WHAT LANGUAGE DID THE BUDDHA SPEAK?

KEYNOTE LECTURE BY PROF. JAN NATTIER

The Proto-History of Buddhist Translation from Gāndhārī and Pāli to Han-Dynasty Chinese

Translation & Transmission Conference 2017, Boulder, Colorado, USA
A Tsadra Foundation Initiative
June 2, 2017

https://m.youtube.com/watch?v=aCCkA12ynLs

Part of the key-note lecture, by kind permission from Prof. Nattier and Tsadra Foundation – transcribed & edited by Borobudur Jinabhumi Team.

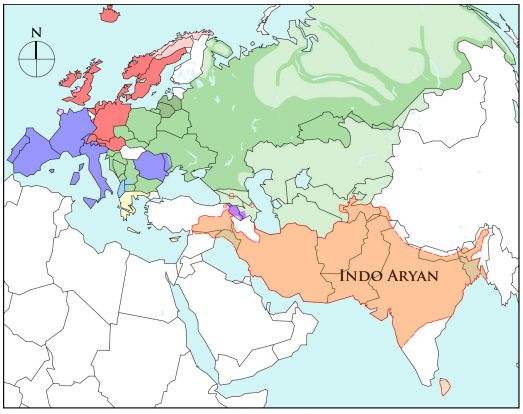
In the early centuries, Buddhist teachings were circulated orally only for several centuries. It was in China, in the northern capital of Luoyang, in the middle of the second century of the Common Era, that we see the first known examples of the translation of Buddhist scriptures into a non-Indian language. But as soon as we raise the question of what the translators were working from, we are immediately forced to deal with the fact that the translation of Buddhist scriptures did not actually begin in China. Instead, it had already begun several centuries earlier, in all probability, during the lifetime of the Buddha himself. This part of the story is rarely told and much of it has only come into clear scholarly focus within the past few decades.

So very basic question is: What language did the Buddha speak?

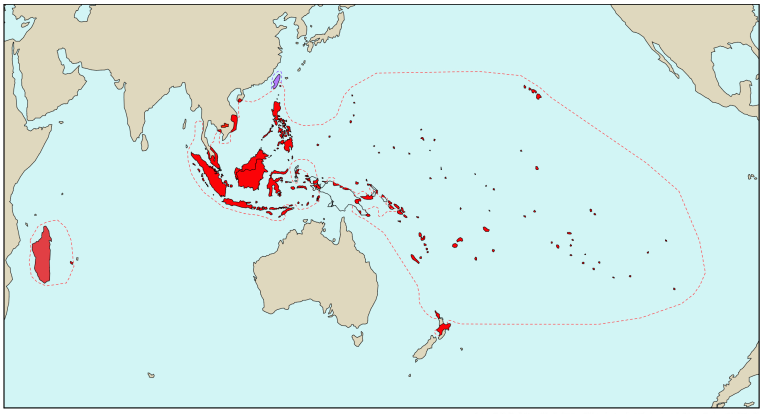
To begin with what we do know, it's well established that the Buddha lived in the northeastern part of the Indian subcontinent, having been born in what today is southern Nepal, and that he spent most of his teaching career in the region known as Magadha. So,

it does seem reasonable to suppose, as specialists do, that most of his teachings were initially delivered in some form of what more technically is called old Magadhi, and it's referred to as “old” to distinguish it from Ardha Magadhi, a much later language that was used for the transmission of scriptures belonging to the Jain tradition.

But old Magadhi was far from the only language spoken in India or even in northern India at the time and it seems likely that the Buddha may have been able to communicate in more than one language as K.R. Norman had suggested in “*The Dialects in Which the Buddha Preached*” (1980) and C. Chowdhury in “*Did the Buddha Speak Pali*?” (2009).



INDO-EUROPEAN LANGUAGES


AUSTRONESIAN LANGUAGES

The Indo-Aryan (Indo-Iranian) languages are divided roughly into three periods: **Old Indo-Aryan** is represented by the Vedic language of the Ṛgveda, and includes Classical Sanskrit; the **Middle Indo-Aryan** (Middle-Indic, Middle Iranian) languages began to appear in inscriptions around the time of the Buddha, 5th century BCE – these are often referred to as Prakrits and have antecedents in archaic OIA languages which existed in parallel to Vedic, but which are no longer extant; **New Indo-Aryan** languages are the modern North Indian languages such as Hindi, Gujarati and Bengali.

