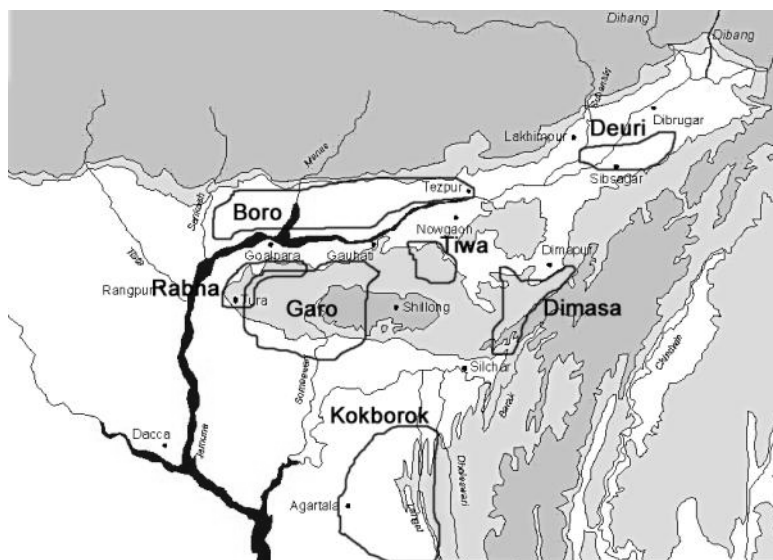


## Discovering Boro-Garo

### History of an analytical and descriptive linguistic category

François Jacquesson

This paper does not require professional linguistic skills on the part of the reader.<sup>1</sup> It is about the history of research on so-called “Boro-Garo” languages, how this started and proceeded; it emphasizes the difficulties in defining human groups and describes some thoughts involved in the pursuit of such definitions.<sup>2</sup>



1. Major Boro-Garo languages, a sketch of the present-day distribution.

#### 1. Boro-Garo, introduction

Using various names, Bodo-Garo, Boro-Garo, Bodo-Koch, or even simply Boro, social anthropologists and linguists define a group of “closely

---

<sup>1</sup> I am delighted to acknowledge the help of Bernadette Sellers, who transformed my erratic speech into decent English; and the stern reluctance of Pascale Dollfus to consider all my adverbs necessary.

<sup>2</sup> For an excellent general book on the history of Tibeto-Burmese Linguistics, up until 1980, see Hale 1982.

related” languages spoken in North-East India.<sup>3</sup> Locally nobody uses terms such as “Boro-Garo”: they are academic coinages, with (in principle) no political consequence. As far as the number of speakers is concerned, it is the most important group of “tribal” languages in the region. It has been identified as forming one consistent group rather early on, under somewhat interesting circumstances. We will examine how this came about.

Eight of these languages are identified in the official 2001 Census of India. The numbers for Bangladesh are of course not given: they are significant for “Garo” and for “Tripuri” (Kokborok). The names in the left-hand column below are those of the Census; some of these are highly debatable.

	in the 7 states	in W Bengal	elsewhere in India	total in India
Boro	1311348	37654	1476	1350478
Garo	887060	1457	962	889479
Tripuri	853196	98	729	854023
Rabha	153714	10967	89	164770
Dimasa	111878	4	79	111961
Koch	29299	1583	237	31119
Deori	27897	6	57	27960
Lalung	27067	0	5	27072

There were, and are, two different practices regarding names and communities. One is the “approach”, for instance, when you discover a country: proper names are given to you, and you have to look for their meaning. There is no proposed hierarchical processing; the categories are not exclusive. You slowly discover ambiguities and homonyms or quasi-homonyms. The leading metaphor is generally the map.

The other practice involves the “census”, when you have to provide a comprehensive picture. Categories are strict, you cannot use two names for the same notion, and hierarchical processing is a must; any ambiguity is forbidden. On the other hand, below a certain level of detail, categorization becomes useless or clashes with the prime purpose. The leading metaphor is usually the tree, and the method “branching”.

These two practices have been used from the very first descriptions, those of Buchanan-Hamilton, and will probably continue.

---

<sup>3</sup> North-East India encompasses 7 or 8 “states” within the Indian Union: Assam (which roughly corresponds to the valley of the middle course of the Brahmaputra River), Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura and Meghalaya. Sikkim is sometimes included.

