e-ISSN 2785-5953

Bhasha
Vol. 3 — Num. 2 — October 2024

Buddhist Hybrid Sanskrit: How Did It Originate?

Hans Henrich Hock
University of Illinois at Urbana-Champaign, USA

Abstract Since Edgerton 1953, Buddhist Hybrid Sanskrit (BHS) has been considered a sanskritization of Prakrit, based on strong linguistic evidence: lexical items are Sanskritic, endings Prakritic. Sanskritization has been argued as motivated by a need to compete with Sanskrit-using brahmins. The issue of how sanskritization might have been accomplished is the topic of this paper. In early AD, Sanskrit was spread by brahmins as language of technical and fine literature. The curriculum of schools imparting Sanskrit instruction started with memorization of a Sanskrit lexicon and a version of Pāṇini's grammar. The link between these was established in later years. Sanskritization of BHS can be explained in terms of early Buddhist students only completing the initial stage of instruction. This would provide them with a Sanskrit lexicon for replacing Prakrit words. However, not yet knowing how to apply the grammatical rules, students would use Prakritic endings. Support for this hypothesis comes from Kapstein's (2018) account of grammatically deficient, but lexically accurate Sanskrit compositions by medieval Tibetans, as resulting from acquiring grammar and lexicon separately, 'with almost no training in practical application'. I conclude by considering the implications of my proposal as well as the similarities and differences between BHS and 'Bilingual Mixed Languages'.

Keywords Buddhist Hybrid Sanskrit. Historical development. Sanskritization. Tibetan parallel. Bilingual Mixed Languages.

Edizioni Ca'Foscari

Peer review

Submitted 2024-06-06
Accepted 2024-11-08
Published 2024-12-18

Open access

Citation Hock, H.H. (2024). "Buddhist Hybrid Sanskrit: How Did It Originate?". *Bhasha*, 3(2), 175-192.

DOI 10.30687/bhasha/2785-5953/2024/02/001

1 Introduction

Since Edgerton (1953) Buddhist Hybrid Sanskrit (BHS), the language attributed to the Lokottaravāda Mahāsāṅghika school of Buddhism, has been considered a sanskritization of an original Prakrit, where 'Prakrit' is understood as a variety of early Middle Indo-Aryan.[1] A likely motivation for this sanskritization has been proposed in earlier literature, namely the desire to compete with Sanskrit-using brahmins in religious disputations at royal courts (e.g. Salomon 1998; Bronkhorst 2010). The manner in which such a massive lexical sanskritization could have been accomplished, however, has not been satisfactorily addressed. This paper attempts to answer that question.

Section 2 presents a brief overview of the characteristics of BHS, followed by a discussion of the similarities and differences between BHS and Epigraphic Hybrid Sanskrit (EHS) in Section 3. Chronological issues regarding BHS are dealt with in Section 4, together with the evidence of EHS. Motivations for the development of BHS are the topic of Section 5. Section 6 discusses the likely mechanism for the sanskritization of BHS. Section 7 examines two texts that show mixtures of Prakrit or Apabhraṁśa and Sanskrit similar to BHS but are sufficiently different to merit separate discussion. Finally, Section 8 presents a summary of my findings, as well as a comparison to 'Bilingual Mixed Languages'.

2 Characteristics of BHS

Buddhist Hybrid Sanskrit is characterized by the following features. The stems of nouns and adjectives almost always are Sanskrit in form. Verbs and function words, by contrast, tend to show various degrees of Prakrit features, and so do nominal and adjectival case endings. In addition, external sandhi generally follows Prakrit rules.

Detailed coverage of the features of Buddhist Hybrid Sanskrit can be found in Edgerton 1953. For present purposes, suffice it to consider examples (1) and (2) from the early text *Mahāvastu* (Marciniak 2019 edition), as illustrations of how Buddhist Hybrid Sanskrit differs from standard Sanskrit.[2] Sandhi (small cap italics), inflectional

This paper is a thoroughly revised and expanded version of a presentation at the 2023 SALA meeting in Venice. I have benefited from comments at that meeting as well as feedback from two anonymous referees. The responsibility for any errors and omissions rests with me.

1 The term, thus, does not (normally) cover the literary Prakrit discussed in detail by Ollett 2017.

2 Edgerton's grammar does not present any textual examples.

endings (bold), and entire forms (italics), especially verbs, are of Prakrit origin, while the majority of lexical items are Sanskrit (unmarked). Sanskrit equivalents are given for comparison below each line; a third line offers a virtual Pāli version that illustrates the difference of BHU from both standard Sanskrit and Prakrit.

(1)	*abhūṣi*	rājā	ikṣvāku	vārāṇasyāṁ	mahābalo
	abhūd/āsīd	rājā	ikṣvākur	vārāṇasyāṁ	mahābalaḥ
	a(b)hosi	rājā	okkāko	bārāṇasiyā	mahābalo
	be.AOR.3SG	king.NOM.SG.M	Ikṣ.NOM.SG.M	Vār.LOC.SG.F	great.strength.NOM.SG.M

'There was a strong king, Ikṣvāku, in Vārāṇasī.'

(2)	so	ca	jīrṇo	*bhavi***tvāna**
	sa	ca	jīrṇo	bhūtvā
	so	ca	jiṇṇo	bhavitvā(na)/bhavittā*
	dem.nom.sg.m	&	old.NOM.SG.M	be(come).CVB
	vepamāne**hi**	gātr**ehi**	rājādvāram	upāgato
	vepamānair	gātrai	rājadvāram	upāgataḥ
	vepamānehi	gattehi	rājabbāraṁ**	upāgato
	trembling.INST.PL.N	limb.INST.PL.N	king.gate.ACC.SG.N	go.up.PFV.PPL.NOM.SG.M

'And he (Indra) turning himself old, […] came to the royal gate with trembling limbs.'

* The latter form is non-canonical.

** This is the expected Pāli form and the general Prakrit version; the usual Pāli form, *-dvāra*, may reflect sanskritization (von Hinüber 1986, 123).