Prakrit is a term that's used to refer to vernacular, that is spoken Middle-Indic languages, in contrast to the language of Sanskrit. It is worthwhile to pay attention to the actual meaning of Prakrit and how it is related to Sanskrit.

Prakrita can be defined as natural or unrefined or unvarnished. “Prakrti” is a famous word in Indian philosophy and the contrast here is with Sanskrit which comes from “Sanskrta”, which in this context means “put together”, in the sense of almost “decorated”, “made beautiful”. Now, it sounds from that pair of terms as though what we're talking about is Prakrits being made into Sanskrit and that did happen later in Indian Buddhist history.

We have early teachings composed in Prakrits, the oral teachings were surely in Prakrit languages, local vernaculars, and then over the course of time, Sanskrit became more and more the language of choice for the transmission of Buddhist teachings and texts. So, for example by the time the Tibetans received this tradition, Sanskrit appeared to be the normal Buddhist language. But things were much more diverse in the early period.

The situation is a little weird. The Prakrits are the daughter languages of Sanskrit. Prakrit is to Sanskrit as the Roman languages for example are to Latin. So, Sanskrit is older as a language but newer as a language of transmission of Buddhist texts. It's as if the early Christian teachings were in French, but this is obviously anachronistic, but as if they were in French and Italian and Spanish and later were put into Latin. So, there's something interestingly backward about what happened language-wise in the history of Indian Buddhism.

So, the many different types of Prakrits, these offspring languages of Sanskrit, were the languages spoken in many different parts of India, especially northern India at this time.

When we look at the broad cultural history of Buddhist translation, there's a prior question: Should we translate at all? Or should we all just learn Sanskrit and recite and study the text in Sanskrit? Or should we all just learn Pali and recite and study the texts in Pali? Do we translate? Or do we keep the sacred texts in their original language?

It’s worth, at the beginning, to look at the translation policy that is attributed to the Buddha himself, in an account contained in the Pali Vinaya, as well as in the different Vinayas belonging to different Nikayas or “schools”.

In the Pali lineage, the lineage that uses the Pali language, that comes to be known as Theravada in the modern world, but also in the Vinayas belonging to a number of other Nikayas’ ordination lineages, we find an account of what the Buddha had to say about this, and because of its spreads and shared across a broad range of canons that probably points to its antiquity.

In this story, we actually get a very clear-cut account of what the Buddha's own translation policy was said to be, immensely valuable passage. So, in this story the Buddha is portrayed as opposing the formalized recitation of his teachings in anything other than the local vernacular. No formalized language, use the local speech.

According to the texts, there were two Brahmin brothers who had been converted to Buddhism and they were concerned - as good Brahmins might be – that monks “of a variety of origins and family backgrounds” were corrupting the word of the Buddha by reciting his teachings each in their own way. “Let us remedy the situation”, the brothers suggested to the Buddha, by putting the words of the Buddha into “chandas”, which in general terms it is associated with the style of recitation used for Vedic texts, that is the religious literature with which the brothers would have been most familiar.

