

MAURITIAN BHOJPURI: AN INDO-ARYAN LANGUAGE SPOKEN IN A PREDOMINANTLY CREOLOPHONE SOCIETY

Philip Baker and P. Ramnah

INTRODUCTION

Varieties of a language called Bhojpuri are spoken in India by some 20 million people on both sides of the Uttar Pradesh - Bihar border in an area stretching from Nepal in the north to Madhya Pradesh and Orissa in the south (Tiwari 1960). In Mauritius, the name Bhojpuri is also applied to a language spoken, to a greater or lesser extent, by a third or more of its population, for the most part people whose ancestors came from north eastern India.¹

Mauritian Bhojpuri (MB) has thus far received scant attention from linguists, with the result that little is known of how this differs from Indian Bhojpuri (IB). While the authors of this paper are not specialists in the study of Indo-Aryan languages,² and do not therefore aspire to provide the definitive answer to that question, we do aim to make a contribution towards that end by setting out some of the more basic facts concerning MB.³ We also examine critically the popularly-held view that the relationship of MB to Hindi is similar to that which holds between two other languages of the island, Mauritian Creole (MC) and French (cf. Bhuckory 1965). We also describe the phonemic correspondences between MB and MC, and vice versa.

Our paper consists of five sections. In section 1, we discuss the phonemic inventory of MB. In sections 2 and 3, we examine aspects of MB nouns and verbs, respectively. In section 4, the basic vocabulary of MB, IB and Hindustani⁴ is compared. In section 5, we look at mutual lexical influence between MB and MC, and this is followed by some concluding remarks.

1. THE PHONEMIC INVENTORY OF MB

We have provisionally identified 41 phonemes of MB but have reservations concerning certain of them as detailed below. Compared with IB, this total is three more than those found in the Sadani dialect by Jordan-Horstmann (1969: 19)⁵ but several fewer than the number identified for Ballia district by Tiwari (1960:3-7).⁶

As we cite both MB and MC examples below, it is convenient to do so in the harmonised orthography proposed for these two languages by Baker and Hookoomsing 1983 (B/H). (Since the latter closely resembles established conventions for transcribing Indian languages in roman characters, this will also facilitate

Papers in pidgin and creole linguistics No.4, 215-238.
Pacific Linguistics, A-72, 1985.

© Philip Baker and P. Ramnah

comparison with IB and Hindustani forms cited further below). The phonemes of MB and MC are set out in Table 1 together with the B/H orthography adopted for their representation in this paper.

Table 1 shows that MB has a 4 × 5 system of stops typical of Indo-Aryan languages. There is an abundance of minimal pairs in which these stops are contrasted initially. For example, the full range of labial stop contrasts is to be found in rather basic items of vocabulary:

contrast		examples			
p	vs. ph	peṭ	<i>belly</i>	vs. pheṭ	<i>mix</i>
p	vs. bh	pāp	<i>sin</i>	vs. bāp	<i>father</i>
p	vs. bh	par	<i>on</i>	vs. bhar	<i>fill</i>
ph	vs. b	phaḷ	<i>fruit</i> ⁷	vs. baḷ	<i>strength</i>
ph	vs. bh	phuḷ	<i>flower</i>	vs. bhūḷ	<i>forget</i>
b	vs. bh	bāt	<i>talk</i>	vs. bhāt	<i>rice</i>

Word finally, however, the distinction between aspirated and non-aspirated stops is not found in the speech of some of our informants. For example, the IB words *jibh tongue*, *dudh milk* and *nokh fingernail* have final aspirated stops in the speech of some MB-speakers, whereas others consistently substitute the corresponding non-aspirated stops, pronouncing these *jib*, *dud* and *nok*, respectively. In the following pages, we will transcribe with a following *h* all final stops which are aspirated in the speech of some of our informants. Further research would be needed to determine whether such loss of aspiration word finally is to be associated with younger rather than older speakers, or with certain geographical areas.⁸

In contrast to its rich system of stops, MB has few fricatives amongst its phonemes, and the status of two of these is doubtful. In Sadani IB, Jordan-Horstmann (1969:25, 28) identifies retroflex [ɽ] and [ɽ^h] as allophones of /ḍ/ and /ḍ^h/, respectively, in the environments /V__V/, /V__C/ and /V__#. In MB, however, the distribution of ḍ and ɽ, on the one hand, and ḍ^h and ɽ^h, on the other, is not entirely complementary. For example, MB haḍi *bone* has a retroflex stop in an environment where one would expect to find a retroflex fricative in Sadani IB. (We have not been able to determine the corresponding IB form.) This may perhaps be related to the fact that the corresponding Hindustani term, haḍḍi, has a geminate consonant. There are also some MB words which have retroflex ɽ where an alveolar r is found in the corresponding IB and Hindustani forms, as in MB chokrā, IB and Hindustani chokrā *boy*. For such reasons, we have tentatively identified both ɽ and ɽ^h as phonemes of MB.

Like Sadani IB, MB has just two nasal phonemes, /m/ and /n/. As noted at the foot of Table 1, phonetic [ŋ], [ŋ] and [ŋ] are heard immediately preceding retroflex, palatal and velar stops, respectively, but are considered allophones of /n/. Morpheme finally, [ŋ] (in unmarked forms) alternates with [ŋg] (in marked forms), as in māng [maŋ] *parting (in the hair)* and mangwā [maŋgwa] *the parting (in the hair)*. This leads us to assume that [ŋ] and [ŋg] are positionally conditioned variants of underlying /ŋg/.

For Jordan-Horstmann, *y* and *w* in Sadani IB are merely allophones of /i/ and /u/ respectively. In MB, the position is somewhat different. /i/ is optionally realised as [i] or [y] in the environments /C(_v__V). (On the use of [y] for IPA

Table 1: The phonemes of Mauritian Bhojpuri and Mauritian Creole												
Consonants		Labial/ Labio-dental	Dental	Retroflex	Palatal	Velar/Glottal						
STOPS -v -asp -v +asp +v -asp +v +asp	MB	p p ^h b b ^h	t t ^h d d ^h	ʈ ʈ ^h ɖ ɖ ^h	ç ç ^h j j ^h	k k ^h g g ^h						
	MC	p b	t d		ç j	k g						
	orthography	p ph b bh	t th d dh	ʈ ʈ ^h ɖ ɖ ^h	c ch j jh	k kh g gh						
FRICATIVES -v +v -v +v	MB		s	ʂ ʂ ^h		h						
	MC	f v	s z			ɣ						
	orthography	f v	s z	ʂ ʂ ^h		h r						
NASALS +v +v	MB	m	n	- ¹	- ¹	- ²						
	MC	m	n		- ³	- ²						
	orthography	m	n									
OTHERS +v +v	MB	w	l r									
	MC	w	l		y ⁴							
	orthography	w	l r		y							
VOWELS MB	i	ĩ			ũ	u						
		e	ẽ		õ	o						
				ʌ	ɤ							
			a		ã							
MB orthography	i	iñ	e	eñ	ã	a	añ	añ	oñ	o	uñ	u
VOWELS MC	i											u
		e			(ə)					o		
			ẽ						õ			
			a				ã					
MC orthography ⁵	i	e	eñ/em	a	(è)	añ/am	oñ/om	o				u

¹In our MB data, retroflex and palatal nasal consonants are found only immediately preceding retroflex and palatal stops, respectively. Both retroflex and palatal nasals are thus regarded as allophones of /n/.
²In both MB and MC, the sequence of phonemes /ng/ is realised phonetically as [ŋg] between vowels and as [ŋ] word-finally.
³In MC, the sequence of phonemes /ny/ is realised as [ny]⁴ between vowels, and as [ɲ] word-finally.
⁴Here and throughout this paper, y represents a palatal glide (IPA [j]), in part to avoid possible confusion resulting from IPA's use of a variant of the same character for a palatal stop.
⁵For details of the use of both ñ and ñ in the representation of nasalised vowels, see footnote 19.