As can be seen, nominal and adjectival roots and stems are Sanskrit in form. Contrast, e.g., the name *ikṣvāku-* with Pkt. *okkāka*, *jīrṇa-* 'old' with *jiṇṇa*, *gātra-* 'limb' with *gatta-*.

However, only a few inflected forms are identical to their Sanskrit counterparts – *rājā* 'king.NOM.SG', *vārāṇasyāṁ* 'in Vārāṇasī', and *jīrṇo* 'old.NOM.SG'. Nominal stems are generally identical to their Sanskrit counterparts, such as *ikṣvāku-* in (1) or *vepamāna-* in (2), an exception being *rājādvāra-* in (2), rather than *rājadvāra-*, with long *ā* vs. Skt. short *a* in the second syllable. Nominal endings, however, are Prakrit, not only those violating Sanskrit sandhi such as *ikṣvāku* rather than *ikṣvākur* and *mahābalo* instead of *mahābalaḥ* but also forms where Middle Indo-Aryan made suffix choices different from contemporary Sanskrit, such as *vepamān**ehi*** rather than *vemapānair/ḥ*.[3]

Especially important are the verbal forms *abhūṣi* 'was' and *bhavitvāna* 'having been/become'. The latter differs from the corresponding Sanskrit form both in the form of the root (*bhavi-* instead

3 Vedic Sanskrit offered an option between *-ebhiḥ* and *-aiḥ* and their sandhi variants, but that was lost in favor of *-aiḥ* in Classical Sanskrit.

of *bhū-*) and in its suffix (*-tvāna* vs. *-tvā*). The former is, in fact, quite complex with an unusual mixture of Prakrit and Sanskrit features. Its retroflex sibilant *ṣ* follows Sanskrit phonology (Prakrit would have dental *s*), and its root shape *bhū-* looks like Skt. *bhū-* in the aorist form *abhūd*,[4] but its overall formation as an *s*-aorist is Prakrit; compare Pali *a(b)hosi* vs. Skt. *abhūd*.[5] A virtual, but unattested Sanskrit counterpart of *abhūṣi* would be *abhūṣīd**, but the final short *-i* of the BHS form is Prakritic. Moreover, the choice of the aorist formation is typical of Prakrit, which merges imperfect and aorist, with heavy dominance of aorist forms; Sanskrit would instead have the imperfect *abhavad* or, preferred in this context, *āsīd* (from a different root).

In the larger text, however, there are instances of sanskritized verb forms, such as *tiṣṭhati* 'stands' vs. Pkt. *tiṭṭhati*. These are mainly found in forms that are relatively similar to Sanskrit in their morphology; forms like *abhūṣi*, whose morphology differs considerably from Sanskrit, remain Prakritic.[6]

Other features, not found in the sample texts of (1) and (2) include those listed in (3) below. (3a) illustrates the common Prakritic extension of the *a*-stem genitive singular ending *-sya,* sanskritization of Prakrit *-ssa,* to the *i*-stems and *in*-stems on the model in (3c).[7] The retroflex sibilant in the rarer variant *riṣiṣya* is a case of hypersanskritization. (3b) provides examples of function words of Prakritic origin or shape; note that BHS *sace* (sometimes partly sanskritized as *saced*) and Skt. *ced* differ in their syntax: *sace* is clause-initial, whereas *ced* is a second-position clitic.

(3)	a.	BHS	Sanskrit		
		vārisya	vāriṇaḥ	'water.GEN.SG.N'	
		riṣisya/riṣiṣya	ṛṣeḥ	'seer.GEN.SG.M'	
	b.	kāci	kāś-cid	'some.NOM.PL.F'	
		sace	ced	'if'	
	c.	Prakrit	*pakkha*:	pakkha-ssa	'wing, side'
			vāri:	X = *vāri-ssa*	'water'

4 Sanskrit forms are cited here in the sandhi version appropriate to the context in (1).

5 However, the root form differs (*ū* vs. *o*), and as Edgerton (1953, 157) notes, there is no exact Prakrit counterpart.

6 Note also, elsewhere in the text, forms such as *āsi* 'be.AOR.3PL' vs. Skt. *āsan* 'be.IMPF.3PL'.

7 On this ending, see also Edgerton 1953, 74.

Bhasha
3, 2, 2024, 175-192
e-ISSN 2785-5953

3 BHS, Epigraphic Hybrid Sanskrit (EHS), and the Issue of Buddhist Sanskrit

Although details differ, a similar mixture of Sanskrit and Prakrit features is found in the language of inscriptions dated (roughly) from the first century BC to the first century AD, a language which has been called Epigraphic Hybrid Sanskrit (EHS) by Damsteegt (1978) and Salomon (1998, 81). For an example see (4). Here again virtual versions in Sanskrit and Prakrit are added for comparison.

(4)	bhaṭārakā *añātiyā*	ca	gatosmi	*varṣāratuṁ* [= varṣā + ratuṁ]
	bhaṭāraka+ājñātyā*,**	ca	gatosmi	varṣartuṁ [= varṣā + ṛtuṁ]
	bhaṭāraka añātiyā	ca	gato mhi	vassāratuṁ
	lord+order.INS.SG.F	&	gone+be.PRS.1SG	rain.season.ACC.SG.M
	mālay**ehi**	ru(d)dha**ṁ**	ut(t)amabhadraṁ	mocayituṁ
	mālayai	ruddham	uttamabhadraṁ	mocayituṁ
	mālayehi	ruddhaṁ	uttamabhaddaṁ	mocayituṁ
	Mālaya.INS.PL.M	besieged.ACC.SG.M	Uttamabhadra.ACC.SG.M	free.CAUS.INF
	tatosmi	gato	pokṣarā**n**i	tatra ca mayā
	tatosmi	gataḥ	puṣkarāṇi	tatra ca mayā
	tato mhi	gato	pokkharāni	tatta ca mayā
	then+be.PRS.1SG	gone.NOM.SG.M	Puṣkara.ACC.PL.N	there & I.INS.SG
	abhiseko	kṛto	tīṇi	gosahasrā**n**i
	abhiṣekaḥ	kṛtas	trīṇi	gosahasrāṇi
	abhiseko	kito	tīni	gosahassāni
	bath.NOM.SG.M	made.NOM.SG.M	3.NOM.PL.N	cow.thousand.NOM.PL.N
	da(t)tāni			
	dattāni			
	dattāni			
	given.NOM.PL.N			

(Nasik Cave Inscription No. II; Bühler 1881, 99-100)

"And by order of the supreme lord I went in the rainy season to liberate the Uttamabhadra chief who was besieged by the Mālayas […] Afterwards I went to the Puṣkaras (Pokṣaras), and I bathed there, and gave three thousand cows". (Bühler's translation)

* This is a virtual Sanskrit form; a more appropriate form would be *ājñayā*.