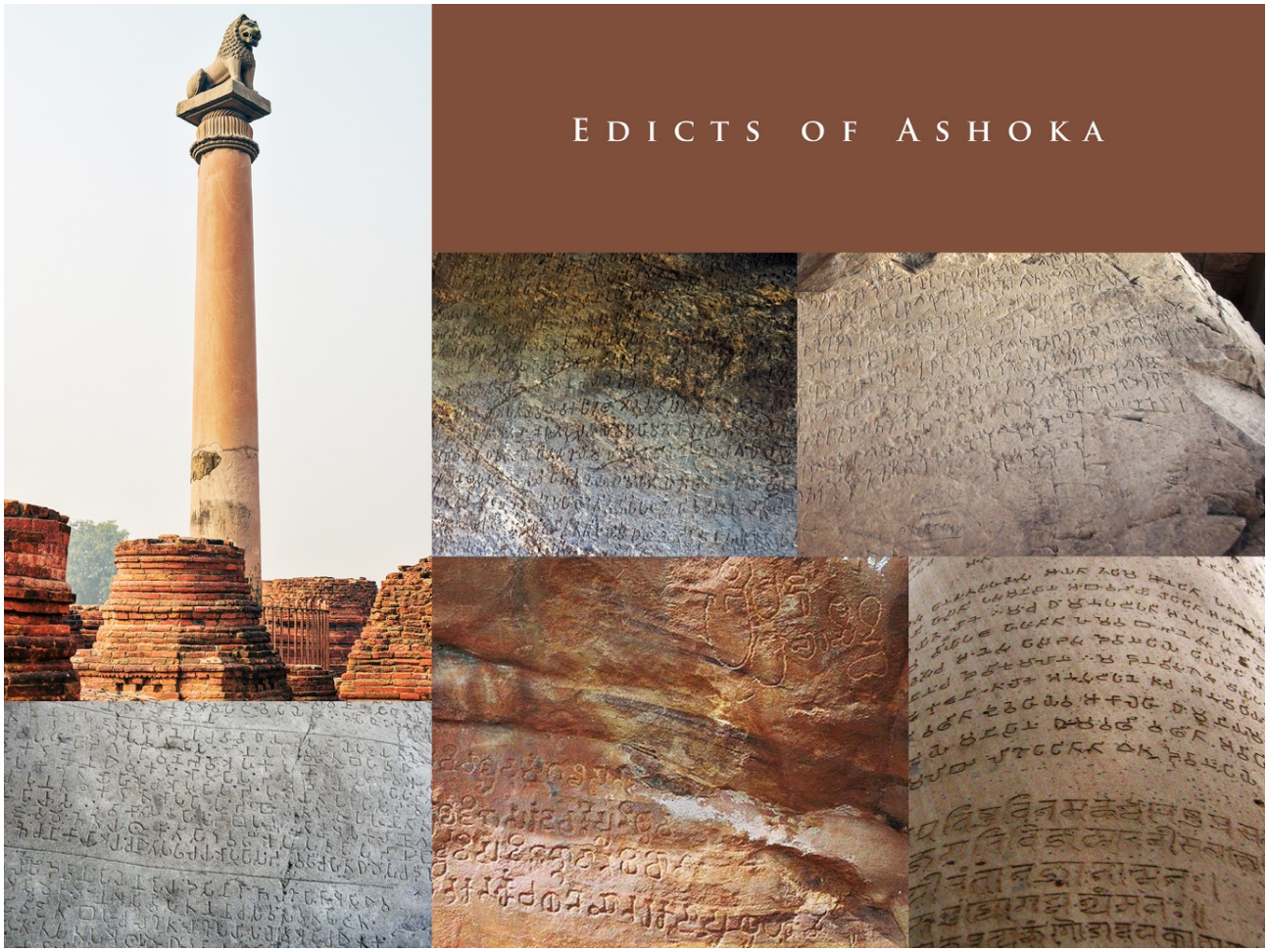
What's really remarkable about the recitation of Vedic texts for scholars of languages, for Indologists especially, is that these were passed down orally, not for centuries, but for millennia, in an ancient language that was no longer spoken even by these Brahmins brothers. In other words, they were reciting classical texts when they were taught Vedic scriptures, and yet that memorization tradition was such that every syllable could be preserved, which is why modern scholars can actually find out from these sources now which are now written, what Vedic Sanskrit was like and talk about the evolution of this language. The brothers seem to have been arguing for doing something more elegant, more standardized, more familiar with the Buddhist teachings, so there wouldn't be – to paraphrase their apparent opinions – there wouldn't be all these various people with their local accents and local dialects messing up this important stuff.

The Buddha's response to the brothers’ suggestion is not at all ambiguous. He forbids the recitation of his teachings in “chandas”, whether that means versified, fancy text, or whether it means actual Vedic language. Whatever it is, he forbids that, but then he says, instructing his followers, “Teach the word of the Buddha each in your own language”.

So, as far back as we can go in the sources that are available to us, the translation policy attributed to the Buddha himself, you can't get much more authoritative than that, the translation policy was “Teach in whatever the local people will understand”. Use the vernacular, or to put in Indian terms, use the Prakrit languages. And that is exactly what the Buddhists seem to have done in these early centuries.