## 2. From narrative to list, and from list to chart

### 2.1. 18<sup>th</sup> century and before

The earliest information we have from Assam comes in the form of maps and narratives. Maps prior to the British ones<sup>4</sup> reveal practically nothing about Northeast India, which was hardly known at the time. The earliest information we have on such groups are from “Muslim” chronicles, generally written in Persian — the court language in India until British colonial times. The first one is the *Tabaqat-i-Nasiri*, by Mihraj us-Siraj, composed and compiled c. 1250.<sup>5</sup>

Muhammad Bakhtiyar, an ambitious general of the Afghan dynasty, conquered Bihar c. 1200. His patron Aybak, from Delhi, thought it wise to push him further east towards Bengal, against the Sena dynasty. In 1204, Muhammad Bakhtiyar established his capital in Gaur. From there, he was tempted to invade “Bhutan and Tibet” and went against Assam, called *Kamrud* (sic, with a “d”). The description of his disastrous campaign provides us with some information about the populations (Siraj 1881: 560-1):

In the different parts of those mountains which lie between Tibbat and the country of Lakhanawati are three races of peoples, one called the Kūnch [N6<sup>6</sup>], the second the Mej (Meg), and the third the Tihārū; and all have Turk countenances. They have a different idiom, too, between the languages of Hind and Turk [N7<sup>7</sup>]. One of the chiefs of the tribes of Kūnch and Mej, whom they were wont to call “Alī, the Mej, fell into the hands of Muhammad-i-Bakht-yâr, the Khalj, and, at his hand also, the former adopted the Muhammadan faith. He agreed to conduct Muhammad-i-Bakht-yâr into those hills, and act as a guide; and he brought the latter to a place where there is a city, the name of which is Burdhan [kot] [N8<sup>8</sup>]

---

<sup>4</sup> The earliest British map for this part of India, Rennell’s, in 1780, goes as far as the border with Assam and includes Goalpara and an area some miles further east. It was published again recently in Deloche 1984. Sketches were drawn during the 1792-1794 British expedition; they have never been published. Rennell’s map gives no information about names of human groups or languages. The 1794 maps give some names.

<sup>5</sup> This important work was translated by Raverty and published in 1881. The Persian text had been published in 1864.

<sup>6</sup> Raverty’s Note 6. In some copies the nasal *n* is left out - Kûch.

<sup>7</sup> Raverty’s note 7. In some of the more modern copies of the text, “Hind and Tibbat”.

<sup>8</sup> Raverty’s note 8. The oldest and best copies generally contain the above, but two add *kot* and one copy gives the vowel points. The *Zobdat-ut-Tawârîkh* also has Burdhan twice. The other compiled copies have Murdhan and Murdhan-kot, and the printed text, in a note, had Durdhan [Wurdhan ?] as well as Burdhan?

Konch, sometimes written Koch, (the same hesitation occurs in Buchanan-Hamilton's manuscripts), is what we today write as Koch. Mej or Meg is the name we write as Mech. We can safely conclude that these names described important groups of people in the 13<sup>th</sup> century, in the area between the Ganges and the Brahmaputra. The relation with Buchanan's "Koch" and "Mech", is obvious, but the kind of entity (ethnical, political etc. ?) implied is not at all obvious.

### 2.2 Buchanan-Hamilton: listing and description

Francis Buchanan<sup>9</sup> (1762-1829) came to India in 1794, as Assistant-Surgeon and with a taste for ichthyology. Most of his time was taken up with special missions and surveys: he went to Ava (Burma) with Capt. Symes in 1795, surveyed Chittagong in 1798 and travelled in southern India, then to Nepal with Capt. Knox in 1802-3. His greatest accomplishment is the survey of Bengal (1807-1814), to which he added a wealth of information about Assam. After that, for one year he took charge of the Botanical Gardens in Calcutta, which he handed over to Wallich (23<sup>rd</sup> Feb. 1815), leaving India forever on the very same day. He then assumed the name Hamilton.

Francis Hamilton left a hoard of manuscripts in the India Office, the complete list of which can be found in Kaye & Johnston 1937. His descriptions are at first geographical, giving *zila* (district) after *zila*.<sup>10</sup> For each *zila*, the description follows the same pattern and a comparative vocabulary concludes each description. His descriptions and lists concerning the Boro-Garo languages and populations are given in the Rangpur manuscripts. These vocabularies are compiled in a special volume.<sup>11</sup> Only a small part of these documents have been used or published. The first extensive use of Hamilton's work is in Martin 1838, *The history, antiquities, topography and statistics of Eastern India*; the 3<sup>rd</sup> volume, pp. 600-696, is about Assam.<sup>12</sup>

In his comparative vocabulary volume for Rangpur, Hamilton had hundreds of words copied (into both Bengali and Latin scripts, with great care) in several languages designated in this way (Rabha, Garo, Kachari, Pani Koch and Mech are Boro-Garo languages):

<sup>9</sup> See Kaye & Johnston 1937. A biographical note, with sources, is available p. 580.