[j], see note 4 at the foot of Table 1.) Thus *dhaniā coriander* and *māi mother*, for example, are liable to be pronounced [d^hʌnya] or [d^hʌnia] and [may] or [mai] in any position. In other environments, [y] or [i] appear to be in complementary distribution, and we are thus satisfied that [y] does not have phonemic status in MB. However, not all occurrences of [y] are to be derived from /i/: the /e/ of verbal inflections of which this is the initial element becomes a palatal glide wherever the final element of the verb stem is a vowel. For example *mar-die*, *mareke to die*, but *khā eat*, *khāeke [k^hayke] to eat*.

With regard to *w*, we are aware of only a handful of morphemes in which [w] alternates freely with [u], such as *dewtā ~ deutā god*. In the great majority of words containing either [w] or [u], no such variation is found. For example, *awrat woman* is consistently [ʌwrat] not *[ʌurat], *dhuān smoke* is consistently [d^huā] not *[d^hwā], and *mwāw kill* is consistently [mwaw] not *[muaw], *[mwau] or *[muau]. While we have not found any minimal pairs distinguished solely by a contrast between [w] and [u], we have also failed to find any pattern(s) of distribution of these two sounds which would suggest that they are allophones. We are thus tentatively inclined to consider /w/ and /u/ to be separate phonemes of MB. In further support of this, it may be said that the rules for deriving the marked forms of MB nouns from their unmarked forms, set out in section 2 below, would be considerably more complicated if surface *w* were to be derived from underlying *u*.

In romanised transcriptions of Indian languages, it is customary to mark long vowels with the macron, and this is often done even where short and long varieties of a particular vowel are allophones of a single phoneme rather than two distinct phonemes. In MB, as in Sadani IB (Jordan-Horstmann 1969), differences in vowel length are subphonemic in the case of each of the eight vowel phonemes written *i*, *īn*, *e*, *ēn*, *o*, *ōn*, *u* and *ūn* in the B/H orthography. Of the remaining four vowels, *ā* and *ā̄n* are perceptibly longer than *a* and *ān* but the more immediately striking difference between the two pairs is that *ā* and *ā̄n* are low peripheral vowels whereas *a* and *ān* are low central vowels (see Table 1). As there are minimal pairs such as *dhān young rice* and *dhan property*, *bhāng cannabis in milk* and *bhang disruption*, *dām price* and *dam courage*, *mān respect* and *man desire*, it is clear that *ā* and *a* are distinct phonemes in MB. Similar examples might be given to illustrate *ā̄n* and *ān* contrasts. However, MB, like IB (Tiwari 1960:104), has a rule whereby *ā̄(n)* in the unmarked forms of nouns is changed to *ā(n)* in marked forms ending *-wā* or *-wan* (see also below). A consequence of this is that *dhanwā* may mean either *the young rice* or *the property*. (No similar ambiguity arises in the case of the other three pairs of examples because one member of each — *bhang*, *dam* and *mān* — is a non-countable abstract noun which cannot occur with either *-wā* or *-wan*.)

2. THE NOUN IN MB

Count nouns in MB have three forms: an unmarked (zero) form which may have a singular or plural reading, according to context, and forms marked by the addition of *-wā* and *-wan* (and variants of these two forms, as discussed below). Both *-wā* and *-wan* are broadly equivalent, semantically, to the definite article of English or the postposed *-la* of MC, but *-wā* is exclusively singular while *-wan* is found only in plural contexts. The marked and unmarked forms of a number of MB nouns are set out in Table 2.

Table 2: Unmarked and marked forms of selected MB nouns

	A	B	C
Specificity	-	+	+
Plurality	∅	-	+
English gloss	_____ (s)	the _____	the _____s
1. <i>bird</i>	ciraiñ	ciraiñwā	ciraiñwan
2. <i>table</i>	latāb	latabwā	latabwan
3. <i>eye</i>	āñkh	añkhwā	añkwan
4. <i>hand</i>	hāñt	hañtwā	hañtwan
5. <i>musical instrument</i>	bajā	bajwā	bajwan
6. <i>fried snack</i>	bhajiā	bhajiawā	bhajiawan
7. <i>greens</i>	bhāji	bhajiā	bhajian
8. <i>tunic</i>	bāju	bajuwā	bajuan
9. <i>person</i>	admi	adamiā	adamian
10. <i>boy</i>	chokṛā	chokaṛwā	chokaṛwan
11. <i>girl</i>	chokṛi	chokaṛiā	chokaṛian
12. <i>water</i>	pāni	paniā	panian
13. <i>knee</i>	thewni	thewniā	thewnian
14. <i>ashtray</i>	sāndrie	sañdriewā	sañdriewan
15. <i>car</i>	loto	lotowā	lotowan
16. <i>egg</i>	aṇḍā	aṇḍwā	aṇḍwan
17. <i>dog</i>	kutā	kutwā	kutwan
18. <i>fire</i>	āg	agwā	agwan
19. <i>cloud</i>	badār	badarwā	badarwan
20. <i>woman</i>	awrat	awratiā	awratian

The unmarked form of MB nouns resembles the 'zero article for nonspecific NP' which Bickerton (1977:58, 1981:56) identifies as a typical feature of Creole languages. Thus Bickerton's translation of his 1977 Guyanese example illustrating its use — *mi go bai buk I shall buy a book or books* (even the speaker does not know which) — applies equally well to its approximate equivalents in both MC (word order: *I + FUT + go + buy + book*) and MB (word order: *I + book + buy + FUT-go*): MC *mo pu al aste liv*, MB *ham liv⁹ kine jāb*. It might thus seem that, in this respect at least, MB had acquired a Creole characteristic. Such

a conclusion would, however, be premature, for two reasons. First, there are many Hindustani nouns for which the distinction between singular and plural is not marked in the direct case. For example, Hindustani ghar may refer to *house* or *houses* (Barz, p.c. Barz adds that plurality is marked in the oblique case, *ghar meñ in the house* versus *gharõñ meñ in the houses*). Secondly, Hertig-Shalická remarks, with reference to the songs she collected in the Ballia district of India, that 'plural is often not marked at all' (1974:129). It may thus be that the potential for unmarked singular nouns to be employed in plural contexts had already been realised in IB before large numbers of IB-speakers reached Mauritius in the 19th century.

Tiwari notes up to three singular and four plural forms of each noun stem in IB. He labels the singular forms 'short', 'long' and 'redundant' (1960:104). The last-named appears to be entirely unknown in MB, while the 'short' and 'long' closely resemble, respectively, the unmarked and marked MB forms set out in columns A and B of Table 2. While Tiwari does not indicate that the 'long' form is associated with definiteness, this was earlier signalled by Grierson: '... the long form is used either in a non-honorific sense or to give definiteness. Thus ghoṛā (short form) a horse, but ghōr'wā the horse (1883:25). With regard to the nouns which in MB have the unmarked forms *ghar house* and *gāi cow*, Tiwari states that there is 'no difference of sense' (1960:107) between the four variant IB plural forms which are, respectively, *gharanh*, *gharanhi*, *gharan*, *gharani* and *gāinh*, *gāinhi*, *gāin*, *gāini*. In MB, these nouns have just one marked plural form, respectively *gharwan* and *gāiwan* (cf. the forms listed in column C of Table 2). While all of Tiwari's plural endings include the consonant *n*, as do the MB forms, the precise source of the latter has yet to be identified.

It will be apparent from Table 2 that the marked forms listed in columns B and C can largely be predicted from the unmarked forms in column A. In fact, with the exception of no.20, the marked forms of all the MB nouns in our data can be derived from three rules, provided that two assumptions are made. The first is that the underlying forms of the marked singular and plural terminations are, respectively, *-wā* and *-wan*. The second assumption is that the underlying forms of numbers 9, 10 and 11 are, respectively, **adami*, **chokaṛa* and **chokaṛi*. The three rules are:

$$1. \quad \bar{a} \rightarrow a / \dots _ \dots + \begin{pmatrix} w\bar{a} \\ wan \end{pmatrix}$$

All occurrences of \bar{a} in the underlying forms of unmarked nouns are obligatorily changed to *a* when the terminations *-wā* or *-wan* are added.

$$2. \quad w \rightarrow \phi / i _ \begin{pmatrix} \bar{a} \\ an \end{pmatrix}$$

The *w* of either of the terminations denoting specificity is deleted wherever it immediately follows *i*.

$$3. \quad a \rightarrow \phi / VC _ CV(N)$$

In nouns of three or more syllables (counting the final syllable of marked forms), *a* is deleted from the penultimate syllable provided this follows, and is followed by, a single consonant.

