

THE BUDDHIST FORUM, VOLUME V

**A PHILOLOGICAL APPROACH
TO BUDDHISM**

The Bukkyō Dendō Kyōkai Lectures 1994

K.R. Norman

THE INSTITUTE OF BUDDHIST STUDIES, TRING, UK
THE INSTITUTE OF BUDDHIST STUDIES, BERKELEY, USA
2012

First published by the School of Oriental and African Studies (University of London), 1997

© Online copyright 2012 belongs to:
The Institute of Buddhist Studies, Tring, UK &
The Institute of Buddhist Studies, Berkeley, USA

British Library Cataloguing in Publication Data

ISBN 0-7286-0276-8
ISSN 0959-0595

CONTENTS

The online pagination 2012 corresponds to the hard copy pagination 1997

Foreword.....	vii
Abbreviations.....	ix
Bibliography.....	xi
I Buddhism and Philology.....	1
II Buddhism and its Origins.....	21
III Buddhism and Oral Tradition.....	41
IV Buddhism and Regional Dialects.....	59
V Buddhism and Writing.....	77
VI Buddhism and Sanskritisation.....	95
VII Buddhism and Aśoka.....	113
VIII Buddhism and Canonicity.....	131
IX Buddhism and the Commentarial Tradition.....	149
X Buddhism and Philology.....	167
Word Index.....	185

ABBREVIATIONS

AMg	Ardha-Māgadhī
AO	Acta Orientalia
B.C.E.	Before Common Era
BHS	Buddhist Hybrid Sanskrit
BHSD	Buddhist Hybrid Sanskrit Dictionary
BHSG	Buddhist Hybrid Sanskrit Grammar
BSO(A)S	Bulletin of the School of Oriental (and African) Studies
BSR	Buddhist Studies Review
BUp	Brhad-Āraṇyaka Upaniṣad
C.E.	Common Era
CP I–V	K.R. Norman. Collected Papers, Volumes I–IV (1990–94)
CPD	A Critical Pāli Dictionary, Copenhagen
CUP	Cambridge University Press
DEDR	T. Burrow & M.B. Emeneau, <i>Dravidian Etymological Dictionary</i> (Revised edition, 1984)
DPPN	G.P. Malalasekera, <i>Dictionary of Pāli Proper Names</i> (1937–38)
EV I, II	K.R. Norman, <i>Elders' Verses</i> I (1969), II (1971)
GD I, II	K.R. Norman, <i>The Group of Discourses</i> I (1984), II (1992)
IJ	Indo-Iranian Journal
IT	Indologica Taurinensia
JA	Journal Asiatique
JBuRS	Journal of the Burma Research Society
JOI(B)	Journal of the Oriental Institute (Baroda)
JPTS	Journal of the Pali Texts Society
JRAS	Journal of the Royal Asiatic Society
m.c.	metri causa
MIA	Middle Indo-Aryan
MRI	<i>Minor Readings and Illustrator</i>

MW	M. Monier-Williams, <i>Sanskrit-English Dictionary</i> (1899)
OUP	Oxford University Press
PDhp	Patna Dharmapada
PE	Pillar Edict
PED	The PTS's Pali-English Dictionary
Pkt	Prakrit
PTS	Pali Text Society
RE	Rock Edict
SBE	Sacred Books of the East
Skt	Sanskrit
SOAS	School of Oriental and African Studies (London)
SWTF	Goerg von Simson (ed.), <i>Sanskrit-Wörterbuch der buddhistischen Texte aus den Turfan-Funden</i> (1973–)
Überbl	O. von Hinüber, <i>Das ältere Mittelindisch im Überblick</i> (1986)
WZKSO	Wiener Zeitschrift für die Kunde Südasiens

Abbreviations of the titles of Pāli texts are those adopted by the CPD.
Editions are those of the PTS, unless otherwise stated.

VI Buddhism and Sanskritisation

In the fifth lecture I referred to the fact that the writing down of the Theravādin canon in the first century B.C.E. seems to have put an end to the Sanskritisation of that canon which had begun in the decades before the reign of Vaṭṭagāmini Abhaya, although, paradoxically, writing seems to have led to the appearance of other texts, which, as far as we can tell now, were actually composed in Sanskrit.

In this lecture I want to consider the reasons for the process of Sanskritisation, the way in which it was effected and the result which it had upon Buddhist texts and Buddhism itself.

I must start with a definition. What exactly do I mean by Sanskritisation? I use the word in two senses. In the first place, I talk, in a broad sense, about the Sanskritisation of Buddhism, when I am discussing a particular phenomenon, namely the way in which Buddhism, which had started as a revolt against the social and religious system which was exemplified by the use of Sanskrit for literary and religious purposes, now began itself to embrace Sanskrit as a medium for the propagation of the *Buddhavacana*.

In the second place, Sanskritisation means the use of Sanskrit in Buddhist texts as a replacement for the dialects of Middle Indo-Aryan in which the Buddha's teachings had previously been transmitted for some hundreds of years. In this sense, the term is applicable to the whole range of Buddhist texts starting from those in a Prakrit which contains a very small amount of Sanskrit, or Sanskrit-like, forms in it, through a range of texts which are in a variety of languages which might be regarded as Sanskritised Prakrit or Prakritised Sanskrit, sometimes called Buddhist Hybrid Sanskrit, to texts which are in pure classical Sanskrit, in accordance with Pāṇinian grammar.¹ The language of the last group is classified by some as Buddhist Sanskrit, because the texts are written by Buddhists about some aspect of Buddhism or Buddhist history, and perhaps contain items of vocabulary which are specifically Buddhist. We can therefore classify Sanskritised texts under three headings: (1) texts written in a Middle Indo-Aryan dialect into which some Sanskritisms have been inserted; (2) texts originally written in a Middle Indo-Aryan dialect which have been

¹ For a survey see von Hinüber, 1989B, 341–67.

translated into Buddhist Hybrid Sanskrit or Buddhist Sanskrit; and finally (3) texts *composed* in Buddhist Hybrid Sanskrit or Buddhist Sanskrit.