<sup>10</sup> Dinajpur, Kaye no. 162; Ronngopur [sic, Rangpur], 163; Puraniya, 164; Bhagalpur, 165; Bihâr & Patna, 166; Shâhâbâd, 167; Gorakhpur, 168.

<sup>11</sup> Ms Eur.G.13 (Kaye no. 169).

<sup>12</sup> The Ms Eur.D.77 contains the *Account of Assam*, copied by S. K. Bhuyan, published by him in 1940 and repeatedly reprinted by the D.H.A.S (see Hamilton 1940), with an index.

English		
Prakṛita banggadēsīya bhāsā	Bengali	desi
Kōchārdēsīya bhāsā	Koch	desi
Rābhājātīya bhāsā	Rabha	jāti
Gārojātīya bhāsā	Garo	jāti
Kāchārījātīya bhāsā	Kachari	jāti
PāniKōch jātīya bhāsā	Pani Koch	jāti
Mechjātīya bhāsā	Mech	jāti
Āsāmdēsīya bhāsā	Asam	desi
Manipūrdēsīya bhāsā	Manipur	desi

The second case for instance, *Kōchārdēsīya bhāsā*, means “language of the country of the Koch”. Some languages (*bhāsā*) are described as characteristic of a country (*dēs*), others as characteristic of a human group (*jāt*). This distinction has a political basis: Bengal, As(s)am, Manipur and Koch were regions, since Manipur and Koch Bihar were then independent kingdoms. Therefore, *jāt* is a default term: those designated as *jāt* languages are those that do not have a political status.

Assam was very much a kingdom. Actually, Buchanan could not enter Assam, which was then closed to foreigners. All his information, as he himself explained, was collected from people he met in Rangpur or closer to the border. His approach came from outside. This explains why the lexicons he was able to collect were either from languages spoken in the Rangpur *zila*, or from languages in Assam but spoken close to the western border — except for Manipur, though the lexicon he compiled for meitei<sup>13</sup> is rather strange anyway.

Apart from this substantial lexicon, Buchanan-Hamilton also wrote (often excellent) descriptions.<sup>14</sup> For instance, he explained that the language of the Koch (country) is very much like Bengali, but that he had found a village where “Pani Koch” was spoken. He correctly assumed his “Pani Koch” language to have been the language of the Koch before most of them “deserted their ancient customs”, and he rightly remarked that this language is not like Garo but very much like Rabha. Buchanan-Hamilton was wiser than most linguists of his time, who still used the Leibnizian idea of tracking ancestry via language: Buchanan had noticed language borrowings and language shift.

Buchanan-Hamilton combined narrative description (based on his own field trips) and listing. Listing vocabularies in order to survey populations

<sup>13</sup> Meitei (or Mete) is the language of central Manipur: Manipur is the name of a country, Meitei of a language.

<sup>14</sup> Notably in Ms Eur.D.74.

was not a new idea: prior to this, Catherine the Great had had the same idea for her Russian empire, and she herself had participated in the venture.

What Buchanan-Hamilton did not do, was to classify the languages.

### 3. From Buchanan to the 1881 Census

#### 3.1. Nathan Brown, 1837

The first outstanding character in the colourful field of North-East Indian British anthropology-cum-linguistics was Nathan Brown. He was the first to publish lexicons of a dozen local Tibeto-Burmese languages, and the creator of local Tai studies.<sup>15</sup> In 1837, he gave two papers to the *Journal of the Asiatic Society of Bengal (JASB)* (Brown N. 1837a & b). One showed that Ahom, the historic Tai language in Assam, retained consonant clusters that had been lost in Shan. One should bear in mind that comparative linguistics as such were quite a new field.<sup>16</sup>

His second paper compared 60 words in 27 languages, among which was Garo. These were the first steps in Tibeto-Burmese fieldwork (the first works by Csoma about the Tibetan language were published in 1834 in Lahore). About Garo, he wrote:

It is difficult to decide from the specimens before us, whether it is to be ranked with the monosyllabic or polysyllabic languages. It probably belongs to the latter. The Garos inhabit an extensive range of hills below Gawahati, and are in a completely savage state. So meagre is their language, that they have not even a term for *horse*, nor do they possess any knowledge of such an animal.