Apart from no.20,¹⁰ rules 1, 2 and 3 will generate all the forms in columns B and C of Table 2, while rule 3 alone will produce all the forms in column A. Some sample derivations are set out in Table 3.

The indefinite article is not distinguished from the number 'one':

ego ber a/one drinking glass
ego gilās a/one metal beaker¹¹

Strictly speaking, ego is not a single morpheme but the fusion of two: ek a/one and go a classifier for which there is no equivalent in English. Go behaves in some respects like a collective or quantitative noun and does not co-occur with the latter. For example:¹²

ego jālmeth a match
ek bwāt jālmeth a box of matches
dugo ālu two potatoes
du libar ālu two pounds of potatoes
tingo sigāret three cigarettes
tin pāki sigāret three packets of cigarettes
ego ber a glass
ek ber pāni a glass of water

The final example shows that ber functions as both a count noun and as a quantitative noun. The classifier go is a feature which MB has inherited from IB (cf. Tiwari 1960:120), and which sets both these off from Hindustani. For example, the equivalents in modern Hindi to the last pair of MB examples have nothing corresponding to go and are: tin sigreṭ and tin pāket sigreṭ. (Both Hindi pākeṭ and sigreṭ are from English, whereas the corresponding MB terms are from French through MC.)

Like IB, but in sharp contrast to most Indo-Aryan languages, MB does not have grammatical gender. It does, however, have separate masculine and feminine forms of nouns for certain categories of people and domestic animals. Most of these belong to one of two groups. In the first group, masculine forms have final ā while feminine forms have final i:

beṭā son beṭi daughter
cācā paternal uncle cāci paternal aunt
ghoṛā stallion ghoṛi mare
murgā cock murgi hen

In the second group, feminine forms add -in to the masculine or unspecified forms:

dulhā bridegroom dulhin bride
kiriol Creole person kiriolin Creole person (f)
musalmān or laskār Muslim musalmānin or laskārin Muslim (f)
cinwā Chinese person cinwāin Chinese person (f)
(cinwawā the Chinaman)

Table 3: Sample derivations of unmarked and marked forms of MB nouns

A unmarked forms									
underlying forms:	bajā	bhajiā	bhāji	bāju	*chokaṛā	*chokaṛi	*adami	hānt	aṇḍa
rule 3					chokṛā	chokṛi	admi		
surface forms	bajā	bhajiā	bhāji	bāju	chokṛā	chokṛi	admi	hānt	aṇḍā
B with -wā termination									
underlying forms:	bajā	bhajiā	bhāji	bāju	*chokaṛā	*chokaṛi	*adami	hānt	aṇḍā
add termination	*bajāwā	*bhajiāwā	*bhājiwā	*bājuwā	*chokaṛāwā	*chokaṛiwā	*adamiwā	*hāntwā	*aṇḍāwā
rule 1	*bajawā	bhajiawā	*bhajiwā	bajuwā	*chokaṛawā			hāntwā	*aṇḍawā
rule 2			bhajiā			chokaṛiā	adamiā		
rule 3	bajwā				chokaṛwā				aṇḍwā
surface forms	bajwā	bhajiawā	bhajiā	bajuwā	chokaṛwā	chokaṛiā	adamiā	hāntwā	aṇḍwa
C with -wan termination									
add termination	*bajāwan	*bhajiāwan	*bhājiwan	*bājuwan	*chokaṛāwan	*chokaṛiwan	*adamiwan	*hāntwan	*aṇḍāwan
rule 1	*bajawan	bhajiawan	*bhajiwan	bajuwan	*chokaṛawan			hāntwan	*aṇḍawan
rule 2			bhajian			chokaṛian	adamian		
rule 3	bajwan				chokaṛwan				aṇḍwan
surface forms	bajwan	bhajiawan	bhajian	bajuwan	chokaṛwan	chokaṛian	adamian	hāntwan	aṇḍwan

As this -in suffix recalls the French feminine suffix written '-ine', it would be as well to demonstrate that this is indeed of Indian origin rather than from French via MC. Tiwari notes this feminine suffix as -inĪ, e.g. *dulahinĪ bride* (1960:107), while Jordan-Horstmann gives the Sadani form as -in, e.g. *bāgh tiger*, *bāghin tigress* and *loharā blacksmith*, *loharāin blacksmith's wife* (1969:63).

3. THE VERB IN MB

As the verbal system of MB is fairly complex, we shall not attempt to do more than give a general outline of its nature in the following paragraphs.

Excluding causatives (see below) and imperatives (not discussed due to lack of sufficient data), there are basically 21 inflected forms (not counting variants of these) of every MB verb, as set out in Table 4.¹³ To the best of our knowledge, all MB verbs have the same set of inflected forms. That is to say, insofar as MB has any 'irregular' verbs, the latter have more than one stem form to which regular inflections are added.

Table 4: Verb forms in MB			
Illustrated with <i>cal-</i> go, walk			
PERSON/NUMBER/STATUS*	PRESENT	FUTURE	PAST
1S	calilā	calab	{calni calli}
1PL	calila sa	calab sa	{calni calli} sa
2S	calele	calbe	calle
2S HON	calela	calba	calla
2PL	calela sa	calba sa	calla sa
2PL HON	calela jā	calba jā	calla jā
3S	calelā	cali	callak
3S HON	calelan	calihan	callan
3PL	calelan sa	calihan sa	callan sa
3PL HON	calelan jā	calihan jā	callan jā
PARTICIPLES	caleke (infinitive)		
	calat (habitual)		
	calal (past)		
	calte (present continuous)		
	calke (past continuous)		
*1, 2, 3 = respectively, first, second and third person S = singular, PL = plural, HON = honorific form			

At first sight, it would appear that Table 4 contains 37 forms rather than the 21 mentioned above. However, it will be seen all the plural forms include the particles *sa* or *jā*. We have tentatively assumed that these particles ought not to be regarded as inflections. In addition, it will be seen that some of the inflected forms occur more than once in the table. When allowance is made for these matters, it will be found that the 21 inflected endings are as follows (in alphabetical order): -ab, -al, -at, -ba, -be, -eke, -ela, -elā, -elan, -ele, -i, -ihan, -ila, -ilā, -ke, -la, -lak, -lan, -le, -li, -ni.

MB also has a causative particle which may be infixed to any of the forms in Table 4 (including variants). The infix in question has, in fact, two forms and these appear to be free variants. This makes a total of 63 inflected forms of each MB verb (and, as stated above, this figure excludes imperatives).

The MB equivalent of certain common verbs of MC (and both English and French) is a combination of a noun and a verb. For example, *to work* is expressed in MB as *kām work* (noun) + *kar to do*. The paradigm of 'present' forms corresponding to 'work' is as follows (S = singular, PL = plural, HON = honorific):

	ham kām karilā	<i>I work</i>
	tu kām karele	<i>you work</i> (S)
	tu kām karela	<i>you work</i> (S, HON)
	u kām karelā	<i>he/she works</i>
	u kām karelan	<i>he/she works</i> (HON)
	hamni kām karila sa	<i>we work</i>
tolog	} kām karela sa	<i>you work</i> (PL)
toholog		
tolog	} kām karela jā	<i>you work</i> (PL, HON)
toholog		
ulog	} kām karelan sa	<i>they work</i>
olog		
ulog	} kām karelan jā	<i>they work</i> (HON)
olog		

The verb forms noted so far do not appear to correspond to any single dialect of IB but rather to a mixture of such dialects (Suchita Ramdin, p.c.; see Domingue 1981, Barz 1980:4). The system sketched here is a good deal less complex than that described by Tiwari 1960 in two respects. First, he notes separate masculine and feminine verbal forms in IB whereas there are no gender distinctions in the MB system. Secondly, he gives 'contemptuous' as well as 'ordinary' and 'honorific' forms for the second and third person, both singular and plural. MB does not have contemptuous forms as such.¹⁴

In addition to the forms noted so far, MB verbs are frequently conjugated with a following auxiliary verb. There are a dozen such auxiliaries in MB, all of which also function as full verbs in their own right, such as those meaning 'reach', 'stay', 'go', 'finish', etc. As they are essentially the same as those given by both Tiwari (1960:182-187) and Jordan-Horstmann (1969:98-102), we will not give examples here.