In this lecture I want to concentrate upon those aspects of Sanskritisation which are exemplified in the first two of these classes. I shall be considering for the most part the problems which arose when texts in Middle Indo-Aryan dialects were converted into other dialects with a Sanskritic content, small or large, or into Sanskrit. I shall not be referring to those texts which were composed as original works in Buddhist Hybrid Sanskrit or Buddhist Sanskrit, except in so far as some of the vocabulary they employ perhaps owes something to Middle Indo-Aryan.

This pattern of Sanskritisation was not restricted to Buddhist texts. We find something very comparable in the language of the inscriptions found in North India. The earliest inscriptions, those of Aśoka, are written in a variety of Prakrit dialects. Those of the early centuries of the Common Era are in a mixture of Prakrit and Sanskrit. From the fifth century C.E. onwards inscriptions are written in classical Sanskrit. If we examine the inscriptions in the middle phase at Mathurā, we find that they do not show a steady progression from Prakrit to Sanskrit. What we know of the languages of North India, during the period covered by the inscriptions, indicates that the local population spoke some dialect of Middle Indo-Aryan. Donors would presumably dictate in Prakrit what they wanted to have carved on their donations, and as it became fashionable to write inscriptions in Sanskrit the scribes would “translate”, to the best of their ability, into that language. An inscription in bad Sanskrit is, therefore, not necessarily older than one in good Sanskrit. The difference in quality may simply represent the ability of one donor to employ a better educated scribe than the other. Similarly, the fact that one Buddhist text is in better Sanskrit than another does not necessarily mean that it was Sanskritised at a later date.

Why did Sanskritisation begin? If we consider the partial Sanskritisation of the language of the Theravādin canon, we must, I think, agree that there is no obvious reason why Sanskritisation should have started in Sri Lanka in the first century B.C.E. We have no information to make us believe that there was a strong pro-Sanskrit movement in the island then. There is, for example, no hint that the Abhayagirivihāra, the rival of the Mahāvihāra, was making use of Sanskrit at this time. The Sanskritisation of Pāli can hardly have started spontaneously, in the absence of any reason, and we must assume that it started under the stimulus of Sanskrit elsewhere, presumably in North India. There was no obvious reason why a language used in Sri Lanka should have been influenced by anything happening in North India, and so we must assume either that the process of Sanskritisation started in the Theravādin canon before it was transported to Sri Lanka, or, more probably, that there was still a strong

connection between North India and Sri Lanka after the introduction of Buddhism to the island.

The Pāli chronicles report that Buddhism was taken to Sri Lanka during the reign of Aśoka. We know that, for secular purposes at least, Prakrit was preferred to Sanskrit at the time of Aśoka, and it was only after his death that Sanskrit began to regain its position of predominance, although we should perhaps note the Sanskritising tendency of the Aśokan scribe at Girnār, even if we are unable to decide whether the Sanskritisms there are features of the local language (which might have been rather archaic) or insertions by a Sanskritising scribe. In favour of the latter, we should note that there is a strong case for believing that the *-r-* forms at Girnār owe more to the scribe's views of what was appropriate than to the actual vernacular spoken in the area.² If we believe that the second century B.C.E. saw the appearance of new Buddhist texts, setting out new teachings, composed by new sects, in Sanskrit, the language of culture and literature, then we might well believe that that was the time when, in competition with these new Sanskrit works, those schools of Buddhism which had hitherto used a Middle Indo-Aryan dialect, began to make changes in the languages of their teachings in order to rival the growing use of Sanskrit by other sects, and to make their teachings available to the same classes of readers.

There were various degrees of Sanskritisation: (1) in its simplest form it involved the partial restoration of Sanskrit phonology, perhaps the restoration of a number of Sanskrit forms to consonant groups, and possibly in addition the restoration of long vowels which had been shortened in Middle Indo-Aryan, e.g. before consonant groups; together with (2) the restoration of Sanskrit morphology, e.g. "correct" verbal and nominal endings; and perhaps (3) the substitution of Sanskrit vocabulary in place of Middle Indo-Aryan vocabulary. This might include the removal of Eastern forms which, as I pointed out in the fifth lecture, represented some of the oldest elements of Buddhist vocabulary, e.g. *āvuso*.

Further levels of Sanskritisation would include: (4) the insertion of the correct Sanskrit sandhi forms, i.e. those demanded by the rules governing word juncture. We find that hiatus is avoided by the insertion of particles, or by the rearrangement of the order of words, or by a change of vocabulary, (5) and, in

² von Hinüber (*Überbl* § 15) writes of an archaising dialect. I would rather think of an archaising scribe. The apparent resemblances to the Girnār (= G) version of the Aśokan inscriptions are not conclusive, since we face there the same problems as in Pāli. The G version represents a "translation" of an Eastern version which we may assume was sent from Aśoka's capital Pāṭaliputra. We have no certain way of telling whether the G version accurately represents the vernacular spoke in Western India at that time, or whether it merely represents the scribe's attempt to produce what he thought was appropriate in the circumstances. It is arguable that the scribe at G was trying to Sanskritise, to the best of his ability, what he had received.

metrical texts, there is the avoidance of irregular metre by changes of word order, or of vocabulary.

When we come to examine the Sanskritisations in Pāli, we have first of all to decide whether they are in fact Sanskritic features which were inserted by the redactors, or whether they might not be original features of the language, i.e. are the Sanskritisms in Pāli retentions or restorations of Sanskritic features? If we thought that such Sanskritisms were remnants of an earlier form of the language of the canon, then we could regard them as genuine archaic features in the language, and we might then define Pāli as one of the oldest (linguistically speaking) of the Middle Indo-Aryan dialects, in as much as it was a form of Sanskrit with some Middle Indo-Aryan developments in it.

