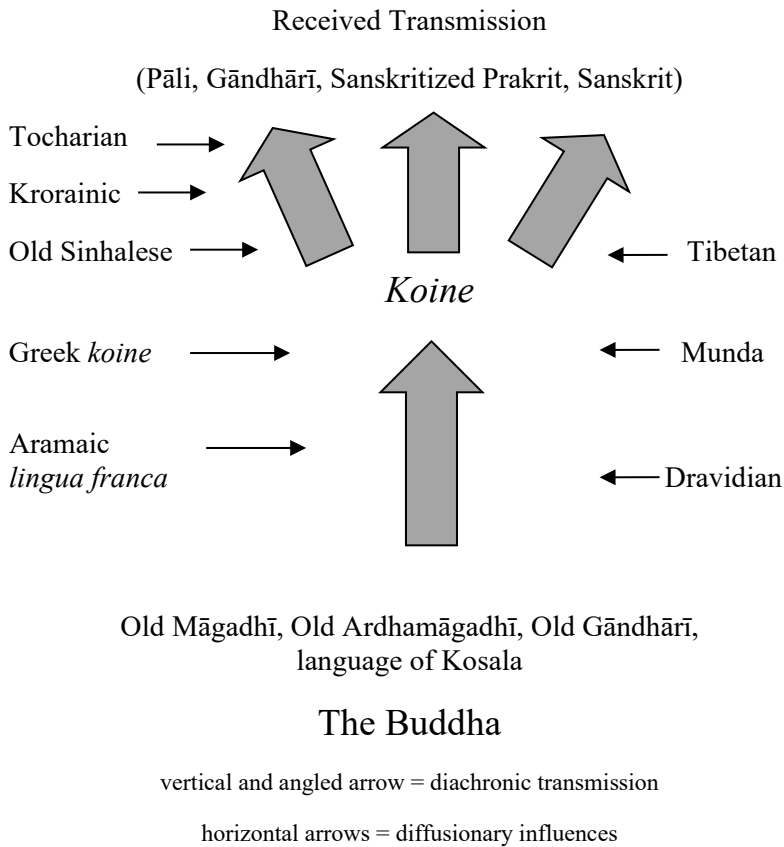
The Buddha no doubt knew many languages and used them for the appropriate audience. For example, there is evidence that the Buddha’s Sakya clan may have been Munda and/or Dravidian speaking (see Levman 2013), and undoubtedly the Buddha spoke in the autochthonous languages when that was all his audience understood. There were also many Indo Aryan dialects in north-eastern India at the time of the Buddha (Māgadhī, Ardhamāgadhī, to name the two most well-known) which the Buddha probably knew as either a native speaker, or as a second/third language learner. By “the language the Buddha spoke” I am specifically referring to the idiom, dialect or *koine* which evolved into Pāli, the only complete record we have of the Buddha’s teachings in an Indic dialect. It would be hard to believe that during his almost fifty years of teaching, he did not also teach in other languages, but Pāli is the only one that has survived. To be wholly accurate I should call this putative *koine* which underlies Pāli “the earliest recoverable language of Buddhism” for we cannot go back any further than this, given the data currently available to us, nor do we have any hard evidence that the Buddha spoke in this idiom; but it is a logical and parsimonious inference based on the evidence. The reader should keep all these factors in mind when I refer to the “language the Buddha spoke”. We will probably never know *exactly* the language the Buddha taught in; but we can reconstruct some of its lexemic content, and we do know that it was very close to Pāli which developed directly from it (Levman 2019).