All MB verbs having the range of inflected forms described above appear to be of Indo-Aryan origin. MB does, however, have some verbs adopted from MC. The latter are all invariable in MB and are conjugated with *kar- do* or an auxiliary verb:

ham	telefone	karli	
<i>I + telephone + do-</i>			FIRST PERSON PAST
<i>I</i>			<i>telephoned</i>
ham	pedāle	karat	hāi
<i>I + pedal do-</i>			HABITUAL PARTICIPLE + <i>be-</i> FIRST PERSON PRESENT
<i>I am</i>			<i>pedalling (on a bicycle)</i>

The Hindustani equivalents of the above examples, in which words adopted from English replace the MC-derived terms in MB, are:

maiñ ne fon kiyā
 I + ERGATIVE MARKER + telephone + did

maiñ ne sāikil calāī
 I + ERGATIVE MARKER + cycle + caused to move

We are grateful to Barz (p.c.) for these examples.

While the MB verbal system is somewhat less complex than that of IB, as described by Tiwari (1960) or Jordan-Horstmann (1969), it has certainly not been radically restructured in any way. The suggestion that MB is 'un créole à base indienne' (Chaudenson 1979:34), a view also taken by Stein (1982:131), thus appears to be unfounded.

4. THE BASIC VOCABULARY OF MB

The basic vocabulary of MB is illustrated in the comparative word list (p.231) where the MB equivalents of the 100 items in the Swadesh list are set out. Of these words, five are of MC origin. Numbers 3 and 76 are both from MC *lapo bark*, *skin* (<Fr *peau*), number 56 is from MC *moñtany* (<Fr *montagne*), number 65 is from MC *sime* (<Fr *chemin*), and number 82 is from MC *ros* (<Fr *roche*).¹⁵ All others are of Indian origin. (A few of them are recognisably cognate with the corresponding French or English terms, a reminder that the latter are distantly related to Bhojpuri.)¹⁶

For the purposes of comparison, IB and Hindustani forms are set out in the word list alongside those of MB. The Hindustani forms are taken from Forbes 1859, a dictionary containing ca. 40,000 entries and compiled at the time when Indian immigration in Mauritius was at its height. For further details, see the notes. The IB forms are drawn from Tiwari 1960, Jordan-Horstmann 1969 and Hertig-Skalická 1974. As none of the latter is a dictionary and as the total number of IB words for which these publications provide an English gloss is probably ca. 3,000, our knowledge of IB items in the list is rather limited. We cannot be sure that the IB forms found are necessarily the most frequent terms for the words in the Swadesh list nor, more importantly, that in the cases where we have not found a cognate form in IB, one does not in fact exist.

In comparing MB with both IB and Hindustani, we will assume that variation between *i* and *iy* in the word list is without significance. As we have been unable to establish an IB form in the case of ten of the words in the Swadesh list and as there are also five MB words of MC rather than Indo-Aryan origin, the total number of forms which may usefully be compared is 85. Of these, MB appears to be significantly closer to IB than to Hindustani in 22 cases (numbers 6, 13, 20, 36, 37, 38, 44, 45, 53, 55, 64, 66, 69, 77, 81, 87, 91, 92, 95, 97, 98 and 100), apparently closer to Hindustani than to IB in nine cases (numbers 18, 26, 27, 28, 35, 42, 61, 78 and 99), and equally close to both IB and Hindustani in the remaining 54 cases. While these figures suggest that MB is lexically closer to IB than to Hindustani, as expected, the list as a whole indicates that all three languages are very closely related, at least in their basic vocabulary. Of the nine cases where MB appears to be closer to Hindustani, it should be noted that the IB forms of numbers 26, 27, 42, 61 and 99 differ from MB and Hindustani only in having a final *-i*, and that Domingue (1981:153) remarks that lack of final *-i* is characteristic of western varieties of IB. It could be, therefore, that only in the case of numbers 18, 28, 35 and

78 is MB significantly closer to Hindustani than IB. That said, we must also acknowledge that the teaching of Hindi in Mauritius and, in particular, the great popularity enjoyed by Hindustani-speaking films in the island, have had some influence on the lexicon of MB in recent decades.

5. MC AND MB: MUTUAL LEXICAL INFLUENCE

Although only five of the MB words in the Swadesh list are of MC origin, there is no doubt that the total number of MC words which have been adopted in MB is very considerable indeed, and could well exceed 1,000.¹⁷ The number of MB words which have been adopted in MC is also considerable. Baker (1982b:314-383) lists some 300 MC words adopted from Indo-Aryan languages and most, but not all, of these are likely to have MB as their immediate source.¹⁸

The phonemic correspondences between MC and MB, and vice versa, can be described most economically by reference to Table 1. In general, MB stops are assimilated to the corresponding MC stops and vice versa, with aspirated MB stops assimilating to the corresponding non-aspirated MC stops and with retroflex MB stops assimilating to MC alveolar stops. There are, however, two partial exceptions. The dental stops in MC are regularly palatalised or lightly affricated before *i* and *y*. MC dental stops in this position are unpredictably assimilated to either palatal affricates or dental stops in MB. By unpredictable, we mean that we cannot immediately offer an explanation as to why MC *pudin pudding* has given MB *pujin* whereas MC *diber butter/margarine* has given MB *diber* (with of course no palatalisation or affrication of the *d* in MB). The other partial exception concerns MB /*ph*/. While there are many cases where MB /*ph*/ becomes MC /*p*/, such as MB *phukni blow-pipe (for making dying embers flare up)* : MC *pukni*, there are at least three cases where MB /*ph*/ instead becomes MC /*f*/, including MC *farata kind of unleavened bread* (MB *pharathā*, Hindustani *parāṭhā*) and MC *fenus milk from a cow which has recently calved* (MB and Hindustani *phenus*). While the normal reflex of MC /*f*/ is MB /*ph*/, it seems a little surprising to find sporadic examples of the reverse process.

MC has four fricatives which lack phonemic status in MB: *f*, *v*, *z* and *r*. The established correspondences concerning the first three of these are MC /*f*/ : MB /*ph*/, MC /*v*/ : MB /*b*/, and MC /*z*/ : MB /*j*/, as illustrated by MB *phātige tired* (MC *fatige*), MB *sabāt thongs (cheap footwear)* (MC *savat*), and MB *joli beautiful* (MC *zoli*). These examples, and many others, are firmly established in MB. In what appear to be more recent adoptions from MC, such consonants are pronounced variably as fricatives and stops. MC *liv book*, for example, is variably pronounced [liv] and [lib] in MB, even by the same speaker. Similar variation is found in words recently adopted by MB from Hindustani. For example, the Hindustani word *fāida profit, gain* heard in many Indian films shown in Mauritius, is now current in MB where it is variably pronounced with initial [ph] or [f].

MC /*r*/ is realised as [ɻ] in the environment /__V but elsewhere is the second element in centring diphthongs ([^ə]) or a vowel lengthener and modifier, for example MC /*ar*/ is pronounced [a:] word finally and when immediately followed by a consonant. Given this diversity of realisation, it is perhaps a little surprising that MC /*r*/ is everywhere assimilated to the alveolar *r* of MB. (There is thus no conflict in the B/H orthographic proposal of graphic *r* for both MB /*r*/ and the phonetically very different MC /*r*/.)