An investigation, however, shows that some of these Sanskritisms in the Theravādin canon are incorrect back-formations, e.g. *atraja*, which is probably a misinterpretation of *attaja* (< *ātmaja*).³ Although some of the forms with the consonant group *br-* are correct, e.g. *brāhmaṇa*, others of them are non-historic, e.g. *brūheti*.⁴ Even some of the forms which are correct, e.g. the absolutive ending *-tvā* and *br-* in *brāhmaṇa*, can be shown to be unoriginal in certain contexts. There is evidence for the absolutive ending *-ttā*,⁵ which we can assume was the regular absolutive ending in the dialect before the restoration of the *-tvā* ending, and both the metre and the etymologies which are given for *brāhmaṇa* in the canonical texts⁶ show that at an earlier time the word occurred with initial *b-*, not *br-*. We can therefore conclude that these forms, and probably all other Sanskritic features, are deliberate attempts at Sanskritisation, made at some time during the course of the transmission of the canon. It is therefore clear that it is not correct to speak of them as retentions. They are features which have been restored to the texts by scribes or reciters who were trying to change into Sanskrit the language which they had received in their exemplars.

A close examination of the Sanskritic features in the Theravādin canon suggests that the process took place in two phases, separated by some centuries.

To the early phase we can allot the Sanskritisation of some consonant groups, e.g. *tv* in the absolutive ending *-tvā*, which I have just mentioned, and *tr* in the suffix *-tra* of the locative pronouns, and the *st* in *utrasta* and *bhastā*, etc.⁷ Some of these changes were probably made as early as the writing down of the canon,

³ See *PED*, s.v. *attaja*.

⁴ Skt *br̥ṃhayati* should have developed > **bumhayati* > **būhayati/būheti* (with the *-um/-ū-* alternation), as in *AMg*.

⁵ See von Hinüber, 1982, 133–40 for absolutives in *-ttā*, and Norman, 1980B, 183, note 21, and 1985, 32–35, for absolutives in *-ttā*. I assume that the latter are m.c., but von Hinüber (*Überbl* §498) lists an absolutive in *-tā*.

⁶ e.g. *bāhita-papo ti brāhmaṇo*, *Dhp* 388.

⁷ See Norman, 1989C, 369–92 (377–79) (= *CP* IV, 46–71 [56–58]).

and they are certainly earlier than the commentaries, because the commentators refer to the variation between the suffixes *-ttā* and *-tvā*. It is noteworthy that *r* is not restored in groups with *p*. The fact that there are no forms with *pr* in Pāli, in contrast to the Gīrnār version of the Aśokan inscriptions, where *pr-* is the most common group containing *-r-*, suggests that, despite all claims to the contrary, there is probably no connection between Pāli and the Gīrnār dialect of the Aśokan inscriptions.

Other Sanskritic features include the retention or restoration of long *-ā-*, e.g. in certain compounds and derivatives of *vāk-* “speech”, even when the resultant form of the word goes against the pattern of the dialect by producing a long vowel before a consonant group, e.g. *vākya*, which by the rules might have been expected to develop into **vakka*, with a short *-a*.⁸ The fact that *vākya* and comparable forms are back-formations is supported by the existence of other derivatives of *vāk* where the expected shortening of long *-ā-* does occur before a double consonant. We can deduce that the redactor did not recognise the existence of the word *vāk* in such compounds, and consequently did not restore the long vowel.

A later phase of Sanskritisation took place under the influence of the grammarians, centuries after the first phase, and probably for the most part after the appearance of the *Saddanīti*, Aggavaṃsa’s grammar of Pāli which was written in the twelfth century. Some of these Sanskritisations are in the field of vocabulary or morphology, and many are concerned with sandhi. Sandhi in Pāli is quite different from Sanskrit sandhi, in that it is much more flexible. Since final consonants have disappeared, sandhi in Pāli consists of the contraction of the final vowel of a word, including a nasalised vowel, with the initial vowel of the following word. Such contractions produce a wide range of crasis vowels, which are not always predictable.

The grammarians, with their extensive knowledge of Sanskrit, tried to bring Pāli more in line with Sanskrit grammar, and copyists inserted into the manuscripts the forms which the grammarians prescribed. Final *-i* and *-u* which had remained in hiatus were (sometimes) changed to *-y* and *-v*. If the final vowel had been *-e* or *-o*, then an indication of this was sometimes given by prefixing *y*

⁸ There are also examples of the retention or restoration of a long vowel before a doubled consonant, which goes against the general rule of two morae. Probably a word such as *dātta* is a borrowing from a North-Western dialect (see Turner, 1973, 424–28 [= *Collected Papers*, 430–35]); Geiger, 1994, § 7 (despite Geiger, *dātta* is quotable from *Mil* 33, 3 foll.). We also find a long *-ā-* occurring before a doubled consonant as a result of the crasis of two vowels in a compound, and also in such words as *yvāssa* (= *yo assa*, *M I* 137,11) and *tyāssa* (= *te assa*, *Dhp-a I* 116, 20) as products of a non-historical sandhi process, as well as *gavāssa ca* (= *gavā assā ca*, *Ja III* 408,21). All these examples seem to be based upon a knowledge of Sanskrit, since they go against the general rules of MIA dialects.

or *v* to the vowel which had evolved by contraction. The fact that these changes are later additions to the texts, and are not the result of the earlier phase of Sanskritisation, is shown by the fact that the manuscripts are rarely consistent about such insertions. If the contraction of final and initial vowels led to a long vowel occurring before a consonant group, this was sometimes kept, even though it offended against the law in Middle Indo-Aryan which demanded a short vowel in this position.

It is clear that the reasons for these two phases of Sanskritisation of Pāli were quite different. It is probable that the reason for the first was the appearance of other schools using Sanskrit, as I have suggested, but the reason for the second is that grammarians who knew enough about Sanskrit to be able to model their grammars upon Sanskrit, and in fact to quote extensively from the Sanskrit grammarians, thought that they should try to make Pāli more like Sanskrit.

Sanskritisation caused problems of a different nature from those which arose in the period of oral tradition. The nature of the development of Middle Indo-Aryan meant that there was an increase in the number of homonyms. A glance at Sheth's Prakrit-Hindi Dictionary (*Pāiasaddamaḥṇavo*) shows, for example, that there are at least ten Sanskrit words which can develop into Prakrit *saya-*, and, if we look at compounds, there are no less than 48 Sanskrit equivalents for the Prakrit compound *para-(v)vāya*. Although a specific context would immediately rule out many of these, it is obvious that a redactor translating into Sanskrit might well have difficulties when trying to decide between homonyms.