** The sign + indicates sandhi applying or failing to apply across linguistic forms.

Here again, many nominal and adjectival stems are Sanskrit in form, such as *varṣā* 'rain, rainy season' vs. Pkt. *vassā* or the name *u(t)tamabhadra-* vs. *uttamabhadda-*. The place name *pokṣarāṇi* for Pkt.

pokkharāni is more complex: its *o* is Prakritic, and its *kṣ* is hyper-Sanskrit, instead of the proper Sanskrit *ṣk*.[8]

Many other forms are lexically Prakritic, such as the forms *abhiseko* and *tīṇi*; the *ti*-stem form *añātiyā* which is both lexically and morphologically Prakritic (the correct Sanskrit form would be *ājñayā* from the *ā*-stem *ājñā*); and the ending of *mālayehi* which is likewise Prakritic (vs. Skt. *-ai(ḥ)*). Beyond that, there are many cases of Prakritic sandhi or lack of sandhi, but also a few Sanskritic sandhi forms, such as *tatosmi* (/tataḥ+asmi/ 'then I.NOM.SG.M'). Elsewhere in this inscription, there is an example of hyper-Sanskrit sandhi (*devatābhy**aḥ*** *brāhmaṇebhyaś ca* 'for the deities and the brahmins', with *-aḥ* before voiced *b-* where Sanskrit would have *-o)*. Relevant verbal forms are rare in EHS, but note the verbal-noun form *bhoj**āpayi**trā* 'causer.to.eat.INS.SG.M' (vs. Skt. *bhojayitrā*), with Prakritic 'double causative' marking,[9] found elsewhere in the text. Other inscriptions offer examples like *sahisya* (vs. Skt. *saheḥ*), with the same Prakritic development as in (3a).[10]

Damsteegt also documents that there were at least three different historical stages of EHS, with sanskritization becoming increasingly stronger in later stages, but with some Prakritic forms, especially the instrumental-plural suffix *-ehi*, stubbornly persisting throughout the entire period. Sanskritization, thus, was a continuing process.

While Damsteegt and Salomon see in EHS a stage of sanskritization, similar to that in BHS, Sen argues that both types of language use reflect a 'Spoken Sanskrit', an 'unstable literary or business language varying according to time and place', and he refers to Buddhist Hybrid Sanskrit as simply 'Buddhist Sanskrit'.

In the case of EHS, Sen's characterization ignores the fact that EHS is a transitional stage in the inscriptional record between earlier Prakrit and later Sanskrit, with sanskritization proceeding through several chronological phases. Salomon (1998) therefore is probably correct in considering BHS, too, to reflect a transitional phase of sanskritization.[11] Moreover, Salomon (1983) demonstrates

8 *kṣ* and *ṣk* both resulted in Ptk. *kkh*, but the correspondence *kṣ*: *kkh* is more common.

9 On this formation see Edgerton 1946.

10 For a detailed study of EHS morphology see Damsteegt 1977.

11 For a discussion of alternative theories see Salomon 1998, 81-6. Yet a different perspective is that of Ollett (2017, 44) who seems to look at the replacement of Prakrit forms like *khattapa* by Sanskritic *kṣatrapa* as involving phonological replacement of clusters like *kh* and *tt* by *kṣ* and *tr,* possibly under the influence of the extreme northwestern Gandhari Prakrit which retained such clusters. However, this proposal fails to explain a large number of other BHS (and EHS) phenomena such as the *au* of *k**au**tūhala* 'curiousness' (vs. Pkt. *k**o**tūhala*); for all the other Middle Indo-Aryan varieties, including Gandhari, changed Old Indo-Aryan *au* to *o*. Phenomena like these can only be explained in terms of lexical transfer from Sanskrit.

that Buddhist Hybrid Sanskrit is only one form of Buddhist Sanskrit – the more or less standard Sanskrit of Aśvaghoṣa clearly differs from the Hybrid Sanskrit of the Mahāsāṅghika school of Buddhism (even though it contains a few partly sanskritized Prakritic forms such as *saced* 'if', probably reflecting technical language of Buddhist argumentation).

4 Chronological Problems and the Testimony of EHS

As noted e.g. by Edgerton (1953, 4), BHS underwent increasing sanskritization 'from the very beginning of its tradition as we know it (that is, according to the mss. we have)'.

Especially instructive is the *Mahāvastu* of which we now have two editions, based on chronologically different manuscripts (Senart 1897 and Marciniak 2019). The edition by Senart, based on later manuscripts, shows a higher incidence of sanskritization than Marciniak's edition, which is based on recently found earlier medieval manuscripts; see e.g. the examples in (5). Example (5a) shows a Prakritic neuter form of the demonstrative *eta-* in Marciniak vs. a Sanskritic form in Senart; (5b) and (5c) exhibit Prakritic loss of final stop vs. Sanskritic presence of the stop (but the Prakritic *kā-* instead of *kāś* is retained even in Senart); and (5d) exemplifies Prakritic final *-o*, while Senart applies Sanskrit sandhi appropriate to the phonological environment. Evidently, later scribes introduced forms that were more in conformity with the Sanskrit norms that they were familiar with. This might also explain the fact that verb forms that are relatively similar to their Sanskrit counterparts as in (6a) are commonly fully sanskritized, whereas forms that are markedly different are not, as in (6b). In some cases, metrical concerns may also have prevented sanskritization.[12] For instance *bhavitvāna*, containing four syllables, cannot be replaced by disyllabic *bhūtvā*.