MB h has no regular reflex in MC, h from all sources being generally deleted in the latter except in a few prestige words such as *haji Muslim who has made the pilgrimage to Mecca* and *holi a Hindu festival*. (Both words may well have been adopted into MC from Indian languages other than MB.)

There is a regular, two way correspondence between each of the four remaining phonemes common to MB and MC: /m/, /n/, /w/, /l/.

In MC [i] and [y] are separate phonemes whereas in MB both are mere allophones, and apparently free variants, of /i/. As might be expected, the reflex of both MC /i/ and /y/ is MB /i/. In the opposite direction, the reflex of MB /i/ is generally MC /y/ in the environments /__V and /V__ and /i/ elsewhere. In a few cases the free variation between [i] and [y] in MB has led to both forms becoming established in MC, e.g. MB *dain witch* is current as both *dayn* and *dain* in MC.

The remaining MB phonemes are /r/, /r̥/ and /r̥h/. /r/ and /r̥/ are both assimilated to MC /r/ in the environment /__V (MB *jirā* : MC *jira cummin*, MB *peṛā* : MC *pera a cake*) while MB /r/ is generally assimilated to MC /l/ in other positions (MB *barphi* : MC *balfi*). Our data do not include any examples of MB words containing /r̥/ in the latter position which have been adopted in MC, nor any examples of MB words containing /r̥h/ in any position which have become established in MC.

With regard to vowels, there is a largely regular, two-way correspondence between each of the vowels written i, e, eñ, oñ, o and u in both MC and MC in the B/H orthography. We are not currently aware of any MB words containing the nasal vowels iñ and uñ which have been adopted into MC. The MB vowels a and añ are regularly assimilated in MC as a and añ, respectively, while the latter MC vowels are variably assimilated in MB as ā or a and āñ or añ, according to the following rules:

In MC words with a single a or añ, the corresponding MB term will have ā or āñ, respectively. For example:

MC	MB	
bwat	: bwāt	<i>box</i>
lapo	: lāpo	<i>skin, bark</i>
pake	: pāki	<i>packet</i>
sigaret	: sigāret	<i>cigarette</i>
fatige	: phātige	<i>tired</i>
sāndrie	: sāndrie	<i>ashtray</i>

In MC words with more than one a or añ, the rightmost of these is assimilated as, respectively, ā or āñ in MB and all others become a or añ. For example:

MC	MB	
latab	: latāb	<i>table</i>
malad	: malād	<i>ill</i>
salad	: salād	<i>salad</i>
marsan	: marsāñ	<i>merchant</i>
kāmpman	: kāmpmāñ ¹⁹	<i>seaside bungalow</i>

It must be emphasised that, where the above are nouns, these correspondences apply to their unmarked forms. As indicated earlier, ā and āñ in unmarked MB nouns are regularly changed to a and añ in marked forms of these nouns, cf. Table 2.

The above correspondences apply to the great majority of MC words adopted into MB, and vice versa. There remain, however, a few exceptions to which we now turn our attention.

The usual MC term for cat is *sat* but there is also a diminutive alternative, *mimi*. The latter may refer in isolation to a cat of either sex. Where sex is relevant, there is an exclusively male term, *matu*, and, in contexts where the latter is employed, *mimi* is understood to be female. All three terms are of French origin. The most frequent term for cat in MB is *mini* which may be employed without reference to the creature's sex. Where sex is relevant, MB has the choice between both Indian-derived *bilār* (male) and *bili* (female) on the one hand, and *maku* (male) and *mini* (female), on the other. That the latter pair derive from MC *matu* and *mini* seems certain, but we can find no obvious reason for the irregular correspondences *t:k* and *m:n*.

The usual MC term for the local variety of tomato is *pomdamur* (<regional Fr *pomme d'amour*) and its phonetically irregular MB reflex is *paldāmun*. We are unable to account for the form of the initial syllable – were the initial consonant aspirated, one might suspect the attraction of *phal fruit* – but the final *n* may perhaps be related to the /n/ ~ /r/ variation found in a number of MB terms such as *gagna* ~ *gagra water-pot*.

Three of the MB words adopted from MC which have been mentioned elsewhere in this paper do not have the vowels predicted by the correspondences set out earlier. *Pāki packet* differs in its final vowel (MC *pake*, <Fr *paquet*). This particular change may perhaps have been favoured by the fact that a substantial proportion of MB nouns have a final *-i* (the great majority of all those which, in former times, had feminine gender in IB, in fact). If that is so it must be noted that it does not apply to all words adopted from MC, cf. *sāndrie ashtray* (MC *saṅdrie*, <Fr *cendrier*). The second case is that of *motaiñ mountain* in which the first vowel is oral rather than nasal as in MC *moñtany* (<Fr *montagne*). As there is a tendency for nasalised vowels in MC to become denasalised in certain positions (see below) – though not yet attested in the MC word *moñtany* – and as our data includes a number of items having an initial nasal consonant in MB in which the following vowel or sequence of vowels is variably nasalised – e.g. *naiñ* ~ *nai not* and *muñ* ~ *mu mouth* – little significance can be attached to the loss of nasality displayed by the *o* of *motaiñ*. The third case is MB *simā path* (MC *sime*). At first sight it appears that *ā* is here derived from *e*, a correspondence not found in any other MB word of MC origin so far as we are aware. Throughout the 19th century, however, the MC word is attested in spellings indicating that the final vowel was nasalised (as it is in its French etymon, *chemin*). Thus, if the word were adopted in the 19th century, the correspondence would be between MC *eñ* and MB *ā*. Phonetically, MC *eñ* has a value, varying from speaker to speaker, in the range [ɛ̃] ~ [æ̃]. This leads us to speculate that (former) *eñ* may have been assimilated to MB *ā* because this was the MB front vowel closest to MC *eñ* in terms of vowel height. Support for this view is also to be found in the MB word for *sugar mill*, *mulā* (MC *muleñ*, <Fr *moulin*). As the majority of Indian immigrants who reached Mauritius in the 19th century were taken there for the specific purpose of working in sugar, there can be little doubt that this was amongst the first words acquired in their new environment. (The loss of nasality exhibited by MC *sime* but not by MC *muleñ* results from a modern trend to replace nasalised vowels with the corresponding oral vowels in the environment /Vm_#, cf. MC *pima chilli* (<Fr *piment*) and MC *legzame examination* (<Fr *l'examen*).

The usual modern MC term for axe is *lars*, phonetically [lɑ:s] (<Fr *la hache*). As the corresponding MB term is *lahās* [lɑhas], it might seem that speakers of MB had adopted this directly from French rather than from MC. We reject this for two reasons. First, apart from one possible exception discussed below, all MB nouns of ultimate French origin which have an initial syllable corresponding to a French article also have this agglutinated article in MC. In other words, MB nouns with an initial syllable wholly derived, ultimately, from a French article, form a subset of the very considerable number of such nouns in MC. We thus assume that the immediate source of the MB form is MC rather than French. Secondly, an alternative, if now fairly rare pronunciation of the MC term for axe is disyllabic *laas*. Just such a pronunciation is consistently implied in the spellings *la hace*, *la-hace* found in 19th century MC texts (Descroizilles 1867, Anderson 1885, Baissac 1888). As graphic *h* is not pronounced in modern French, such spellings do not in themselves suggest that the MC word was formerly pronounced *lahas* rather than *laas*. However, in 1975, Stein recorded a man in Rodrigues who pronunciation of this word was [lahas] (Corne and Stein 1979:70). As the Mauritian dependency of Rodrigues was settled, from the end of the 18th century, almost exclusively by speakers of MC (Baker 1982a:207-208, 1982b:857-858), and as Rodrigues did not receive any Indian indentured labourers during the 19th century, the Creole spoken there must be essentially old MC. The clear implication is that the pronunciation [lahas] was probably formerly current in Mauritius too. If so, MB *lahās* would be the regular reflex of old MC *lahas* (cf. the correspondences between MC and MB described earlier), suggesting that this word was adopted by speakers of MB in the 19th century.