Problems arose even in the limited amount of Sanskritisation which we find in Pāli. The problem for the redactor was to decide which of two or more possible forms to choose when Sanskritising. There is sufficient evidence for us to deduce that in the language of the Theravādin canon before it was Sanskritised the absolutive ending was *-ttā*, as I have mentioned. This would be identical with the form in Ardha-Māgadhī, the language of the Jain canon. A form such as *kattā* < Sanskrit *kṛtvā* "having done" was therefore identical in form with, i.e. was a homonym of, *kattā*, the agent noun "a doer" from the same root *kṛ-*. There is a *pāda* which occurs twice in Pāli which in its earlier form contained, we can deduce, the word *kattā*. In one context *kattā* was taken as a verb and Sanskritised as *katvā*. In the other the tradition interpreted it as the noun *kattā*, and so retained it in that form and explained it as a noun.⁹

Exactly the same happened with the word *chettā*, from the root *chid-* "to cut". Here, in one and the same context, the tradition is ambivalent. In the edition of

⁹ *katvā ti ... karitvā*, Ja II 317, 21' (ad *katvā*, 317, 14*); *kattā kāraṅko*, Ja IV 274, 9' (ad *kattā*, 274, 2*).

the canonical text we find the absolutive *chetvā* “having cut”, but the commentary reads the noun *chettā* “a cutter”, and explains it accordingly.¹⁰

When a Sanskritising recensionist was faced with the word *santa*, he had the option of assuming that it was the present participle of the verb “to be”, which gets the meaning of “good”—you will know that the feminine of this (*sāti*) is used of the good woman who is so devoted to her husband that she immolates herself upon his funeral pyre—or the past participle of the verb *śam* “to rest”, i.e. the equivalent of Sanskrit *śānta* “at rest, peaceful”. In some cases it is possible, even probable, that in a Middle Indo-Aryan dialect *santa* was intended to be ambiguous, but in Sanskrit one or other of the two meanings has to be preferred.

The same problem can arise with the Middle Indo-Aryan word *dīpa*, which may be the equivalent of Sanskrit *dīpa* “lamp” or *dvīpa* “island”. As a parallel to the verse in the Dhammapada¹¹ “a wise man should make a *dīpa* which the flood will not overwhelm”, the Sanskrit Udānavargaha¹² has *dvīpa* “island, but the Chinese version¹³ has “lamp”, showing that it is based upon a Sanskrit version which had *dīpa* “lamp”.

A verse in the Dhammapada¹⁴ tells us that someone who has *pīti* in the *dhamma* sleeps happily. The word *pīti* is ambiguous since there are two words *pīti* in Pāli, one from Sanskrit *pīti* “drinking” from the root *pā* “to drink”, and the other from *prīti* “joy” from the root *prī* “to please”. It seems very likely that both meanings are intended in this verse. The second *pāda* tells us that he sleeps with a clear mind, perhaps unfuddled by drinking the *dhamma* as opposed to the intoxication he would have experienced if he had drunk strong drink. In the last *pāda* of the verse, however, we read that the wise man delights (*ramati*) in the *dhamma*,¹⁵ which suggests that *pīti* is also to be taken as “joy”. The Pāli commentator presumably did not see the possibility of the word play, and explains it only as drinking the *dhamma*.¹⁶ The redactor of the Sanskrit Udāna-varga had to choose between *pīti* and *prīti*, and perhaps because of the idea of “delight” in the last *pāda*, or perhaps because he was following a different commentarial tradition, he decided to read *prīti*.¹⁷

¹⁰ *chetvā*, Th 1263; *chettā chedako*, Th-a III 199, 11.

¹¹ *dīpaṃ kayirātha meddhāvī yam ogho nābhikīrati*, Dh 25.

¹² *dvīpaṃ karoti meddhāvī*, Udāna-v 4.5.

¹³ Quoted by Brough, 1962, 209 (ad GDhp 111).

¹⁴ Dh 79.

¹⁵ *dhamme sadā ramati paṇḍito*, Dh 205.

¹⁶ *dhammapīti ti dhammapāyako dhammaṃ pivanto ti attho*, Dh-a II 126, 15.

¹⁷ The Gāndhārī Dharmapada is written in a dialect which retains many of the consonant groups of Skt, and there we find *dhama-pridi*, GDhp 224. See Brough, 1962, 244.

There was some doubt about the Sanskritisation of the Pāli word *nekkhamma*. As the *Pali-English Dictionary* points out, although the word is derivable from a hypothetical Sanskrit form **naiṣkramya* “going out”, from the root *kram-* “to go”, i.e. departing from the householder’s state to the houseless way of life, there are also word plays which seem to be based upon the meaning “the state of being without desire” (**naiṣkāmya*) from the root *kam-* “to desire”. I see no reason to doubt that we are in fact dealing with two homonyms here. In Buddhist Sanskrit, however, we find only the form *naiṣkramya*,¹⁸ even in conjunction with the word for desire (*kāmeṣu*).¹⁹

We sometimes find that Pāli and Sanskrit technical terms do not correspond, and in such cases we may surmise that the lack of correspondence may be due to Sanskritisation from a homonym. The Buddhist Hybrid Sanskrit equivalent of Pāli *sammappadhāna* “right effort” is *samyakprahāṇa* “right abandoning”. It has been suggested²⁰ that the Sanskrit version is an incorrect backformation from an Middle Indo-Aryan form **samma-ppahāna* which might stand for either *-pradhāna* or *-prahāna*. In view of the fact that such word-plays are very common in Indian literature, we cannot exclude the possibility that the authors of the passages which contain such words were making using of the ambiguity, and, if this was so in this particular case, we can see that neither the Pāli nor the Sanskrit form is capable of giving the double meaning which the author intended.