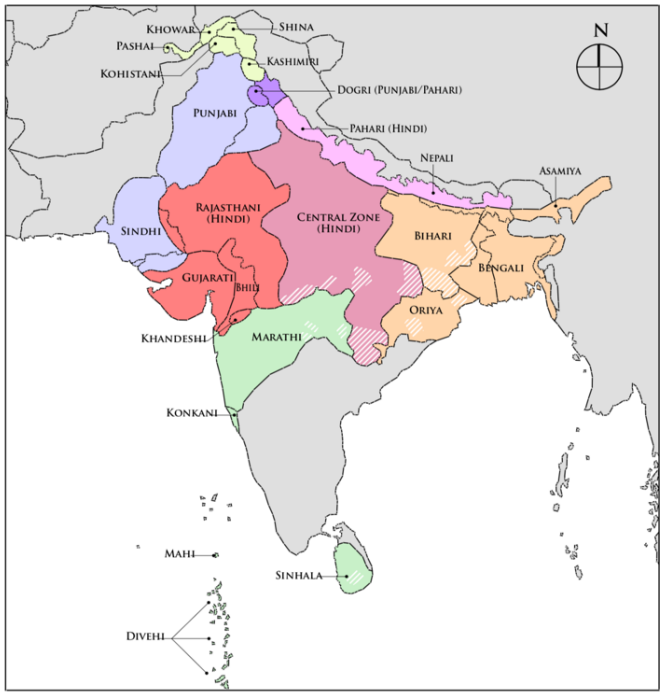
As already mentioned, there were there many different Prakrits being used in different parts of India, and we can actually obtain some concrete information about what they were like from the inscriptions of King Ashoka, who had messages to his subjects, carved on stone

and erected throughout his realm. Ashoka's Mauryan Empire, included virtually the whole of the Indian subcontinent, the largest unified Empire until the coming of the British in India.

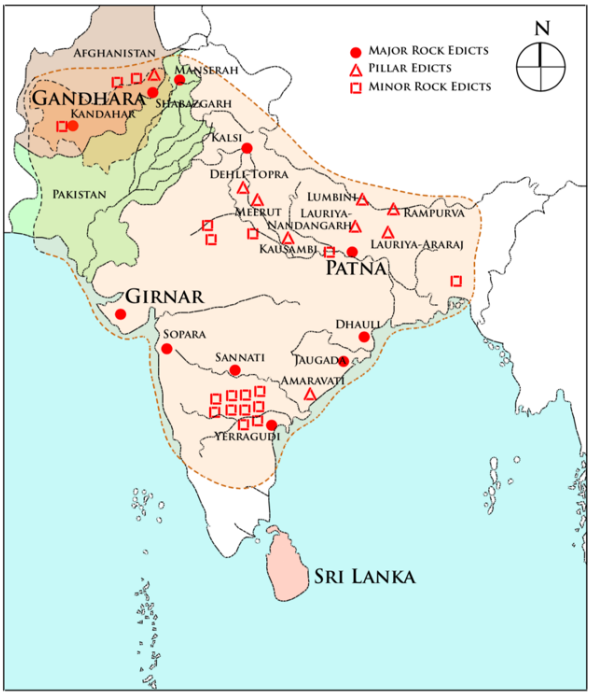


Ashoka controlled a huge amount of territory and his inscriptions have been found in a huge amount of that territory as well. Normally, when we talk about Ashoka, we dive right into the content, what a wonderful king he was and how he gave up violence when he became a Buddhist, etc. That's all immensely important, but what would be interesting to know is what languages were those inscriptions in and this is where scholars can use this material because they are in a variety of Prakrits and even in some non-Indian languages.

Ashoka's realm extended up into what today is Pakistan and even eastern Afghanistan and in these regions there were inscriptions carved in Aramaic and Greek as well as in Prakrit.



MAJOR INDO- ARYAN LANGUAGES



EDICTS OF ASHOKA

Aramaic is the language Jesus spoke, it's a Semitic language from the Middle East. It becomes understandable when we recognize that Aramaic had been used as the language of civil administration by the Achaemenid Persian Empire. That's why it was there. This is the remains of a Persian practice. Persian of course being Indo-European, using a completely unrelated language for their civil administration, their written communications.

The use of Greek of course is part of the legacy of Alexander the Great, who invaded this region, dying about a century before the time of King Ashoka.