While this concern over horses is indeed funny, the question about syllables is not. The polysyllabic character of Assamese was well-known, and considered typical of Western languages, while the monosyllabic feature was considered diagnostic of Eastern ones, such as Chinese or Thai. The border between the West and East could therefore be defined by linguistic experts, so examining Garo in this respect was meaningful. He computed percentages of a common lexicon between languages. For Garo, his only Boro-Garo language, he finds more correlation with Jili, Singpho, northern Tangkhul.

---

<sup>15</sup> One of the first books about Tai languages was James Low's *Grammar of the Thai or Siamese Language*, Calcutta, Baptist Mission Press, 1828. This book was probably well known to Nathan Brown.

<sup>16</sup> The first comparative essay by Franz Bopp, *Über das Conjugationssystem*, was published in 1816; the first version of his *Vergleichende Grammatik* was published in 1833.

Nathan Brown began comparing languages on a large scale. However, his concern with Boro-Garo was still restricted. The next phase was of course to feed the impulse with data.

### 3.2. *From 1840 to 1850: fighting with hierarchies*

In 1840, another *JASB* paper is Capt. Fisher's "Memoir of Sylhet, Kachar, and the adjacent Districts". This represents the southern point of view, since the British were actually more familiar with Bengal and Arakan. The paper is about the economy and agriculture, but ends with notes on the local people: the first information we are given about Dimasas and "Tipperas". He does not say anything about languages, but adds:

The people of Tippera are said to have the same origin as the Kacharis, and the similarity of religion, customs, and appearance, makes this probable. It may be added, that the Rajas of both countries have formerly acknowledged the connexion; the Tippera family being described as a younger branch of the ancient royal family, which in their expulsion from Kamrup established itself independently in the country which it formerly held as an appendage.

Family ties are similarly touched upon the following year, 1841, in the same *JASB*, with Lt Phayre's "An Account of Arakan", again a view from the south.

The people called Mrung, by the Arakanese, announce themselves as descendants of persons carried away from Tipperah several generations back by the Arakan kings. They were first planted on the Le-myoo river, with the view I suppose of cutting off their retreat to their own country; but when Arakan became convulsed in consequence of the invasions of the Burmese, they gradually commenced leaving the Le-myoo, and returning through the hills towards their own country. For a time they dwelt on the Kola-dan; now, none are to be found on any part of Arakan, save on the Mayu in its upper course, and only a few stragglers there. Many still reside, I understand, on the hills of the E. frontier of the Chittagong district. By a reference to a few words of their language, given in the appendix, those acquainted with the language of the Tipperah tribes will be able to decide whether the tale the Mrungs tell of their descent be true or not.

Phayre's guess is right, as far as the language is concerned: the Mrung lexicon is close to what we now call Kokborok (and was then dubbed Tipperah). This will be demonstrated by Lewin twenty-six years later, in 1867. The lexicon is here taken as evidence, through the native narrative, of their descent. This is one of the numerous details that show the slow but clear racialisation of language concerns, when we progress through the 19<sup>th</sup> century.

Hodgson saw his 200-page book, *Essay on the Kocch, Bodo, and Dhimal tribes*, 1847, as the first step in a complete description of all tribes of India.

The book contains lexicons (Hodgson 1847: 11-103), and sketch grammars of Bodo and Dhimal (105-140). His Koch is (he probably did not read Buchanan's work) what he calls "corrupt Bengali". He makes the first attempt at a definition of a Boro-Garo grouping (p. 151-2):

The Bodo are still a very numerous race (...) in the eastern marches from Gauhati to Sylhet, they are less numerous only than the Garos, Rabhas and Hajongs, not to mention, that the two last, if not all three, are but Bodos in disguise. I look upon the Rabha as merely the earliest and most complete converts to Hinduism, who have almost entirely abandoned the Bodo tongue and customs, and upon the Hajongs or Hojai Kacharis of Nowgong, as the next grade in time and degree of conversion, who now very generally affect a horror at being supposed confreres in speech or usages with the Bodo, though really such. Nor have I any doubt that the Garos are at least a most closely affiliated race, and no way connected with the monosyllabic-tongued tribes around them. I do not, however, at present include the Garos, or Rabhas, or Hajongs among the Bodo, who are now viewed as embracing only the Mechs of the west and the Kacharis of the east and south; and, so limited, this race numbers no less than 150,000 to 200,000 souls.