(5)		Marciniak	Senart	Sanskrit	
	a.	etaṁ	etad		'this.ACC.SG.N'
	b.	yāva	yāvat		'as long as'
	c.	kāci	kācit	kāścit	'some.NOM.SG.F'
	d.	upāgato	upāgataḥ	upāgataḥ	'(having) come up'
		tato	tataḥ		'then, after that'

12 Much of the *Mahāvastu* is in verse.

(6)		Prakritic	Sanskritized	Sanskrit	
	a.	bhāṣati	bhāṣate		‘speaks’
		vardhati	vardhate		‘grows’
	b.	abhūṣi		āsīt/abhavat/abhūt	‘was’
		bhavitvāna		bhūtvā	‘having been’
		bhavāhi		bhava	‘be IMPV’

This evidence raises the question of whether the original BHS of the early centuries AD might have been even less sanskritized – a question that for lack of attestations is difficult to answer. However, the late BC/early AD evidence of sanskritization in EHS is ‘set in stone’, not subject to the predilection of later scribes, and early BHS texts like the *Mahāvastu* are remarkably similar (making allowance for differences in genre) to the later phases of EHS. This evidence supports the view that even in its earliest, first-century AD stages, BHS would have undergone a significant degree of sanskritization.

5 Motivation and Institutionalization

In the early centuries AD, Sanskrit spread as the language of statecraft and of technical and fine literature throughout South Asia. Bronkhorst (2010) plausibly argues that this spread was propelled by Sanskrit-using brahmins who were experts in these matters. My own work (Hock 2019) suggests that an important vehicle for imparting the knowledge of Sanskrit consisted in brahmin-dominated, Sanskrit-medium schools.

As regards the use of Sanskrit by Buddhists, whose original texts were in different forms of Prakrit, it has been suggested that a switch toward Sanskrit was motivated by a desire to acquire competence in Sanskrit in order to compete with Sanskrit-speaking brahmins and to defend the Buddhist faith against brahmins in disputations at royal courts; see e.g. Salomon 1998, Bronkhorst 2010. The hybrid language of one school of Buddhism, then, might be considered to be an intermediate stage in this sanskritization process, comparable to EHS.

For reasons that are not recoverable, texts reflecting this intermediate stage became institutionalized and their Hybrid Sanskrit was adopted as the sacred language of one school of Buddhism, just as Pali was the sacred language of Theravada Buddhism, and Sanskrit was the sacred language of Brahmanism and became the language of most forms of Buddhism. In fact, Bronkhorst (1993) presents evidence that just as brahmins rationalized the difference between the Vedic language of their sacred texts and the Sanskrit of their current usage by declaring the two to actually be the same language, so members of the Mahāsāṁghika branch of Buddhism claimed that BHS and Sanskrit actually are the same language; and just as the

brahmins were able to account for Vedic pecularities in terms of rules in Pāṇini's grammar, Mahāsāṁghikas employed some of the same rules to justify peculiarities of BHS.

6 Mechanism

While sanskritization must clearly be recognized as a historical process, the term 'sanskritization' itself is merely descriptive, its mechanism is left unaccounted for.

We can safely rule out the idea that Buddhists employed brahmins competent in Sanskrit to sanskritize their texts. If they had done so, the brahmins would surely have produced grammatically proper Sanskrit texts. A more likely account for the mechanism underlying sanskritization would proceed along the following lines.

As noted earlier, an important vehicle for instruction in Sanskrit consisted in brahmin-dominated, Sanskrit-medium schools. Now, the curriculum of such schools started with the memorization of a thesaurus of Sanskrit words and of a simplified version of Pāṇini's grammar. A link between these memorized texts was established in later years through composition and the study of texts. The existence of such a form of instruction into the early twentieth century was confirmed by me in a 1980-81 research project on spoken Sanskrit, funded by the American Institute of Indian Studies.

Against this background, the mainly lexical sanskritization of BHS can be explained as the result of some of early Buddhist students only completing an elementary level of Sanskrit instruction.[13] This would leave them with a ready-made Sanskrit lexicon for replacing Prakrit words – hence the massive lexical sanskritization. As regards grammar, however, they would not yet have a full grasp of how to apply the memorized rules in practice – hence they would tend to use Prakrit sandhi and inflectional forms. Occasional hypersanskritizations further attest to the fact that Sanskrit grammar was only incompletely grasped.[14]

13 A reviewer objects that '[s]urely *some* of the Buddhist students would have gone beyond the first year and mastered 'proper' Sanskrit. It is hard to see how the weak attempts by beginners would have been codified into a literary vehicle.' Now, it is in fact true that 'some Budhhists' went beyond an elementary study of Sanskrit; but they produced a different literary vehicle – that of the bulk of Indian Buddhist tradition and also of poets and dramatists such as Aśvaghoṣa, whose language (as noted in Section 3) is more or less 'proper Sanskrit'. The problem is that Buddhist Hybrid Sanskrit cannot possibly accounted for as resulting from more than elementary exposure to Sanskrit grammar.

14 The evidence of EHS suggests a limited increase in familiarity with aspects of Sanskrit grammar during the late centuries BC and the early centuries AD.

The result, then, evidently became institutionalized as the language of the Mahāsāṁghika branch of Buddhism. Other branches, except for the Theravādins who used Pali, followed the general trend to full acquisition of Sanskrit.