The diagram below (Figure One) is a graphic reconstruction of the diachronic transmission of the *Buddhadhamma* as hypothesized here. It is, like any diagram, an oversimplification (and certainly not to scale!), as the *koine* is almost coeval with the dialects which bring about its genesis, arising along with the need for communication and understanding, and the varying degrees of diffusionary influence of the native languages (Tibetan, Munda and Dravidian), the IA languages (Tocharian, Krorainic and Old Sinhalese) and the foreign ones can not properly be shown. Old Sinhalese (OS), for example, although geographically distant from the other languages had an

important influence on the language of early Buddhism because of Asoka's son Mahinda's early translation of the *Tipiṭaka* ("the three baskets of the Buddha's teachings") commentaries into this language in the third century BCE (*Cūlavamsa* 37, 228; Ps 1, 1²¹⁻²⁴; presumably the *suttas* were also translated into OS as well), and their re-translation into Pāli by Buddhaghosa in the fifth century CE. To complete the diagram, one might also put Vedic underneath (or alongside) the Prakrits, with the understanding that Vedic itself contains many Middle Indic dialect features (Levman 2014, 55); and Sanskrit might be shown obliquely off to the side to denote Pāṇini's standardization of the language in the middle first millennium BCE. That Pāli is closer to Vedic than it is to Pāṇinian Sanskrit¹⁰ was argued in 1875 by Robert Childers who wrote one of the earliest Pāli English dictionaries and noted the many forms found in Pāli which came from Vedic but had not survived in Pāṇinian Sanskrit (xii, n. 2)¹¹. Pāli also contains archaisms which predate even Vedic, suggesting that the original dialect on which Pāli is based is very old (Oberlies 2001, 6-7; examples he gives include the words *idha*, "here" *kiṇāti*, "he buys", and *(a)sāta*, "(mis)fortune, (un)pleasant"). In Oberlies' new *Pāli Grammar*, he asserts that "Pali goes back to a Vedic vernacular situated most probably (south-) east of Arachosia near the Bolan Pass (2019, 35), which he calls a "Nebenmundarten" ("nearby dialect") of the Ṛgvedic main dialect (p. 21).

¹⁰ Vedic and Pāṇinian Sanskrit are both known as Old Indic (OI) and the Pāli and the Prakrits are known as Middle Indic (MI). Vedic and Pāṇinian Sanskrit are basically the same language; the latter is simply the standardization and formalization of the former, which took place in the 5th or 6th century CE. The major difference between the two is that Vedic contains some archaic forms which are not found in Sanskrit (but are sometimes found in Pāli and the Prakrits). Vedic dates from the mid to late 2nd millennium BCE, and because it was a liturgical, sacred language, did not change very much at all – that is, it was deliberately preserved in its ancient form by the brahman priests – between then and its standardization by Pāṇini. I use the word "Vedic" or "Vedic Sanskrit" when the word in question can be definitively traced to the Vedic writings. The word Sanskrit (Skt) is the more general term which in my usage usually refers to post-Vedic works (from Pāṇini's time forward), but in some instances (which will be clear from the context) refers to the language as a whole which includes Vedic and its Sanskrit progeny. The word Sanskrit itself simply means "well-formed, perfected, polished" (< Vedic, *saṃskṛta*)

¹¹ For other examples of Prakrit forms in Vedic, see Wackernagel 1896 [2005]: §37b; PG §6; Bloomfield & Edgerton 1932. For interchange of surd and sonant mutes (voiceless and voiced stops), see Chapter 2, p. 26f; for loss/interchange of aspirated stops and non-aspirates, see Chapter 3, p. 47f. For interchange of *-m-* > < *-v-* cf. Vedic *-vant* and *-mant*, both with same meaning ("consisting of"). For assimilation of consonants to one double consonant see §406-§421.

Figure One: The Transmission of the *Buddhadhamma*

Because of phonetic simplifications to eliminate dialect differences, the *koine* had a lot of words which could have several different meanings. Voiced and unvoiced intervocalic stops, for example, were often replaced with a glide -y-, to avoid phonemic confusion on meaning, as some dialects preferred voicing stops and others not. It was up to the hearer then to back-translate the glide into the relevant stop in her/his dialect; and since there were often many possibilities this was usually determined by context and individual choice, which automatically led to variation. The same practice was also followed with aspirated stops which were replaced with an aspirate only (-h-). This kind of variation is the principal reason why scholars maintain that the Buddha did not speak Pāli per se, but an earlier dialect or

koine on which Pāli was based. When this *koine* was later formalized as Pāli, there were naturally many differences, depending upon each person's interpretation of glides or aspirates and other simplifications in the *koine* (for which, see Levman 2016, pp 12-14). There was also significant variation between traditions (Pāli, Gāndhārī and the various Buddhist Sanskrit forms), though all derived from a common source (see inset on **veha*, page 82 below).