Those are bizarre and wonderful examples, the Aramaic and Greek, and there are even bilingual ones, and they can be compared to the Prakrit to help translate one another. It's a very rich corpus of multilingual material, but one of the most invaluable legacies of King Ashoka is that his inscriptions have allowed scholars to map out the variety of different Prakrits spoken on the Indian subcontinent during his time.

Just staying with northern India, it's possible to distinguish the linguistic and phonological features of Eastern versus Western Prakrits in the middle of the third century, so depending on whether you use the short chronology, the long chronology, about a century or two after the time of the Buddha.

Where does Pali fit into this picture? It's certainly to be classified as one of the many Middle-Indic languages, it belongs to that general territory, linguistically. And thus, in a certain sense it's a Prakrit. But when we map Pali onto the picture provided by the Ashoka inscriptions, the Indological specialists discover something very peculiar, for Pali does not fit the normal profile of an ordinary Prakrit.

Instead, as the great Pali scholar Oskar von Hinüber has shown, it exhibits a mixture of eastern and western features, including some terminology that's thought to go back to old Magadhi itself, even older language. The majority of the characteristics of Pali in fact line up with - what we know from the inscriptions - would be Western Prakrits, the Girnar Prakrit, not Eastern, and this already would call into question whether it was anything close to the language the Buddha spoke since the Buddha lived and taught in the eastern part of this region. But it's a mixture, it contains some Eastern and a majority of Western Prakritic features.

Pali, in other words, is a hybrid language containing a peculiar mixture of dialectal forms. To sum up the work of Oskar von Hinüber and others, this fact of the hybridity of Pali, has some striking implications.

It's clear that it was not the language the Buddha spoke.

But not only did the Buddha not speak Pali, no one spoke Pali. Pali is, as Oskar von Hinüber calls it, an artificial language cobbled together from a variety of spoken languages to come up with what he called it a “lingua franca”, for the purpose of disseminating Buddhist teachings across dialectal or perhaps language lines. So, it's a compromise language made up of pieces from different languages in India at the time.

It's meant to be a scriptural lingua franca, but used originally and broadly only for religious texts, not only for the canonical scriptures attributed to the Buddha, but later for commentaries and texts like the Dipavamsa from the third or fourth century or the Mahavamsa from fifth century, composed to tell the history of Buddhism and of course, the history of Sri Lanka in the process, composed in Sri Lanka much later.

So Pali is already representing a bit of a divergence from the language policy mentioned at the beginning. It's no one's language, and ideally everyone's language, at least as a

language for the study and recitation of Buddhist teachings and texts. So, it might be a language of compromise, this would be another way to put it.

Now, again there's another rather surprising implication of this, maybe even shocking to some people who consider Pali is the real Buddhavacana, and that is that the texts in Pali cannot in any sense be described as the originals.

Texts in Pali are not original teachings in anyone's dialect. They are translations, or at least transpositions, into different wording using various Prakrit elements. They're “translations”, from one very closely related dialect to another. The tasks are much easier, obviously, than if you're translating from an Indian language into Chinese or Tibetan, but it's a kind of translation then there.

So, already with Pali, we see evidence of the attempted change of religious text from one, or probably many different languages being used for this, even in the region of northeastern India, into a “fixed” scriptural language.

The word “Pali” is the subject of a very old misunderstanding, not ours, we inherited it. This is a misunderstanding in medieval Asian traditions. The word doesn't appear in the Pali Tripitaka / Tipitaka. It's a commentarial word, it comes from a later layer of Buddhist literature, and when it does appear, it's the phrase that appears in “pali bhasha”, just the Pali language.

It sounds and looks like the name of a language like the French language, the Sanskrit language or the Chinese language, but it's not. Pali, doesn't mean a language name, it means texts, specifically, it means scriptural religious texts, not necessarily written down, but so Pali Basha is the language of the texts, which again reminds us this is not a spoken vernacular, this is a textual language created for a very specific purpose.

So, at some point, Oskar von Hinüber has traced this, K.R. Norman has also talked about it. Asian Buddhists began to misunderstand this also as the name of the language. They forgot that Pali meant texts, and a particular kind of text, and they began to take it in parallel to other such expressions as the name of the language itself. If that language have had a name, it would probably have been Buddhist Hybrid Prakrit or something of that sort.