Unexpected support for the likelihood of the present hypothesis comes from Kapstein's account of grammatically deficient, but lexically accurate Sanskrit compositions by Tibetans during the Middle Ages:

> Errors such as this [...] were likely due in part to the practice of teaching vyākaraṇa [grammar] and abhidhāna [practice] quite separately, [...] with almost no training in practical application. (2018, 470)

7 Some isolated Mixed Sanskrit varieties similar to BHS and EHS

Beside BHS and EHS there are several other forms of language use that exhibit a mixture of Prakrit (or Apabhraṁśa) features and Sanskrit, but these are limited to individual texts, not parts of broader attested textual traditions; and they exhibit features sufficiently different from BHS and EHS to merit separate discussion. One of the these is the so-called *Patna Dhammapada* (Cone-Ānandajyoti Bhikku 2017, Ānandajyoti Bhikku 2020), the other the mathematical "Bakhshali Manuscript" (Hoernle 1887, Hayashi 1995).

The *Patna Dhammapada* (PDh), considered to be associated with the Sāmmatīya branch of Buddhism, has been variously characterized as Buddhist Hybrid Sanskrit, Prakritic, Buddhist Prakrit, and Sanskritized Prakrit (Cone-Ānandajoti Bhikkhu 2017, 4), but Karpik (2023, 78) considers it 'an important milestone in the Sanskritisation of Pali'.

There are in fact a number of features shared with BHS, but there are also differences. Consider the sample passage in (7). (Prakrit forms are in italics, Prakrit inflectional endings in bold, Prakrit sandhi in small cap italics; unambiguously Sanskrit forms are underlined; plain roman indicates forms that could be both Prakrit and Sanskrit). Here again a virtual Sanskrit version is added for comparison, as well as a real Prakrit counterpart (from the Pali *Dhammapada*).

(7)	ākroś**i**	maṁ	avadh**i**	maṁ
	ākrośan	mām	avadhiṣur	mām
	akkocchi	maṁ	avadhi	maṁ
	abuse.PST.3PL	I.ACC.SG	hit.PST.3PL	I.ACC.SG
	ajini	maṁ	ahāsi	me
	ajayan	mām	aharan	me
	ajini	maṁ	ahāsi	me

Bhasha
3, 2, 2024, 175-192
e-ISSN 2785-5953

defeat.PST.3PL	I.ACC.SG	take.away.PST.3PL	I.OBL
ye		tānı	upanahyanti
ye		tāny	upanahyanti
ye	ca	taṁ	upanayhanti
REL.PRON.NOM.PL.M		DEM.ACC.PL.N	combine.PRS.3PL
veraṁ	tesaṁ	na	śāmyati ‖ 5 ‖
vairaṁ	teṣāṁ	na	śāmyati
veraṁ	tesaṁ	na	sammati
hatred.NOM.SG.N	DEM.GEN.PL.M	neg	quiet.down.PRS.3SG

'They abused me, they hit me, they defeated me, they robbed me; Those who combine these (actions), their hatred does not cease.'

As in BHS, some forms are sanskritized, such as *ākroś-* (vs. Pali *akkocch-*) and *śāmyati* (vs. *sammati*), and many others could be either Prakrit or Sanskrit, such as *na* NEG. However, many parts of example (7) are purely Prakritic, such as *veraṁ tesaṁ* 'their hatred' (vs. Skt. *vairaṁ teṣāṁ*); and throughout the text the ratio of Prakritic vs. Sanskritic forms is much higher than in BHS.[15]

Morphologically, past-tense forms exhibit the Prakrit generalization of the aorist, where Sanskrit would tend to use the imperfect (or the perfect), as in *ajini* 'they defeated' vs. Skt. *ajayan*; moreover, the endings of these forms are Prakritic, as in *-i* vs. Skt. *-an*. Forms that are relatively similar to Sanskrit, however, tend to be sanskritized; e.g. *upanahyanti* (vs. Pali *upanayhanti*). Elsewhere in the text we find Prakritic function words such as *sace* 'if' (vs. Skt. *ced*). External sandhi follows Prakrit norms, as in *tāni**i*** *upanahyanti* vs. *tān**y*** *upanahyanti*; similarly *man**o**śreṣṭhā* 'having mind as their leader' vs. *mana**ḥ**śreṣṭhā* elsewhere in the text. There are also some instances of hypersanskritisms, such as *ku**ś**īdaṁ* 'idle' (vs. Skt. *ku**s**īdaṁ*, Pali *ku**s**ītaṁ*) and *śīghra**śś**o* 'fast horse' (Skt. *śīghrā**śv**a*, Pali *sīgha**ss**o*).

Following Karpik (2023), these facts could be interpreted as indicating an early stage of the sanskritization found in BHS texts, in which case it might be tempting to attribute the greater degree of sanskritization in BHS to Sanskrit-proficient scribes. However, it is also possible that the *Patna Dhammapada* reflects a different development from the one underlying BHS, with a more minimal amount of sanskritization, in a different branch of Buddhism than the one associated with BHS.

15 Terminology central to Buddhism, such as *dhamma* '(cosmic) order; proper conduct' (Skt. *dharma*), *kamma* 'action' (Skt. *karman*), *bhikkhu* 'monk' (Skt. *bhikṣu*), *nibbāṇa* 'release' (Skt. *nirvāṇa*), remains unsanskritized in virtually all cases. On the other hand, as in Pali, *brāhmaṇa* 'brahmin' always appears in Sanskritic form (vs. Pkt. *bamhana*).

The possibility of different, independent sanskritization receives some support from the Bakhshali Manuscript (BM) whose base is not Prakrit (including Pali) but more likely Apabhraṁśa, a later form of Middle Indo-Aryan.

Like the *Patna Dhammapada*, the Bakshali Manuscript has been variously characterized. Hoernle labels its language as "the literary form of the ancient Northwestern Prakrit", with "a strange mixture of what we should now call Sanskrit and Prakrit forms" (1887, 10), 'an imperfect sanskritisation of the vernacular Prakrit' (1887, 14). Kay (1933, 11) describes it as "an irregular Sanskrit', whose 'peculiarities of spelling, sandhi, grammar […] are exceedingly common in the eleventh and twelfth centuries found in north-west India" (without, however, giving examples of other texts with the same characteristics). Hayashi (1995, 15) notes that the language 'has a number of peculiarities in common with the so-called Buddhist Hybrid Sanskrit', but also has 'strong affinity with Apabhraṃśa […], and, as expected from the find-spot (Bakhshālī near Mardan) of the manuscript, also with Old Kashmiri' (with reference to Grierson 1929); a final part of BM, however, is in an entirely different dialect, which is not discussed below.