Sometimes commentators are aware of the fact that a word can have two meanings, and they incorporate both senses in their explanations. It is suggested by some²¹ that the Buddhist tradition deliberately capitalised on, or even exploited, any ambiguity which a Middle Indo-Aryan form might have, but it seems to me that using such phrases as “deliberately capitalised” is perhaps the wrong way of looking at the matter. In my view it is rather taking advantage of the situation. The speakers of Middle Indo-Aryan dialects were not likely to know that a given item in their vocabulary had developed from two or more Sanskrit words. They merely knew that it had two (often quite different) meanings, and the commentaries tried to make use of this fact when giving their explanations. It was only when someone proficient in Sanskrit tried to Sanskritise that the need to choose between two possible Sanskrit antecedents arose. We should not be too surprised if a recensionist made the wrong decision,

¹⁸ See Sasaki, 1986, 3.

¹⁹ Mvy 6444 “renunciation as regards desires”, quoted in BHSD.

²⁰ See Gethin, 1992B, 70.

²¹ Gethin, 1992B, 72.

especially since,²² in the example I have just mentioned, Sanskrit *pradhāna* does not have the meaning “endeavour”.

There are many other similar examples of wrong backformation in partially Sanskritised dialects, such as the language of the Theravādin canon and the Patna Dharmapada, and also in more highly Sanskritised languages. It should be noted that it is not always the Sanskrit version which is incorrect, and we can surmise that there was sometimes a commentarial tradition available to the translators, which gave information about the meaning and interpretation of the words they were concerned with.

The need to choose between homonyms quite often leads to a situation where puns and explanations which worked in Middle Indo-Aryan no longer worked in Sanskrit, or even in a Middle Indo-Aryan dialect with partial Sanskritisation. In the Dhammapada the word *brāhmaṇa* is explained by means of forms from the root *bāh-* “to remove”. A brahman is one “who has removed, got rid of, his sins”.²³ The connection between *brāh-* and *bāh-* is not as close as one would like for an etymology, but the answer to the problem is to deduce that this explanation was originally formulated in a Middle Indo-Aryan dialect where the development from *brāhmaṇa* was **bāhaṇa*, so that the pun on the meanings of the two verbs, which in their Sanskrit forms are *brh-* “to be strong” and *brh-* “to remove”, worked perfectly. The Sanskritisation to *brāhmaṇa* which occurred in Pāli, as I have already noted, had already spoiled the pun. Another pun found in the Dhammapada which does still work in Pāli, is that found in the etymology which is given for the word *samaṇa* “ascetic”. One is called a *samaṇa* “because one’s sins have been put to rest”,²⁴ where the *sam-* of *samaṇa* is punningly linked with the root *sam-* “to be at rest”. In a Sanskritised form, however, *samaṇa* becomes *śramaṇa*, which is from the root *śram-* “to make an effort”, while the Sanskrit form of the root “to rest” is *śam-*. The pun which worked well in Pāli is lost when we have *śramaṇa* explained by *śam-*.²⁵

The concept of the *pratyeka-buddha* is well known as the middle element in the triad: *buddhas*, *pratyeka-buddhas* and *śrāvakas*. The word is usually translated as “a *buddha* (awakened) for himself”, but this usage of *pratyeka* (*pacceka* in its Pāli form) is unusual, and it is not at all clear how it could acquire the meaning which is given for it. The concept of a similar type of Buddha is also found in Jainism, but there, in Jain Prakrit, the name is *patteya-buddha*. The variation in the name suggests that the concept of *pratyeka-buddhas* was borrowed into both Buddhism and Jainism, and we can therefore look for an

²² As Gethin notes (1992B, 71).

²³ *bāhitapāpo tii brāhmaṇo*, Dhṛp 388.

²⁴ *samītatā hi pāpānaṃ samaṇo tti pavuccati*, Dhṛp 265.

²⁵ *śamītatvāt tii pāpānāṃ śramaṇo hi nirucyate*, Udāna-v 11.14.

etymology for the word outside Pāli. It seems most likely that the earlier form of the word was *pacceya-buddha*, where *pacceya* is to be derived from Sanskrit *pratyaya* “cause”. A *pacceya-buddha* was “one who is awakened by a specific cause, a specific occurrence (not by a Buddha’s teaching)”, and both the Buddhist and the Jain tradition give lists of the occurrences which caused the awakening of the four most famous persons in this class. The Pāli form *pacceka* is an incorrect translation, probably a hyper-form, of *pacceya*, and *pratyeka* is a Sanskritisation of *pacceka*.

Sanskritisation led to hyper-forms, as a result of misinterpretations. Many Buddhist Sanskrit texts are entitled *sūtra*. To anyone who comes to Buddhist studies from classical Sanskrit studies, this name comes as a surprise, because, in Sanskrit, *sūtra* literature is a specific genre of literature, composed in prose, usually of a very abbreviated and concise nature, while Buddhist *sūtras* have an entirely different character. This difference is due to the fact that the word *sūtra* in Buddhist Sanskrit is a Sanskritisation of the Middle Indo-Aryan word *sutta*, which is probably to be derived from Sanskrit *sūkta*, a compound of *su* and *ukta*, literally “well-spoken”.²⁶ It would be a synonym for *subhāṣita*, which is the word used of the *Buddhavacana* by the emperor Aśoka, as we shall see in the seventh lecture, when he said: “All that was spoken by the Lord Buddha was well-spoken”.

Another wrong back-formation centres around the same word *sutta*. There is a verse in the Pāli Dhammapada where we are exhorted to *chinda sotam* “cut off the stream” (of *saṃsāra*), and the Sanskritisation in the Udāna-varga gives the same sense: *chindhi srotah*. The Patna Dharmapada, however, was doubtless dependent upon a Middle Indo-Aryan version which, through an orthographic variation, must have had the reading *sutta*, via *sōtta*. This was then Sanskritised as *sūtra*, so the *pāda* in that version tells us to “cut off the thread”, presumably interpreted as the thread of rebirth, since we find in Pāli that craving (*taṇhā*) is described as “the seamstress” (*sibbanī*), which joins us to *saṃsāra* by means of death and rebirth.²⁷

One of the best-known examples of Sanskritisation is the word *bodhisattva*, which is a back-formation from Middle Indo-Aryan *bodhisatta*. Anyone who knows anything about Sanskrit will realise that the translation which is commonly given for this word “one destined to be a Buddha”, or “one destined for awakening”, is, to say the least, unlikely, and we should note that Monier-Williams gives the basic meaning as “one whose essence is perfect

²⁶ See Walleser, 1914, 4, note 1; von Hinüber (1994, 132, note 28) follows Mayrhofer (1976, 492, s.v. *sūram* [online editor’s note: “ū” has an additional é above it]) in thinking this proposed etymology is unnecessary.