The possible exception concerning agglutinated French articles alluded to earlier is MB *butik shop*, modern MC *labutik* (<Fr *la boutique*). In 19th century MC texts, however, this word is attested both with (Baissac 1880) and without (Chrestien 1822) an agglutinated article. Speakers of MB may thus have adopted the non-agglutinated form at a time when this was current in MC (rather than adopting this from French). There is one other MB word in our data which at first looks to be direct from French rather than through MC — *libar pound* (*weight*) (<Fr *livre*) for which the corresponding modern MC word is *liv*. Clearly the final syllable of the MB form must derive from a French or MC pronunciation of the term in which a final *r* was audible. Although MC today lacks any word-final consonant clusters in which the last element is *r*, we cannot be entirely sure that this was always the case. For example, the French verb *suivre follow* is found in five different texts dating from between 1818 and 1888 as '*sivre*' or '*sivré*'. (The modern form of this verb is, among different groups of speakers, either /swiv/ or /sivre/). There are also a few attestations of words which in French have a final consonant cluster of which *l* is the final phoneme, being written in MC texts with *r* instead of *l*, for example *diabre* for French *diable* (Chrestien 1822:22). It seems most unlikely that such a change would have been made if MC *r* did not have some phonetic reality in this position. While the above in no way proves that MB *libar* has MC rather than French as its immediate source, the possibility that it might be from MC cannot be excluded.

Thus far, our examples have concerned apparently irregular sound correspondences between MC and MB but we will conclude with a semantic distinction, now obsolete in MC but attested in 19th century MC texts, which is very much current in MB. In MC texts from the first half of the 19th century there are two verbs which might be glossed *to be tired* but one, *las/e*, might be better defined as *to be*

fed up with, bored with while the other *fatig/e*, really means to be *physically tired*. This distinction is clear from the first attestation of each:

Moi lassé coudre ton cimize (Chrestien (1822) 1831:9)
I'm tired of darning your shirt

... porté-li dans Port comment çoçon, ... Pour li n'a pas trop fatigué,
 Et moi croire' li va bien vendé (Chrestien 1831:39)
Carry it [a donkey] to town like a pig... So it won't get tired, and I think it should sell well.

In modern MC, *las/e* has become entirely obsolete and *fatig/e* is now employed in both senses. In MB, however, both *lāse* and *phātige* are current and many speakers preserve the semantic distinction between them found in Chrestien's MC publications:²⁰

baïṭhal baïṭhal lāse hogāili
I'm tired of sitting (i.e. bored)

calte calte phātige hogāili
I'm tired of walking (i.e. physically tired)

CONCLUDING REMARKS

In the preceding pages we have examined only a few aspects of MB but we believe that these are sufficient to demonstrate that MB is not a radically restructured form of IB and has certainly not undergone anything comparable to the process of creolisation. We strongly suspect that a more detailed study would show that, in morphology and syntax, MB was a continuation of a range of IB dialects and related speech forms of north-east India from which dialectal differences had gradually been eroded and from which some inflected forms relating to social conditions no longer observed in Mauritius had been eliminated. (Some examples which support this view are to be found in Domingue 1981). With regard to lexicon, however, MC has undoubtedly had a very major impact on MB providing the latter with a reservoir of terms to draw on for virtually all facets of life in Mauritius encountered by MB-speakers which were unknown in the Bihar their ancestors left in the 19th century.

If the view of MB expressed in the preceding paragraph is essentially correct, then the widespread (in Mauritius) notion that the relationship between MB and Hindi is broadly comparable to that between MC and French is erroneous. MC is very definitely *not* a continuation of a range of French dialects but an independent language which originated in Mauritius out of a multilingual situation in the 18th century (Baker 1982b:806-860).

A comparison of Mauritian Bhojpuri (MB), Indian Bhojpuri (IB)
and Hindustani using the 100 basic words of the Swadesh list

No.	ENGLISH	MB	IB ^{2 1}	HINDUSTANI ^{2 2}
1.	<i>all</i>	sab	<i>T</i> sab	sab
2.	<i>ash</i>	rākhi	<i>T</i> rākh	rākh
3.	<i>bark</i>	lāpo	?	post
4.	<i>belly</i>	peṭ	<i>T</i> peṭ	peṭ
5.	<i>big</i>	baṛā	<i>T</i> baṛā	baṛā
6.	<i>bird</i>	ciraiñ	<i>T</i> cirāī	ciṛiyā
7.	<i>bite (v)</i>	kāṭ-	<i>T</i> kāṭ- <i>cut</i> ^{2 3}	kāṭ-
8.	<i>black</i>	kariā	<i>T</i> kariā	(kariyā), kālā
9.	<i>blood</i>	khun	?	(khun), lohu
10.	<i>bone</i>	haḍi	?	haḍḍi
11.	<i>breasts</i>	dudh, ^{2 4} (chāti)	<i>J</i> chāti	chātī
12.	<i>burn (v)</i>	jarā-	<i>H</i> jar-	(jār-), jalā-
13.	<i>claw (n)</i>	nokh	<i>J</i> nokh [<i>finger</i>]nail	(nakh), nākhun
14.	<i>cloud</i>	badār	?	(bādal), abr
15.	<i>cold (a)</i>	ṭhandā	<i>J</i> ṭhandhā	ṭhandhā
16.	<i>come</i>	āw-	<i>T</i> āwē	ā-
17.	<i>die (v)</i>	mar-	<i>T</i> mar-	mar-
18.	<i>dog</i>	kutā	<i>T</i> kukur	kuttā, (kukur)
19.	<i>drink (v)</i>	pi-	<i>T</i> pī-	pī
20.	<i>dry (a)</i>	sukhal	<i>T</i> sukhal <i>became dry</i>	sūkhā
21.	<i>ear</i>	kān	<i>T</i> kān	kān
22.	<i>earth</i>	māṭi	<i>T</i> māṭi	(māṭī), miṭṭī
23.	<i>eat</i>	khā-	<i>T</i> khā-	khā-
24.	<i>egg</i>	aṇḍā	?	aṇḍā
25.	<i>eye</i>	āñkh	<i>T</i> āñkh	āñkh
26.	<i>feather</i>	pāñkh	<i>T</i> pāñkhi	(pankh), par
27.	<i>fire</i>	āg	<i>H</i> agār, <i>D</i> āgi	āg
28.	<i>fish</i>	machi	<i>J</i> machri	(machī), machī
29.	<i>flesh</i>	gos	?	goṭ
30.	<i>fly (v)</i>	ur-	<i>T</i> uṛ-	uṛ-
31.	<i>foot</i>	goṛ	<i>T</i> goṛ	(goṛ), pāñw
32.	<i>full</i>	bharal	<i>T</i> bhar <i>to fill</i>	bharā