This is not very clear for people unfamiliar with the local names of people and places. As he admits in the end, his way of seeing the "Bodos" is twofold: he starts by using "Bodo" to designate a wide range of people ("a numerous race"), then wonders if some others are not "Bodos in disguise". He ends on a cautionary note and refrains from unmasking the dubious tribes, registering only the Mechs and Kacharis, which is indeed better from the present linguistic point of view. However, Hodgson does not cite any linguistic argument here. His approach cannot be taken as a classification, but rather traces more or less tightly linked circles with the Bodo as the centre:

centre	farther away	outside
Bodo	Rabha	Koch
Kachari	Hajong (=Hojai)	Lalong (=Lalung)
Mech	Garo	

Only in a footnote (p. 142 †) does he give more accurate sources:

Fifteen in sixty words of Brown's Vocabulary are the same in Garo and in Mecch, and the whole sixty or nearly so in Kachari and Mecch. Again, the Kacharis call *themselves* Bodo, and so do the Mecch; and lastly the Kachari deities Siju, Mairong and Agrang are likewise Mecch deities - the chief ones too of both people, to whom I restore their proper name. These are abundant proofs of common origin of Garos also.



Robinson knew Assam well, as his 1841 *Descriptive Account of Assam* shows. He had an informed and wise opinion about the respective position of Assamese and Bengali, the reduced influence of Tai languages, the importance of comparing not only lexicon but also grammar and about the importance of archaeological research for a better understanding of the past peopling of Assam. Robinson's position concerning the importance of writing tones and the existence of grammar even in Chinese are excellent. In his 1849 *JASB* paper, he gives a grammatical sketch of Garo:

The Garos have no traditionary legends whatever that may serve to enlighten us on the subject on their origin. Their remote situation, and their physical appearance, together with their modes and customs, so diverse from those of the Bhotias, would at first militate against the supposition that they were in any way connected with the Cis-Himalayan tribes. This connection however is now made apparent from the strong affinity existing between the language of the Garos and the several dialects spoken by those tribes. Though these present several modifications, they may nevertheless be traced to the same radicals, so as to prove that an essential affinity existed in their primitive structure, thus affording historical evidence of such a nature as it is impossible for either accident or design to have falsified.

The scope and orientation are of particular interest. He says that whatever the physical or social anthropology, language shows the link between Tibetans (Bhotias) and Garos. The same theme is emphasised regarding the Boros ("Kacharis or, as they term themselves, Borros", 215-223):

An examination into their language however furnishes abundant proof of their intimate connection with the tribes of the Cis-Himalayas. A large proportion of their vocables are identical with those of the Garos, and almost all the rest can be traced to some dialect of the Thibetan, while the idiom of the language and the peculiarities of its grammar show abundant traces of descent from a common origin. Closely connected with the Kacharis, among the inhabitants of the plains, are the Hojai Kacharis, the Kochis (including the Modai Kochis, the Phulgurias, and Hermias), the Mechis, and the Rabhas.

At the end of his paper, he gives about 250 words in five languages (Bhotia, Changlo, Garo, Kachari, Miri).

It is clear that Robinson, just as Hodgson, describes the "connection" at two different levels. The relationship he describes between Boro and Garo or "some dialect of the Thibetan" is described as an "intimate connection" and then interpreted as "traces of descent from a common origin". Yet, the arborescence metaphor is not explicit; his phrasing rather evokes a common pool. Then come the "closely connected" ones: Kachari, Hojai, Kochi, Mechi, Rabha. These latter ones obviously form a tighter unit. Yet,

the hierarchy is not explicit either: there is no overt scheme of levels of implication by “families” or “branches”.

In 1850, Nathan Brown published nine lexicons, four on Tai languages, four on Naga languages, and the first ever data about Deori (or Deori), which he called Chutia (“The Chutia is the language of one of the old tribes of Assam, now nearly extinct.”), opening a long debate - since the “nearly extinct” language is still very much alive. I re-edited Nathan Brown’s Deori lexicon in my book about the language (Jacquesson 2005). His paper does not indulge in any comparison, only in presenting data.