²⁷ See *EVI* 663.

knowledge”,²⁸ which would be more appropriate as an epithet for a Buddha than for a *bodhisatta*. It is noteworthy that the Pāli commentaries did not assume that the second element of the compound was the equivalent of Sanskrit *sattva*. They give derivations either from the root *sañj-* “to be attached to”, which should have given a Sanskrit form *bodhiṣakta* “one attached to *bodhi*”, or from the root *śak* “to be able” which should have given a Sanskrit form *bodhiśakta* “one capable of *bodhi*”.²⁹

We should note, incidentally, that wrong Sanskritisations are not restricted to Buddhist texts. One of the Sanskrit words for “pearl” is *muktā*, which literally means “released”. It is, however, more than likely that the Sanskrit form is a wrong Sanskritisation from a Middle Indo-Aryan form *muttā*, which occurs, for example, in Pāli, so that its Sanskritisation is an example of a folk etymology: a pearl is called “released” because it is released from the oyster. It has been suggested that Middle Indo-Aryan *muttā* is to be derived from Sanskrit *mūrta*³⁰ “coagulated, shaped, formed”, but in view of the wide-spread existence of a word *muttu* in Dravidian,³¹ with the meaning “pearl”, it is perhaps more likely that Middle Indo-Aryan *muttā* is a loan-word from Dravidian.

Resolved consonant groups, those consonant groups which have been separated into their constituent members by inserting an epenthetic (*svarabhakti*) vowel between them, were Sanskritised by the removal of the *svarabhakti* vowel, e.g. Sanskrit *kriyā*; “action” became Middle Indo-Aryan *kiriyā*, and was then Sanskritised back to *kriyā* again. Sometimes misunderstanding on the part of the translators led to forms such as *parṣat* being Sanskritised³² from Middle Indo-Aryan *parisā*, which was derived from Sanskrit *pariṣat*; on the assumption that *-i-* was an epenthetic (*svarabhakti*) vowel, and similarly *nyāma* “rule”, or “regulation” was, in a similar way, wrongly back-formed from Middle Indo-Aryan *niyāma*, which was identical with Sanskrit *niyāma*. This type of error can be seen in the oldest translations which we have in India, i.e. the Aśokan inscriptions, where one of the scribes wrote *hveyu* in place of the form *huveyu*, from the root *bhū-*, which he had received in his exemplar, presumably in the belief that *-u-* was a *svarabhakti* vowel.

Although I said earlier that the fact that one Buddhist text is in better Sanskrit than another does not necessarily mean that it was Sanskritised at a later date, nevertheless, in some cases we can see how Sanskritisation developed. We have, for example, an edition of an earlier version of the Udānavarga from Chinese

²⁸ MW, s.v. *bodhisattva*.

²⁹ See Bollée, 1974, 27–39 (36, note 27).

³⁰ See von Hinüber, 1993, 113, referring to Lüders, 1940, 179–90.

³¹ See *DEDR* 4959.

³² See Sander, 1985, 144–60.

Turkestan,³³ which is closer in its form, especially with regard to word order, to its Middle Indo-Aryan original than the later version from Chinese Turkestan,³⁴ which, in its attention to metre, its observance of sandhi rules, etc., represents a more correctly Sanskritised version of the text.³⁵

It is sometimes thought that because a more Sanskritised version of a text is in better, i.e. more correct, Sanskrit, it must also give a better, i.e. more correct, interpretation of the underlying Middle Indo-Aryan text, and we should therefore interpret a Middle Indo-Aryan text in the light of a later Sanskritised version. This, however, is not necessarily a good principle to follow. Since all the versions we have of Hīnayāna texts are translations of some earlier version, the correctness or otherwise of a Sanskrit version depends upon the aids which the Sanskrit translator had available to him when he was making his translation. Where there was doubt about the way to interpret, and hence to translate, a word or passage, a Sanskrit translator who had better guides to interpretation than someone translating into a Middle Indo-Aryan dialect, was likely to produce a Sanskrit version which was more correct than the Middle Indo-Aryan version. If the Middle Indo-Aryan redactor had better guides, then it is likely that the Middle Indo-Aryan version would be more correct. If neither of them had any help, then their translations would be based upon their ability to guess the meaning, and neither translator was likely to be consistently better than the other at guessing.

Reviewers and others sometimes point out that in my translations or studies of Middle Indo-Aryan words and works I have not taken account of the Chinese or Tibetan or Khotanese, or whatever other, version they think is important. This is perfectly true, for a number of reasons: (1) I am deliberately confining myself to translating or dealing with the Middle Indo-Aryan material; (2) I was probably ignorant of the Chinese, Tibetan or Khotanese version's existence; (3) I could not handle it even if I knew about it; (4) such versions are translations, usually from Sanskrit, which is itself a translation, from some Middle Indo-Aryan version. It may therefore be interesting, as an interpretation which one tradition has put upon the material it has inherited, but it is in no way an authority, to be followed slavishly. It very often happens that such versions are incorrect, because of errors or misunderstandings in the tradition.

A study of the metre of the Buddhist Hybrid Sanskrit version of Upāli's verses has recently been made.³⁶ A Pāli version of these verses occurs in the Majjhima-nikāya, and portions of a Buddhist Hybrid Sanskrit version of the verses were

³³ Nakatani, 1987.

³⁴ Bernhard, 1965.

³⁵ See Bechert, 1991A and de Jong, 1974, 49–82 (53).