Isipatana

an example

Take, for example, the compound describing the famous park where the Buddha preached his first sermon, which Lévi notes as an example of intervocalic consonant weakening (1912: 499–500); the name has been preserved in two forms, as *isi-patana* (“descent of the seers”) and *isi-vadana* (“speaking of the seers”); however, as Caillat has shown (1968: 177–183), the underlying form which gave rise to both these possibilities was *isi-vayana* which itself was derived from Skt *ṛṣya-vṛjana* (“antelope enclosure or pasture”), where *ṛṣya* > *isi* and *vṛjana* > *vajana* > *vayana*. The translators had misconstrued *isi* as derived from Skt *ṛṣi* (“seer”) and *vayana* as derived from Skt *patana* (“descending,” the weakening of *-p-* > *-v-* being a common Old Indic > Middle Indic change), or *vadana* (“talking”). Interestingly, the commentary retains the correct etymology of the compound (*migānaṃ abhaya-dāna-vasena*, “on account of giving a fearless retreat to animals”) while inventing fake etymologies for *isi-patana* and *isi-vadana* (Levman 2014: 394–396)

In the example above, the *koine* form transmitted was *isi-vayana*, from the Vedic *ṛṣya-vṛjana*. Without going into all the differences between Vedic and Prakrit, in the second word of the compound, the intervocalic *-j-* was changed to a consonantal glide *-y-* (and the vocalic *-r-* > *-a-*, a typical Prakrit change), with the meaning “enclosure.” The first word *ṛṣya* (“antelope”) which becomes *issa* in Prakrit (with *r-* > *i-* and *-ṣy-* > *-ss-*) somehow became *-isi-* (“seer”) which then prevented the translator from understanding the second word *-vayana* (< *-vajana*) as “enclosure (for antelopes)” predisposing him/her to translate it as either *patana* (“descent (of the seers)”) or *vadana* (“speaking (of the seers)”). The original meaning was however, preserved in the commentary.

If all this seems very complicated, don't worry, it is complex. The Buddha taught in this *koine*, and his disciples, on his injunction, memorized what he had to say. There was no writing in those days; everything was done orally. Writing may have been used for governmental and trade purposes in the Buddha's time, but we have no record of it until the 3rd century BCE with the Asokan edicts, and religious teachings were not written down at all. The Brahmans memorized the Veda and the Buddha's disciples memorized his teachings. It was not until the 1st century BCE that his teachings were committed to writing.

When the transmission wasn't understood, as in the example above, various attempts were made to clarify his meaning, by applying the rules in reverse.

So if a consonantal stop changed to a glide, then one changed the glide back to a stop, based on one's interpretation of the passage. But different interpretations were possible and often one mistake (*isi-*) lead to another (*-patana*). The example here is not uncommon, but most are not as egregious as this, and certainly most did not affect meaning significantly in a doctrinal sense. But some did, as we shall see.

Did the Buddha speak Pāli?

It is impossible to know exactly what language the Buddha spoke. In my 2016 study I suggested it was a *koine*, or inter-language which eliminated most of the common dialect differences, so that someone speaking a Middle Indic dialect in the east of India could understand someone speaking one in the west. This *koine* was then transformed into Pāli (and other dialects) some time after the Buddha's passing. So no, the Buddha didn't speak Pāli, but he spoke something close to it. An example of the kind of alteration of meaning that can enter the language is the inset box on page 11, where the word in the *koine* was *isi-vayana* which was misunderstood by the tradition as to its meaning, even if the correct phonemes were translated.