An examination of the 'peculiarities' of the text suggests that Hayashi's comparison with Apabhraṁśa, the latest stage of Middle Indo-Aryan and the (near-)ancestor of the Modern Indo-Aryan languages, points in the most likely direction.[16]

As in the other mixed varieties, let us start with a short sample of BM; see (8), with similar formatting as for PDh.; for the forms with dotted underline see the discussion below. Here again, an attempt at a Sanskrit version is added for comparison.[17] The grammatical glosses refer to what may be the correct Sanskrit equivalent.[18]

16 Hayashi's comparison with the Old Kashmiri text described by Grierson is less likely to be correct. Grierson describes the language of that text as similar to learnèd texts in Modern Indo-Aryan languages, with a heavy amount of lexical borrowing from Sanskrit, but with the grammatical endings of Old Kashmiri (1929, 77).

17 I don't feel competent to add a virtual Apabhraṁśa version, especially since this is a technical, mathematical text. Where individual virtual Apabhraṁśa forms are cited, the information they are based on comes from Tagare 1948.

18 Problems with the grammatical endings of the BM sometimes make it difficult to guess what the correct Sanskrit equivalent should be.

(8)	[rā]japutr**o**	dvay**o**	*k**e**ci*	nṛpati**s**sevy**a**	
	rājaputrau	dvau	kaucin	nṛpatisevyau	
	Rajput.NOM.DU.M	2.NOM.DU.M	some.NOM.DU.M	king.master.NOM.DU.M	
	santi	vai**ḥ** |			
	sto	vai			
	be.PRS.3DU	particle			
	*mekā*syāhne	*dvaya**ṣ**ṣa**d**bhāg**ā***	dvitīyasya		
	ekasyāhne	dviṣaḍbhāgo	dvitīyasya		
	1.GEN.SG.M + day.LOC.SG.N	2.[1]/6.part.NOM.SG.M	2[nd].GEN.SG.M		
	*div*arddhika**ṁ**				
	dvyardhikasya				
	1[1]/2.GEN.SG.M				
	prathamena	dvitīyasya	daśa	dīnāra**s**	dattavān |
	prathamena	dvitīyāya	daśa	dīnārā	dattāḥ
	first.INS.SG.M	second.DAT.SG.M	10	dīnāra.NOM.PL.M	given.NOM.PL.M
	kena kālena	samatā**ṁ**	gaṇayitvā		
	kasmin kāle	samatā	gaṇayitvā		
	what time.LOC.SG.M	equality.NOM.SG.M	calculate.CVB		
	vada+āśu	me ||			
	vadāśu	me			
	tell.IMPV.2SG+quickly	I.obl			

'Two Rājputs are the servants of a king. The wages of one per day are two and one-sixth of the other one and one half. The first gives to the second ten *dīnāras*. Calculate and tell me quickly, in what time will there be equality (in their possessions)?' (Hoernle's translation)

As can be seen, almost all lexical items are in Sanskritic shape. Contrast, e.g., the Apabhraṁśa numeral *doṇṇi* or *beṇṇi* '2' with the *dvayo* of our text, Ap. *pahila-* 'first' with *prathama-*, Ap. *pu**tt**a* 'son' with *-pu**tr**o,* or Ap. *v**aahi*** 'say IMPV' with *v**ada***.

However, the case usage, such as *dvitīyasya* (GEN) 'to the second' for *dvitīyāya* (DAT) and *kena kālena* (INS) 'at what time' for *kasmin kāle* (LOC) reflects Prakrit or Apabhraṁśa grammar. Similarly, the use of plural instead of dual forms, as in *keci* (PL) 'some' vs. Skt. *kau-cin* (/kau-cid/ DU) is attributable to Prakrit or Apabhraṁśa influence.[19]

More significantly, several forms and constructions suggest that the composer's (or composers'?) grasp of Sanskrit was limited.

19 Some peculiarities of BM do not seem to be explainable in terms of any other known language use. These include the doubling of word- or stem-initial consonants in (8) in such forms as *nṛpati-**s**sevya* for *nṛpati-sevya-*, as well as the use of *m, r,* or *s* as hiatus breakers. There is also the form *meka-* 'one', which Hayashi (1995, 35) plausibly explains as reanalyzed from *ekameka-* 'one-another'.

Consider the form *dvaya-* 'two'. The correct Sanskrit formation is *dvau* (and inflectional variants) and it is used only in the dual; Skt. *dvaya-* means 'of two kinds, double' and is normally used only in the singular or the plural, not in dual reference.[20]

Further, the verb form in *prathamena* (INS) *daśa dīnāra dattavān* 'the first one […] gave' is ungrammatical in Sanskrit, which would require the *ta*-participle form *dattā(ḥ)* instead, agreeing with *dīnārā(ḥ)*.[21] The past participle in *-tava(n)t* is construed as active, requiring a nominative subject (*prathama(ḥ)* NOM *daśa dīnāra dattavān* 'the first one gave'); the past *ta*-participle is construed as ergative, with instrumental agent marking (*prathamena* INS *daśa dīnāra dattavān* 'by the first one was given' = 'the first one gave'). Apparently the composer did not understand this difference and, focusing on the fact that both the *tava(n)t-* and the *ta*-form indicate past tense, used the *tava(n)t*-participle instead of the *ta*-participle because it better fits the meter (which calls for the scansion ⏑⏑– in this position). Note that Apabhraṁśa only uses the reflex of the *ta*-participle for past-tense reference.[22]

Further evidence that the composer(s) were insufficiently familiar with Sanskrit is the hyper-Sanskritism *vai**ḥ*** for Skt. *vai*.