No.	ENGLISH	MB	IB	HINDUSTANI
33.	<i>give</i>	de-	<i>T</i> de-	de-
34.	<i>good</i>	achā	<i>D</i> acchā	acchā
35.	<i>grease (n)</i>	carbi	<i>T</i> carabi	carbī
36.	<i>green</i>	hariar	<i>J</i> hariar	(hariyā), harā
37.	<i>hair</i>	bār	<i>T</i> bār	bāl
38.	<i>hand</i>	hānt	<i>J</i> hānth/hāth	hāth
39.	<i>head</i>	kapār	<i>J</i> kapār <i>forehead</i>	(kapār), sar
40.	<i>hear</i>	sun-	<i>T</i> sun-	sun-
41.	<i>heart</i>	dil	?	dil
42.	<i>horn</i>	sing	<i>T</i> siṅgi	sīng
43.	<i>hot</i>	garam	<i>T</i> garam	garam
44.	<i>I</i>	ham	<i>T</i> ham	(ham), maiñ
45.	<i>kill</i>	mwāw-	<i>T</i> muāw- ²⁵	mār-ḡālna ²⁶
46.	<i>knee</i>	thewni	<i>J</i> ṭhehuna	(theūnā), ghūṭnā
47.	<i>know</i>	jān-	<i>T</i> jān-	jān-
48.	<i>leaf</i>	patā	<i>T</i> pattā	pattā
49.	<i>lie (down)</i>	leth-	?	leṭ-
50.	<i>liver</i>	karijā	?	(kalījā), jigar
51.	<i>long</i>	lambā	<i>T</i> lambā	lambā
52.	<i>louse</i>	ḡhil	?	(ḡhil), jūñ
53.	<i>man</i>	marad, admi	<i>T</i> marad, <i>J</i> admi	(mard), ādmī
54.	<i>many</i>	bahut	<i>T</i> bahut	bahut
55.	<i>moon</i>	cān	<i>T</i> cān	cānd
56.	<i>mountain</i>	motaiñ (pahār)	<i>T</i> pahār	pahār
57.	<i>mouth</i>	muñ	<i>T</i> muñh	muñh
58.	<i>name</i>	nām	<i>T</i> nām	nām
59.	<i>neck</i>	galā (gardan)	<i>T</i> galā <i>throat</i>	(galā) gardan
60.	<i>new</i>	nañwa	?	nayā
61.	<i>night</i>	rāt	<i>T</i> rāti	rāt
62.	<i>nose</i>	nāk	<i>T</i> nāk	nāk
63.	<i>not</i>	naiñ	<i>T</i> nāhiñ	nahiñ
64.	<i>one</i>	ek, ego ²⁷	<i>T</i> ek, <i>T</i> ego ²⁷	ek
65.	<i>path</i>	simā	<i>J</i> baṭ	bāṭ
66.	<i>person</i>	admi	<i>T</i> admi	ādmī
67.	<i>rain</i>	pāni	<i>T</i> pāni	(pānī), meñh

No.	ENGLISH	MB	IB	HINDUSTANI
68.	red	lāl	<i>T</i> lāl	lāl
69.	root	jaṛi	<i>T</i> jari	jaṛ
70.	round	gol	<i>T</i> gol-gāl <i>round-making</i>	gol
71.	sand	bālu	<i>T</i> bālu	(bālū), ret
72.	say	bol-	<i>T</i> bol-	bol-
73.	see	dekh-	<i>T</i> dekh-	dekh-
74.	seed	biā	<i>T</i> bia-han <i>seed-corn</i>	(biyā), bīj
75.	sit	baiṭh-	<i>T</i> baiṭh-	baiṭh-
76.	skin	lāpo, (camri)	<i>T</i> camṛā	camr-ā or -ī
77.	sleep (<i>v</i>)	sut-	<i>T</i> sut	(sūt-), so-
78.	small	choṭa	<i>T</i> choṭ	choṭā
79.	smoke (<i>n</i>)	dhuwān	<i>T</i> dhuān	dhūnān
80.	stand (<i>v</i>)	kharāho- ²⁸		kharā-ho- ²⁸
81.	star	terengan	<i>J</i> tairgan	taraī, (tārā)
82.	stone	ros	<i>T</i> pathal	pathar
83.	sun	ghamān, (suruj)	<i>T</i> ghām <i>heat of the sun</i> , <i>T</i> suruj	- ²⁹ , sūraj
84.	swim (<i>v</i>)	nahā-	<i>T</i> nahā- <i>to bathe</i>	(nahā-), ³⁰ pair-
85.	tail	poñchi	<i>J</i> poiñch	(ponc), dum
86.	that	haw	<i>J</i> ohe	wuh
87.	this	hai	<i>T</i> haī	yīh
88.	thou	tu	<i>T</i> tu	tū
89.	tongue	jibh	<i>J</i> jibh	jībh
90.	tooth	dānt	<i>T</i> dānt	dānt
91.	tree	gāchi, gāñchi	<i>J</i> gāch, gāñch	(gāch), darak <u>ht</u>
92.	two	du, ³¹ dugo	<i>T</i> du, ³¹ dui	(du), do
93.	walk (<i>v</i>)	cal-	<i>T</i> cal-	cal-
94.	water	pāni	<i>T</i> pāni	pānī
95.	we	ham, hamni	<i>T</i> ham, hamāni	ham
96.	what	konchi	<i>T</i> kaun	kaun
97.	white	ujar	<i>T</i> ujar	(ujjal), safed
98.	who	ke, kon	<i>T</i> ke, kaun	kaun
99.	woman	awrat	<i>D</i> aurati	'aurat
100.	yellow	piar	<i>J</i> piar	(pilā), zard

NOTES

- ¹MB is also the usual home language of some people living in rural Mauritius whose ancestors originally came from western or southern India. In addition, Bhojpuri is spoken as a second or additional language by some people of Chinese or part African descent who live in predominantly MB-speaking areas.
- ²Baker has worked on Mauritian Creole (MC) since the late 1960s and this has led to an interest in the influence of MC on MB and vice versa. Ramnah is a native speaker of MB who has advised Baker on a number of matters relating to MB at various times since 1972. Ramnah has also spent three years as a student in India and is thus familiar with modern spoken Hindi.
- ³The authors wish to thank all those who commented on earlier drafts of this paper or who provided additional data for it, especially Richard Barz, Chris Corne, Theo Damsteegt and Anand Syea.
- ⁴Hindustani, Hindi and Urdu are names for different styles of speech but are grammatically substantially identical. We employ the word Hindustani throughout the remainder of this paper, for several reasons. First, because we use for the purposes of lexical comparison *A dictionary of the Hindustani language* compiled at the time when Indian immigration in Mauritius was at its height (Forbes 1859). Secondly, because Hindustani is free of the religious connotations of Hindi (Hinduism) and Urdu (Islam); MB being spoken natively by both Hindus and Muslims. Thirdly, because in Mauritius, as in India, Hindustani is associated with films and popular culture generally, and it is this style of speech, rather than that of Hindi or Urdu which are now taught formally in certain Mauritian schools, which has thus far had a greater influence on MB.
- ⁵Provided one overlooks the fact that she treats aspirated stops as 'monophonemic consonant clusters' (i.e. as sequences of plosive + /h/), the differences are that w is considered an allophone of /u/ and that ɾ and ɾh are classed as allophones of, respectively, ɖ and ɖh.
- ⁶Tiwari does not identify phonemes as such but does give examples of minimal pairs suggesting that some speakers may accord phonemic status to some of the following: h, mh, nh, ɾ, ɳ, ɳh, lh, rh.
- ⁷Also pronounced phar by some MB speakers.
- ⁸Our research has been conducted mainly with people living in Flacq district.
- ⁹Also pronounced lib. As will be discussed in section 5, the v of MC words adopted into MB is often variably pronounced as v or b, even in the speech of the same informant.
- ¹⁰The marked forms awratīā and awratian are those which would be expected if the unmarked MB form were *awratī. Note that auratī is attested in IB (see word list) and that the loss of final -i is reported to be a feature of western dialects of IB (see section 4 below).
- ¹¹The word gilās, though of ultimate English origin (glass) may well have been established in IB before the massive immigration of Indians in Mauritius in the 19th century began, cf. Bihari gilās 'It is modelled closely on the lines of a European peg-tumbler, but is of metal' (Grierson (1885) 1928:131).
- ¹²All the quantitative nouns in these MB examples are of ultimate French origin through MC.