When these oral traditions were transposed or translated into Pali, this is hardly the kind of linguistic task, or the dramatic change that we see in for example the translation of the

Hebrew Bible into Greek, the Septuagint, or the transfer of the teachings of early Christians from Aramaic again into Koine Greek, spoken Greek, the New Testament. These are shifts from one language family to another. Within India, what we're seeing is shifts among Prakrit, shifts from one Middle-Indic language to another as Buddhism extended into further regions. So, when did Pali become a church language? We don't know exactly, it's impossible to give us a definite date, but there's one very important benchmark in the history of the Pali tradition, and that is that the Pali scriptures were written down toward the end of the first century of the Common Era, sometime during the 20s we're told in a later traditional source. Why were they written down? This was done in Sri Lanka. The usual argument has been that there was a fear that as long as the tradition is oral, if the last monk who can recite the Digha Nikaya dies in a famine, you've lost it. If there's famine, if there's civil war, if there's an epidemic and your tradition is dependent on the living memory of people reciting these texts, that could be the end of the line for part or all of your canon, ultimately. So that's the normal reason that's given for this.

Steve Collins more recently has argued that this was actually, an intra-Nikaya feuding and that those who wrote these Pali texts down were trying to get an upper hand in the battle. Well, they certainly did. The Pali tradition has grown larger than life in its impact in the contemporary world and in the intervening centuries.

But be that as it may, what is important to pay attention to here is what a radical departure this is from earlier Buddhist practice, not only Buddhist practice, but religious practice throughout India.

Sacred texts were not written down.

We're familiar with how the Vedic tradition was memorized and passed down orally, but it's not simply that they were good at this. It is on the one hand that if you pass it down orally, you can determine very specifically who gets access, and not everyone was eligible to get access to the Vedic text. So, you can keep a kind of esoteric aspect to the transmission if you do this. You don't lose the text in some unqualified person who picks it up and reads it, finds it.

But more than this, there's a fascinating passage in a Vedic text, or later Vedic text, the *Aitareya Aranyuka*, that actually describes a list of things that disqualify you for performing Vedic rituals, in other words you have to go and purify yourself before you can then enter

this sacred realm. And the list has a lot of stuff that we would expect: touching a corpse, coming into contact with blood, having sexual intercourse, but the next one is: writing, including erasing writing. All of these were considered polluting activities from a Brahmanical perspective. Now that tells you something about how appalling it would have been to write down the Vedas, or the Upanishads or for anyone who is part of this world, to think about taking a sacred text and putting it into writing.

The *Aitareya Aranyuka* is quite unambiguous about this being a polluting activity. So what an amazing divergence, or deviation from Indian common practice this was, to write down the sacred texts of Buddhism. Note where it took place. It did not take place in Magadha, it did not take place in Madyadesa, the central area. It did not take place in the subcontinent. It took place in Sri Lanka, far from the Brahmanical heartland, out of the area where those constraints would have been fully in place.

That’s interesting and it becomes more interesting when we look at the other place that we know of Buddhist texts being written down. Amazingly, it's at about the same time, in the first century BCE and at the complete opposite end of the Indian subcontinent, that is in northern Pakistan and southern Afghanistan, in a region known as Gandhara and that language is the Gandhari language.

For Pali, we have legend, tradition, history maybe, a story of the writing down of the canon in the 1st century BCE. The manuscripts that are available to us are well over a thousand years later. There are two pages apparently, from an 8th to 9th century that were found in Nepal. Otherwise, it's mostly from the 15th century onward, and as Oskar von Hinüber points out, the climate is a factor, things and particularly palm leaves don't survive so well if they're not buried in a desert or somewhere at least a little colder and less humid.

For Gandhari, we have 1st century BCE manuscripts carbon dated to that old. These are the oldest Indian manuscripts, not carvings on stone, the oldest Indian literary text we could say, written on birch bark, that was the medium of choice in this area, and in the Gandhari language. So, we have actual physical evidence of Buddhist sacred texts being circulated in writing at the other end of the Indian cultural sphere. Gandhari is another Prakrit, it's from this area where Greek and Aramaic were being used, and interestingly enough, it's written in a script derived from Aramaic. It's written from right to left but it's a Middle-Indic language, it's a Prakrit, so it's a cousin of Pali but very, very different obviously, it’s a northwestern style of Prakrit.