Hoernle (1887, 15) and Hayashi (1995, 26-55) note that there is a great amount of confusion as regards case, gender, and number. As it turns out, this 'confusion' is characteristic of Apabhraṁśa, where sound change and analogy led to the loss of distinction between masculine and neuter, the attentuation of a formal distinction of feminine gender, and the reduction of the case system (Tagare 1948, 27, 105-6). More specifically, in his listing of morphological peculiarities, Hayashi (1995, 26-55 with general reference to Tagare 1948) gives numerous examples of case endings that are found in Apabhraṁśa, but not in earlier Prakrits, as well as endings that are found in both. These include the masculine/neuter nominative/accusative singular endings ***-a**, -aṁ, -**ā**, -āḥ, -āṁ, -**u**, -o* and the corresponding plural

20 The second occurrence of *dvaya-* seems to be a dittography; something like *d(v)iṣaḍbhāga-* would better fit the meter.

21 A reviewer suggests that *daśa dīnāra dattavān* should be interpreted as singular, because *daśa* is singular; but *dīnāra* would have to be in the plural (*dīnārā(ḥ)*) and the verb should agree with that plural.

22 A reviewer comments that "[m]ixed passive/active constructions are common in OIA and MIA languages; see especially Jamison 2000". To my knowledge, such structures are not at all common in either Old Indo-Aryan or the mainstream of Middle Indo-Aryan. Jamison (2000), to be sure, notes structures that are superficially similar in the extreme northern Niya Prakrit, but the morphosyntax differs considerably from Sanskrit or the mainstream of Middle-Indo Aryan. For instance, forms of 'be' are cliticized to first- and second-person forms of the *ta*-participle, leading to the loss of ergative alignment; third-person antecedents may be marked in the absolute case or the instrumental; instrumental marking is often found in traditional formulas or seems to be employed as a kind of 'differential agent' marker.

endings ***-a***, *-aṁ, -aḥ,* ***-ā***, *-āṁ,* ***-e***, *-o* (where bold indicates exclusively Apabhraṁśa variants). Forms like *rājaputr**o***, *dīnār**a***, *samatā**ṁ***, thus, can be accounted for as having Apabhraṁśa morphology. (Note that the final short vowel of *dīnār**a*** is guaranteed by the meter).

There remain certain forms that suggest that in some cases the composer(s) used correct Sanskrit grammatical forms and thus showed at least some familiarity with Sanskrit grammar. Consider e.g. *santi* 'they are' in example (8). While the singular present forms of the verb *as-* 'to be' are attested in Apabhraṁśa, the third plural *santi* is not found in the extensive list of attested present forms in Prakash (1975, 260-7). Perhaps, then, *santi* replaced a metrically equivalent Ap. *honti* 'they are' (from the Sanskrit root *bhū*), but perhaps *santi* was still used in Apabhraṁśa but did not happen to be found in the sample texts examined by Prakash (even for the earlier Prakrits, Pischel (1900, 350) notes that *santi* is 'selten' (rare).[23]

We can thus conclude with Hayashi that the base language of the Bakhshali Manuscript is most likely to have been Apabhraṁśa and that the composer(s) of the text had access to the Sanskrit lexicon but were insufficiently informed about Sanskrit grammar.[24]

Although details differ, the language of BM, thus, exhibits a similar pattern of sanskritization to BHS (and EHS), but with a different base language – Apabhraṁśa vs. Pali/early Prakrit.

Most important, however, the example of BM clearly shows that the sanskritization of BHS (and the *Patna Dhammapada*?) is not unique, but that similar developments could arise independently, presumably because of the overarching prestige of Sanskrit for religious and scholarly discourse.

8 Conclusions

Buddhist Hybrid Sanskrit reflects incomplete sanskritization, mainly lexically and especially in noun and adjective stems. The process was most likely motivated by trying to defend Buddhism against Sanskrit-using brahmins in disputations at royal courts. A likely account for the mechanism of sanskritization is that it is the result of incomplete learning in a school system which required memorization of a Sanskrit thesaurus and a formal grammar of Sanskrit, and that the educational process did not continue to later stages at which the

23 Note however that the plural *santi* vs. the expected Sanskrit dual *staḥ/sto* reflects the Prakrit and Apabhraṁśa loss of the dual category.

24 A reviewer suggests to consider the language of the BM to be 'Vernacular Sanskrit', along the lines of Salomon 1989. However, doing so should not preclude attempts at trying to determine a vernacular, non-Sanskrit base for this form of language.

two components would be firmly linked with each other. The memorized thesaurus, then, would serve as a source for replacing Prakrit words with Sanskrit ones. Function words and inflectional suffixes, as well as verb forms that strongly differed from Sanskrit, however, to a great extent were not replaced. The result, then, is similar to general tendencies in linguistic borrowing, with function words and inflectional affixes less likely to be adopted (Hock 2021, 414). Occasional hypersanskritizations further attest to the fact that Sanskrit grammar was only incompletely acquired. The entire process is similar to what is found in the EHS of late first century BC/early centuries AD inscriptions, in the *Patna Dhammapada* and the Bhakshali Manuscript, as well as in medieval Sanskrit compositions in Tibet. Unlike these forms of language use, however, early Buddhist texts composed in BHS came to be institutionalized and their language was canonized as a defining feature of a specific school of Buddhism.

In many ways, then, BHS resembles the kinds of languages that have been referred to as 'Bilingual Mixed Languages' (BML), discussed e.g. in Bakker & Matras 2003 and the contributions to that volume. BMLs are commonly characterized as having undergone 'relexification' of one language based on the lexicon of another, as a deliberate process of creating a separate linguistic identity. The fact that sanskritization in BHS was largely confined to the lexicon could be considered an example of relexification, and the fact that it appears to have been motivated by an attempt at competing with Sanskrit-speaking brahmins could be considered a deliberate action. However, it is by no means clear that sanskritization itself was a deliberate act of creating a separate linguistic and social identity. Rather, the separate linguistic identity of BHS seems to be the result of the institutionalization of a product of incomplete acquisition, in terms of lexical borrowing from Sanskrit without comparable acquisition of the grammar. But even in the case of paradigm examples of 'Bilingual Mixed Languages', the creation of a separate linguistic and social identity may well have been an after-effect of massive lexical borrowing (Hock 2021, § 13.6).