- ¹³ The forms set out in this table are a revised and amended version of a list drawn up by Suchita Ramdin and her colleagues at the Mahatma Gandhi Institute (Mauritius) in consultation with Baker in April 1983.
- ¹⁴ However, as Domingue (1971:35) rightly notes, some of the MB 'ordinary' forms correspond to what, in IB, are 'contemptuous' forms.
- ¹⁵ MB also has IB-derived terms for each of these five items but the latter are less frequent than the corresponding MC-derived terms included in the comparative word list. With regard to no.56, it appears that *motaiñ* is generally applied only to the hills and mountains of Mauritius and that IB-derived *pahār* is applied to other mountains (seen in films and photographs in magazines, etc.).
- ¹⁶ Particularly striking are no.38 *hānt* (cf. Eng hand), no.58 *nām* (cf. Eng name, Fr nom), and no.90 *dānt* (cf. Fr dent).
- ¹⁷ A clear idea of the extent of MB borrowing from MC will emerge when the dictionary of MB currently being compiled at the Mahatma Gandhi Institute, under the direction of Ms Suchita Ramdin, has been published.
- ¹⁸ The presence of Indian traders in Mauritius dates from the mid-18th century. Such people are thought to have come mainly from the west of India. MC terms for goods imported by these traders may well have their source in languages such as Gujerati, Konkani, etc.
- ¹⁹ The B/H orthography employs vowel + *ñ* or *ṁ*, according to etymological criteria, to represent nasalised vowels in MC. The basic aim is to preserve a constant sequence of roman letters for morphemes in which nasalised vowels alternate with the corresponding oral vowel + nasal consonant, as e.g. *noṁ* [nõ] *name* (noun) and *nom*, *nome* [nom], [nome] *name* (verb). This convention means that wherever a bilabial plosive immediately follows a nasalised vowel in MC (within the same word), the latter is represented by vowel symbol + *ṁ*. In MB, sequences of nasalised vowel + bilabial plosive occur almost exclusively in words adopted from MC. Indeed, we are currently aware of only one exception (though there may well be a few others), *sāmp* *snake* (cf. Hindustani *samp snake*).
- ²⁰ Virtually all of our informants were under the impression that *lāse* was a word of Indian origin, in contrast to *phātige* which they knew to be of MC origin. This is evidence of how firmly established *lāse* is in MB. However, one of our younger informants claims that it is mainly older people who employ this word and that he personally uses *phātige* in both the senses illustrated in these examples.
- ²¹ IB forms are taken from Tiwari 1960 (*T*), Jordan-Horstmann 1969 (*J*), Hertig-Skalická 1974 (*H*) or are those reported by Damsteegt (*D*; letter of 8 May 1985, citing publications not available to us), as indicated by these abbreviations. As vowel length is not phonemic with respect to *ī*, *ē*, *ō* and *ū*, the macron has been removed from these letters in citing forms from Tiwari in order to make them more readily comparable with those from other sources.
- ²² All forms are taken from Forbes 1859 and, unless enclosed between brackets, are the first Hindustani word he gives as equivalent to the English word in the Swadesh list. (Forbes gives several Hindustani words in translation of each English term in the English-Hindustani part of his dictionary. As these are not listed in alphabetical order, it has been assumed that they are in order of apparent frequency.) Forbes' *ch* and *chh* have been systematically

replaced with c and ch respectively in line with modern practice. His use of a comma to indicate that e.g. both e and u are to be pronounced separately in the, $\bar{u}\bar{n}\bar{a}$ has been eliminated. The infinitival termination $-n\bar{a}$ has been deleted from all his verbs to facilitate comparison with MB and IB forms.

²³As $k\bar{a}\bar{t}$ - means *cut* and *bite* in both MB and Hindustani, it seems likely that it also has the meaning *bite* in IB.

²⁴dudh (IB dudh) is the usual word for *milk* but is also far more frequent than *chāti* in the sense of *breasts*.

²⁵ $\mu\bar{a}\bar{w}$ is the stem of the double causative form of *mar-* to die (cf. no.17).

²⁶The first element means *killing*. The second element is a verb to *throw down* which, in contrast to the usual policy, is cited here with its infinitival suffix $-n\bar{a}$.

²⁷As indicated in the text, *ego* is the fusion of two morphemes. There is no corresponding form or forms in Hindustani.

²⁸The Hindustani forms consists of an adjectival element $khar\bar{a}$ *erect* and the stem of the verb to *be*, *ho-*. MB informants felt $khar\bar{a}ho-$ to be a single word but, even if this is so, it is clear that it derives from the same two elements.

²⁹Hindustani has a related verb, $gham\bar{a}n\bar{a}$ to *bask in the sun*.

³⁰Glossed as to *bathe*, to *wash* by Forbes, it apparently does not have the sense of *swim*. In view of this, Tiwari's gloss of *bathe* may also mean *get washed* rather than *swim*. In MB, however, $nah\bar{a}$ means both *swim* and *get washed*.

³¹As in the case of 'one' (see note 27), *dugo* is really two morphemes but occurs with far greater frequency than *du* alone. For IB, Tiwari indicates *dui* as the usual form and *du* as a dialectal variant. He does not indicate what the combined form with the classifier $-go$ is (*duigo* or *dugo*). Barz (p.c.) confirms that *duigo* is current but adds that, while he has not personally heard **dugo*, this might well occur in some dialects.

BIBLIOGRAPHY

ANDERSON, Samuel

1885 *L'Évangil selon S. Matthié (dan langaz créol Maurice)*. London: British and Foreign Bible Society.

BAISSAC, Charles

(1880) *Étude sur le patois créole mauricien*.

1976 Geneva: Slatkine Reprints.

(1888) *Le Folk-lore de l'Ile Maurice*.

1967 Paris: Maisonneuve et Larose.

BAKER, Philip

1982a On the origins of the first Mauritians and of the Creole language of their descendants: a refutation of Chaudenson's 'Bourbonnais' theory. In Baker and Corne 1982:131-259.

BAKER, Philip

- 1982b The contribution of non-Francophone immigrants to the lexicon of Mauritian Creole. Ph.D. thesis, University of London, School of Oriental and African Studies.

BAKER, Philip and Chris CORNE

- 1982 *Isle de France Creole: affinities and origins*. Ann Arbor: Karoma.

BAKER, Philip and Vinesh Y. HOOKOOMSING

- 1983 *Lortograf-linite*. Mauritius.

BARZ, Richard K.

- 1980 The cultural significance of Hindi in Mauritius. *South Asia* 3/1:1-13.

BHUCKORY, Somdath

- 1967 *Hindi in Mauritius*. Port Louis.

BICKERTON, Derek

- 1977 Pidginization and creolization: language acquisition and language universals. In Albert Valdman, ed. *Pidgin and creole linguistics*, 49-69. Bloomington and London: Indiana University Press.

- 1981 *Roots of language*. Ann Arbor: Karoma.

CHAUDENSON, Robert

- 1979 *Les créoles français*. Paris: Nathan.

CHRESTIEN, François

- (1822) *Les essais d'un bobre africain*.
1831 Port Louis.

CORNE, Chris and Peter STEIN

- 1979 Pour une étude du créole rodriguais. *Etudes créoles* 2:58-84.

DESCROIZILLES, Henri Charles

- 1867 *Navire fine engagé or the Mauritius in danger*. Mauritius.

DOMINGUE, Nicole Z.

- 1971 Bhojpuri and Creole in Mauritius: a study of linguistic interference and its consequences in regard to synchronic variation and language change. Ph.D. thesis, University of Texas.

- 1981 Internal change in a transplanted language. In *Studies in the Linguistic Sciences* 11/2:151-159.

FORBES, Duncan

- 1859 *A dictionary of the Hindustani language*. London: Wm. H. Allen and Co.

GRIERSON, George A.

- 1883 *Seven grammars of the dialects and subdialects of the Biharī language*. Calcutta.

GRIERSON, George A.

(1885) *Bihari peasant life.*

1926 Patna: Government Printing, Bihar and Orissa.

HERTIG-SKALICKÁ, Jita

1974 *Fifty Bhojpuri folksongs from Ballia District.*

Basle: Aku-Fotodruck.

HOLM, John

f/c *Creole languages [provisional title].* Cambridge: Cambridge University Press.

JORDAN-HORSTMANN, Monika

1969 *Sadani. A Bhojpuri dialect spoken in Chotanagpur.* Wiesbaden: Otto Harrassowitz.

STEIN, Peter

1982 *Connaissance et emploi des langues à l'Ile Maurice.* Hamburg: Buske.

TIWARI, Udai Narain

1960 *The origin and development of Bhojpuri.* Calcutta: The Asiatic Society.