And here, in the case of both Pali and Gandhari, we can see that the text as we have them are translations. In Gandhari there are words that came over in forms that aren't allowable in Gandhari. Oskar von Hinüber points out the same thing in Pali. So, we can see kind of traces of what these might have come from. It’s a remarkable picture. These are the only two Prakrits for which we have concrete evidence that they were written down in the Indian cultural sphere.

And as we'll see they both spread way outside their home turf. So, the Pali story is fairly straightforward: from Sri Lanka it eventually gets exported to South East Asia and becomes the church language there. The Gandhari story is becoming clearer in recent years with current research showing more and more evidence that the early textual tradition arriving in China may not have been by any means exclusively Gandhari, but certainly there's a lot of evidence that some of the texts were in Gandhari.

Some of the others may have been oral, we know there's a mixture of oral and written material coming to China. So, sometimes we can see in Chinese translations concrete evidence, like a transcription in pronunciation of a term that points to Gandhari and doesn't line up with other Prakrits.

We have Pali going out from South Asia, from Sri Lanka to Southeast Asia, obviously by the sea route basically, and then in the northwest, Gandhari making its way along the Silk Road to China.

And there have been texts found in Pakistan and Afghanistan in recent years, but already almost a century ago we had evidence of Gandhari from eastern central Asia, that is the Tarim Basin, the southern part of what's now Xin Jiang in the People's Republic of China, one of those being Gandhari *Dharmapada* version of the *Dhammapada*. It's clear that Buddhist texts were being exported through this region to China.

So, Pali and Gandhari that's what we know about of languages in which Buddhist literature was transmitted and written down, and the fact that it was written down has helped a lot to preserve these outside India, and the fact then that they were exported beyond their native region meant that other cultures kept them alive when Buddhism and some of these areas began to die out.

So again, at the extremes of the Indic cultural sphere, coming into China, we have evidence of awareness of Buddhism already in the first century in secular sources. History and poetry mentioned these Buddhist monks and so forth. But the translation of Buddhist texts into languages other than the language of India began in the middle of the second century. It's a remarkable account, it's just one of the amazing things in cultural transmission in world history.

It's interesting to note about the Sanskrit word "pustaka", in Pali it's "pitika". Pustaka appears to be an Iranian loan word in Sanskrit. It's a word that appears first in Buddhist literature, and mainly in Buddhist literature, and is used for a category that didn't exist in India before, that is written religious texts. They didn't have anything to call them, they needed a word and they called them "pustaka".

So, if we read those passages in for instance the 8000-line perfection of wisdom, the *Astasahasrika Prajnaparamita Sutra* about writing the text, it's the word they use there is make it into a book, made into a pustaka when you write it down.

What does pustaka mean, or what the Iranian antecedent means? It means leather, hides. Iranians wrote on hides, cow hides. Traditionally the Zoroastrian scriptures were recorded on cowhide, so this was a writing material in that part of the world. So, a pustaka was something on hide.

The Buddhists did not follow in those footsteps, they did not write their scriptures on hide. In the northwest they wrote them on birch bark, in Sri Lanka they wrote at them on palm leaves, as they did in the Indian mainland. But this is a really good example of how foreign that concept was. They had to borrow a word that it's original meaning probably was forgotten quite quickly in India, from the Iranian cultural sphere, from their neighbors, just slightly to the north, to describe something that did not exist in India before: written religious texts.

Jan Nattier
Hua Hin, Thailand

***

Jan Nattier did her undergraduate work in comparative religion (specializing in Buddhism) at Indiana University, where she also began graduate training in the Department of Uralic and Altaic Studies. She completed her PhD at Harvard University under the Committee on Inner Asian and Altaic Studies (specializing in classical Mongolian and Tibetan). She has taught at Macalester College, the

University of Hawai'i, Stanford University, Indiana University, and the University of Tokyo, in addition to serving as a member of the International Research Institute for Advanced Buddhology (Soka University). Her monographs include *Once Upon a Future Time: Studies in a Buddhist Philosophy of Decline* (Asian Humanities Press, 1991), *A Few Good Men: The Bodhisattva Path According to the Inquiry of Ugra* [*Ugraparipṛcchāsūtra*] (University of Hawai'i Press, 2003), and *A Guide to the Earliest Chinese*.

***

Borobudur Jinabhumi. Revision: January 2019.