With these reservations, then, BHS could be added to the set of hybrid languages subsumed under the label 'Bilingual Mixed Languages'.

Bibliography

Ānandajoti, B. (2020). *A Comparative Edition of the Dhammapada: Pāḷi Texts with Parallels from Sanskritized Prakrit, Edited Together with a Study of the Dhammapada Collection*. 4th rev. ed.
https://www.ancient-buddhist-texts.net/Buddhist-Texts/C3-Comparative-Dhammapada/Comparative-Dhammapada.pdf

Bakker, P.; Matras, Y. (eds) (2003). *The Mixed Language Debate: Theoretical and Empirical Advances*. Berlin; New York: Mouton de Gruyter.

Bronkhorst, J. (1993). "Buddhist Hybrid Sanskrit: Its Original Language". Mishra, K.N. (ed.), *Aspects of Buddhist Sanskrit = Proceedings of the International Symposium on the language of Sanskrit Buddhist Texts* (Sarnath, 1-5 October 1991). Sarnath; Varanasi: Central Institute of Higher Tibetan Studies, 396-412.

Bronkhorst, J. (2010). "The Spread of Sanskrit". Franco, E.; Zin, M.; Davidson, R.M. (eds), *From Turfan to Ajanta: Festschrift for Dieter Schlingloff on the Occasion of His Eightieth Birthday*, vol. 1. Bhairawaha; Rupandehi: Lumbini International Research Institute, 117-39.

Bühler, G. (1881). "Inscriptions". *Archaeological Survey of Western India*, vol 4. Bombay: Government Central Press.

Cone-Ānandajoti Bhikku (2017) = *Patna Dhammapada: Patna dhamma verses*.Transcribed by M. Cone; translated (and edited) by Ānandajoti Bhikku.
https://www.ancient-buddhist-texts.net/Texts-and-Translations/Patna-Dhammapada/index.htm

Damsteegt, Th. (1978). *Epigraphical Hybrid Sanskrit*. Leiden: Brill.

Edgerton, F. (1946). "Indic Causatives in *-āpayati* (*-āpeti, -āvei*)". *Language*, 22, 94-101.

Edgerton, F. (1953). *Buddhist Hybrid Sanskrit Grammar and Dictionary*, vol. 1. New Haven: Yale University Press.

Grierson, G.A. (1929). "The Language of the *Mahā-Naya-Prakāśa*: An Examination of Kāshmīrī as Written in the Fifteenth Century". *Memoirs of the Asiatic Society of Bengal*, 11(2), 73-120.

Hayashi, T. (1995). *The Bakhshālī Manuscript: An ancient Indian mathematical treatise*. Groningen: Egbert Forsten.

Hinüber, O. von. (1986). *Das ältere Mittelindisch im Überblick*. Wien: Österreichische Akademie der Wissenschaften.

Hock, H.H. (2019). "Foreigners, Brahmins, Poets, or What? The Sociolinguistics of the Sanskrit 'Renaissance'". *Journal of Asian Linguistic Anthropology*, 1(2), 10-7.

Hock, H.H. (2021). *Principles of Historical Linguistics*. 3rd rev. ed. Berlin; Boston: de Gruyter Mouton.

Hoernle, R. (1887). *On the Bakshali manuscript*. Vienna: Alfred Hölder.

Jamison, S.W. (2000). "Lurching Towards Ergativity: Expressions of Agency in the Niya Documents". *Bulletin of the School of Oriental and African Studies*, 63(1), 64-80.

Kapstein, M. (2018). "Other People's Philology: Uses of Sanskrit in Tibet and China, 14th-19th Centuries". *L'espace du sens: Approches de la philologie indienne/The space of meaning: Approaches to Indian philology*.
https://hal.archives-ouvertes.fr/hal-02190087

Karpik, S. (2023). "Light on Epigraphic Pali: More on the Buddha Teaching in Pali". *Journal of the Oxford Centre for Buddhist Studies*, 23, 41–89.

Kay[e], G.R. (1933). "The Bakshali Manuscript: Early Hindu Mathematics: A Study in Mediaeval Mathematics". *Indian Historical Researches*, 24(1).
https://ia601505.us.archive.org/4/items/in.ernet.dli.2015.33644/2015.33644.The-Bakhshali-Manuscript--Vol-1.pdf

Marciniak, K. (ed). (2019). *The "Mahāvastu"*: *A New Edition*, vol. 3. Tokyo: Soka University.

Ollett, A. (2017). *Language of the Snakes: Prakrit, Sanskrit, and the Language Order of Premodern India*. Oakland: University of California Press.

Pischel, R. (1900). *Grammatik der Prakrit-Sprachen*. Strassburg: Trübner.

Prakash, R. (1975). *Verb-Morphology in Middle Indo-Aryan*. Delhi: Munshiram Manoharlal.

Salomon, R. (1983). "The Buddhist Sanskrit of Aśvaghoṣa's Saundarananda". *Wiener Zeitschrift für die Kunde Südasiens*, 27, 97-112.

Salomon, R. (1989). "Linguistic variability in Sanskrit". Caillat, C. (ed.), *Dialectes dans les littératures indo-aryennes*. Paris: Collège de France, 275-94.

Salomon, R. (1998). *Indian Epigraphy: A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the Other Indo-Aryan Languages*. New York; Oxford: Oxford University Press.

Sen, S. (1977). "On Buddhist (Hybrid) Sanskrit". *Bulletin of Tibetan Studies*, 1, 5-8. https://api.repository.cam.ac.uk/server/api/core/bitstreams/acc1569f-bd84-494b-87cd-d2b0f37432f6/content

Senart, É. (ed.) (1897). *Mahāvastu avadānam: Le Mahâvastu*, vol. 3. Paris: Imprimerie Nationale.

Tagare, G.V. (1948). *Historical grammar of Apabhraṁśa*. Poona: Deccan College Postgraduate and Research Institute.