³⁶ See Norman, 1993, 113–23, where the relevant information is given.

discovered in Chinese Turkestan and published in 1916. Since then other fragments have been identified and published. The verses are of great interest, for a number of reasons. They give a list of 100 epithets of the Buddha in a metre which occurs only very rarely in Pāli—the old *āryā* or old *gīti* metre. An examination of the structure of the verses of the Buddhist Hybrid Sanskrit Upāli-sūtra shows clearly that the transforming into Sanskrit was made from a Middle Indo-Aryan exemplar in a way which was almost entirely mechanical. We can see that of the features of Sanskritisation which I mentioned earlier, the redactor restored consonant groups, replaced non-historic case endings with more correct endings, and made some changes of vocabulary. He did all this without any regard whatsoever for the metre, and although he made some changes to the sandhi, the frequent examples of hiatus have been reproduced almost without exception. It is obvious that the verses of the Buddhist Hybrid Sanskrit version were originally composed in a Middle Indo-Aryan dialect, but the differences in word order, and indeed in the order of the verses, show that it was a tradition which differed somewhat from the Pāli.

There is also a Chinese version of these verses, which seems not to be based on the Pāli version or any of the Sanskrit versions which we possess. Some of the epithets in the list differ from those we find in the other versions, although it is not clear whether the differences are due to variations in the tradition or to the Chinese redactor's inability to understand his exemplar. It is very probable that the Chinese translation was made from a Buddhist Hybrid Sanskrit version, which in turn was made from a Middle Indo-Aryan version, probably in a dialect of the North-western Gāndhārī variety, which must in turn have come from another Middle Indo-Aryan dialect, which may have been the dialect used by the Buddha, but was probably not. And yet, despite the long history of translations in the tradition, each of which would have given opportunity for errors to arise, people are very surprised when I say that the Chinese, Tibetan or Khotanese, or whatever version it was, might not have been all that useful to me, if I had been able to read it. I must make it clear that I am talking about Hīnayāna texts whose origin may be presumed to have been in India.

Sanskritisations from the Gāndhārī dialect are always likely to be of doubtful value, because of the wide discrepancy between orthography and pronunciation in that dialect. It is clear that sometimes, at least, the orthography was carried over into translations, irrespective of whether it made sense or nonsense. There is a story found in the Chinese version of the Vinaya of the Mūlasarvāstivādins, where Ānanda heard a monk reciting a Dharmapada-verse which ended with the words, “It were better that a man live only for one day, and see a water-heron”. Ānanda's efforts to persuade the monk that the verse should have ended with the words “and see the principle of coming into existence and passing away”, were unsuccessful. This Chinese version was obviously following a tradition based

upon a Sanskrit form **udaka-baka*, which could only come from a Gāndhārī-type dialect³⁷ which inserted a non-historic *-k-* in place of a glide *-y-*, in the compound *udaya-vyaya*³⁸ “arising and passing away”.

A recent study of the language of the earliest Chinese version of the *Saddhāmapuṇḍarīka*³⁹ suggests that a number of the idiosyncrasies of the language of that text owe something to the fact that the text seems to have been based not upon the Sanskrit versions which we know now, but upon some other version in a dialect which was very similar to the Gāndhārī dialect. The frequent confusion between *jñāna* “knowledge” and *yāna* “vehicle” found in the text seems to indicate that there must have been a variant *jāna* in place of *jñāna* in the donor dialect, and a development of this with *-y-* in intervocalic positions in compounds would have produced forms such as *Buddha-yāna* in place of *Buddha-jñāna*.⁴⁰ Similarly, the fact that *bhoti*, the contracted form of *bhavati* “he becomes”, appears to have been Sanskritised as *bodhi* “awakening” implies a translation from a dialect which, like Gāndhārī, articulated the aspirate so weakly that it could be written or omitted at will.

Among the features which are very common in Middle Indo-Aryan metrical texts is the way in which verses, although hypermetric, can be made to scan by assuming the very common phenomenon of the resolution of syllables, under the musical influence which allows, in certain positions in many metres, the replacement of a long syllable by two short syllables. When *pādas* showing resolution are Sanskritised without any further change, and Sanskrit forms, etc., are restored, the redactor cannot compensate for such resolution, and a verse results which has to be categorised by a modern editor as hypermetric,⁴¹ or irregular.⁴² We can sometimes see that the first stage of Sanskritisation produced verses which barely approximated to Sanskrit in form and metre, and a second stage of Sanskritisation was required to produce something which was more correct, by changes of vocabulary or word order.

Another common feature in Middle Indo-Aryan metrical texts is the way in which forms are evolved *metri causa*, by the lengthening or, less commonly, shortening of syllables, e.g. we find *nirūpadhi* instead of *nirupadhi* “without sub-strate”, or *anodaka* as a replacement for *anudaka* “without water” in the cadence of a *śloka* verse, where the metre requires the short-long-short-long pattern of syllables. When such metrical adjustments are Sanskritised, the earlier

³⁷ cf. *udaka-vaya*, GDhp 317.

³⁸ DhP 113.

³⁹ Karashima, 1992.

⁴⁰ Karashima, 1992, 266 and 1991, 607–43.

⁴¹ See Matsumura, 1989, 80 (ad [26]).

⁴² See Matsumura, 1989, 81 (ad [32]); 82 (ad [35]).

Sanskritised Prakrit version writes forms which are still not correct Sanskrit, e.g. *niropadhi* and *anodaka*. In a more correct Sanskritised version we find that a different way of dealing with the problem is evolved, often involving the use of *vrddhi* vowels, e.g. *niraupadhi*.⁴³

What effect did all these various aspects of Sanskritisation have on Buddhist texts? It is obvious that, by definition, the Sanskritisation of Buddhist texts made them more like Sanskrit, which meant that ambiguous Middle Indo-Aryan phonology and morphology was replaced by precise Sanskrit forms, ambiguities of sandhi and grammar were removed, and the metre was repaired, where it was recognised. This also had a negative side. Where a Middle Indo-Aryan form was ambiguous, the Sanskrit redactor may have chosen the wrong alternative. Where the ambiguity was intentional, perhaps to make a deliberate play upon words, e.g. when giving an etymology based upon a homonym, the Sanskrit redactor could only translate the homonyms into Sanskrit, where they were not homonyms, thus losing the point of the pun. If the redactor did not understand his exemplar, and was forced to translate mechanically or by guesswork, then the results were unpredictable, but were almost certainly wrong.

In the third lecture I spoke of the Waxing Syllable Principle, the way in which words in stock phrases are often arranged in order according to the number of syllables in each word, and I considered the help which this principle might be considered to have given in the memorisation and recitation of texts. The Waxing Syllable Principle was sometimes violated in Sanskritisation, because Sanskritisation changed the syllable count by removing *svarabhakti* vowels, or by changing the vocabulary, i.e. by replacing Middle Indo-Aryan words by Sanskrit synonyms.⁴⁴ The Waxing Syllable Principle was a feature of oral literature, intended as an aid to memorisation and recitation. Since Sanskrit and writing went hand in hand, the fact that Sanskritisation made memorisation more difficult was perhaps of less importance than it might have been at an earlier time.

What effect did Sanskritisation have upon Buddhism? I mentioned in the fourth lecture the suggestion made by some scholars that in the often-quoted passage from the Vinaya, where bhikkhus ask for permission to translate the Buddha's teachings, the word *chandaso* means "into Sanskrit". If that is correct, we must note that the Buddha forbade translation in this way. Centuries later, however, that instruction (if it was indeed his instruction) was ignored, and Buddhists works were indeed translated into Sanskrit. Whether those who made the translation were aware that they were ignoring the Buddha's command, is an interesting question.

⁴³ Udāna-v 6.10.

⁴⁴ See von Hinüber, 1993, 101–13 (109).

The Sanskritisation of Buddhist texts meant that their readership or intelligibility was not limited to monks who were acquainted with the dialect in which the texts had hitherto been transmitted, and it must, therefore, have made them more accessible to a wider range of readers. If a monk of any individual school could read and understand the (more or less) Sanskritised version of the teachings of his own school, he would be able to read the teachings of other texts too, if in his travels he came to a *vihāra* belonging to another school, or met a monk of another persuasion. The effect which such increased communication between sects and schools had is a matter for speculation. It has been shown, for example, that the Sarvāstivādins and Mūlasarvāstivādins possessed a number of different versions of the same text, e.g. the Udānavarga, and also remained in close contact, so that a Sarvāstivādin version of a text might be influenced by a Mūlasarvāstivādin version of the same text.⁴⁵ Such influence could more easily be effected if both versions were in Sanskrit, than if they were in two different dialects of Middle Indo-Aryan.

This takes us back to the reason for Buddhism making use of the vernacular languages in the first place. This must have been two-fold: as a rebellion against the brahmanical caste and their language, and to make the Buddha's teaching available to the common people. The brahmins were the communicating channel between men and gods. They alone knew the rituals and the sacrificial methods to intercede with the gods on behalf of their patrons. The Buddha, by denying the power of the gods, removed the need for a priestly caste to intercede, and taught a system whereby every man could gain his own salvation, either immediately or in a future existence. His teaching was available in the language of the people. As the teaching became codified and its language became ossified, and as the vernacular languages continued to develop, a gap arose between the language of the teachings and the language of the people. Already by the second century B.C.E. we can see that there was a considerable difference between Pāli and Sinhalese Prakrit, as it appears in inscriptions, and a century or two later the difference was so great that, apart from Pāli loan words in Sinhalese Prakrit, one would be hard pressed to see that the two Middle Indo-Aryandialects were related. If, therefore, the *Buddhavacana* was to be propagated in a form which the common people could understand, then bhikkhus would have to translate the texts which they had memorised and recited in Pāli into the vernacular languages of the people. Ultimately, as we know, the Pāli canon had to be translated into Sinhalese (Prakrit). Something similar must have been the experience of all sects and schools of Buddhism.

Although the language of the Gāndhārī Dharmapada was archaic, and approximated to the spoken language of two centuries before, nevertheless we

⁴⁵ See Schmithausen, 1987, 304–81.

may assume that it was still intelligible, for the most part, to the people living in the Gandhāra area and along the southern Silk Road in Chinese Turkestan where it was found. Texts written in this language, or something very similar, were, however, then translated into a Prakritised Sanskrit and then into Buddhist Sanskrit by the monks living in the Turfan area on the northern Silk Road. We may wonder whether anyone, other than those monks, could read the scriptures in that form.

If the teachings were no longer intelligible to the masses, then it did not matter much which language they were in, and it is clear that Sanskrit, the language of culture and literature, was an obvious choice. It was the literary language of North India, at least, and translation into that language meant that educated people—mainly monks, but probably some laymen as well—could read their own scriptures and also those of any other school, if they were so minded. The result was, however, not hard to predict. Buddhism started to become an academic study, where only the educated, who had learned Sanskrit, had access to the teachings in their written form. Bhikkhus, studying in their *vihāras*, became more and more remote from the people. We know of only one Buddhist work devoted to the life and duties of a layman, the *Upāsakajanālaṅkāra*. When Buddhism came under attack from the iconoclastic Muslim invaders, and the *vihāras* and their libraries were destroyed, Buddhism in the land of its founder virtually disappeared.

There is a very obvious contrast with Jainism. There the monks and the laymen were more closely integrated. Texts dealing with the duties of laymen are numerous. Many translations of canonical texts were made, but into vernacular languages, not Sanskrit, and there were also commentaries written in vernacular languages. Although the Jains did use a form of Sanskrit, which resembled the Sanskrit used by Buddhists in its dependence upon Middle Indo-Aryan forms, Sanskrit never gained the importance for Jains which it had for Buddhists. Although both Buddhism and Jainism had royal patrons, Jainism also had close links with, and was less remote from, the people, and we find, for example, that the Jains made great use of popular literature, with Jain versions of the *Mahābhārata* and the *Rāmāyaṇa*. It has been argued that it was as a consequence of its support among the common people that Jainism, although hit no less than Buddhism by the invaders, survived in North India. Buddhism did not.