#### 3.4. *The wider field, 1866-1874*

The biological metaphor enters our sphere in 1866, in a special issue of the *JASB*, where Campbell published an *Ethnology of India*, which included notes about Boro-Garo people, plus a lexicon of the Mech language.<sup>17</sup>

The people of the very lowest hills of Bhootan and of all the low country at their foot are of another race, the Meches or Mechis (before alluded to in marking the boundaries of the Indian Aborigines), who are apparently the same as Hodgson’s “Bodo”. They are, it appears, now quite ascertained by their language to be Indo-Chinese of the Lohitic or Burmese branch of the Turanian family,<sup>18</sup> a connexion which their physiognomy confirms. They seem to be a good sized, fair, but rather yellow-looking people. They are described as rude in their agriculture (using the hoe, not the plough), and erratic in their habits, but good-natured and tolerably industrious. They profess a kind of debased Hinduism, but are very omnivorous in their habits.

Biological metaphors (“sister language”, “family of languages” etc.) had been in use since the late 18<sup>th</sup> century, but never systematically, nor did it convey any specific methodology. Things changed with the *vergleichende Grammatik*, and Bopp’s first preface to his famous book (1833) has a different ring. The same kind of assessment can be found in authors such as Max Müller, in the 1850s. Yet, the real transformation had to wait for August Schleicher in the 1860s, who harboured the idea that languages evolve like species and, like them, differ from each other. Schleicher’s teaching — although his main book, the *Compendium* (Schleicher 1866), was translated into English only in 1874 — or at least the spirit of it, spread far and wide very quickly.

---

<sup>17</sup> In Appendix B.

<sup>18</sup> The term “Turanian” was coined by Max Müller and became popular after the 2<sup>nd</sup> edition (1855) of his influential book, *The Languages of the seat of war in the East, with a survey of the three families of language, Semitic, Arian, and Turanian, with an appendix on the missionary alphabet, and an ethnographical map, drawn by Augustus Petermann*. The term “Turan” itself dates back, at least, to the *Shah Nameh* by Firdousi, where it described the Steppe dwellers, the traditional enemies of the Iranians.

Lewin published *The Hill Tracts of Chittagong and the dwellers therein* in 1869. It has several pages about the “Tipperah” (the Boroks) and the Mrungs. Lewin explains that Mrung is a name given to the Tipperah by the Arakanese people. Lewin also read Phayre, 1841, as mentioned above, compared Phayre’s data with his, and showed them to be the same language. This is the first demonstration of the wide geographical extension of Boro-Garo speaking populations: Lewin showed that the approach from Rangpur in Buchanan’s time, the approach from Cachar during the British advance from the plains of Bengal, and the approach from the Chittagong Hill Tracts eventually faced the same phenomenon.

In 1873, Captain Butler<sup>19</sup> published his “Rough Comparative Vocabulary of some of the dialects spoken in the “Naga Hills District”. The paper exhibits the same quality of data aimed at a future synthesis. He gives several hundreds of words in 7 languages (Assamese, Kachari, Mikir, Kuki, Angami naga, Rengma naga, Kutcha naga). His “Kachari”, the only Boro-Garo language in the group, still has the /əi/ diphthongs, a feature that most Dimasa dialects, and Haflong’s among them, does not have (it has /i/ for the same words). Since the same words in both Boro and Kokborok carry this diphthong, this dialect proves to be a link between Boro and Kokborok. I could demonstrate (Jacquesson 2006: 288) that this dialect still survives and is none other than what is now called Riang or Bru. The work of those first pioneers is still of great use.

### 3.5. Systematicians and the era of classification

Two *JRAS* essays lay the foundations of the classification of Boro-Garo languages. One is by E. J. Brandreth dated 1878, “On the non-Aryan languages of India”, the other by G. B. Damant in 1880, “Note on the locality and Population of the Tribes dwelling between the Brahmaputra and the Ningthee Rivers”. Sten Konow, when writing the Tibeto-Burmese volumes of the *Linguistic Survey of India*, writes that he started from Damant’s classification.

Many predecessors had not only gathered the material, but made successful comparisons and clever groupings. Much of the work had been done: (1) Mech, Hojai, Bodo or Boro, and Kachari from Cachar Hills are closely related languages - this is right. (2) Garo, Rabha, Koch and Tiperah are often cited as more or less close languages - right again. If the group itself, or its core, was quite clearly identified, the outer margin remained very vague, and its description fluctuated greatly from author to author.

---

<sup>19</sup> There are two John Butlers. The father is usually called Major Butler and the son (who died in January 1876 from a wound received in a Naga ambush) Captain Butler. Major Butler published two interesting books, *A Sketch of Assam*, 1847, and *Travels and Adventures in the Province of Assam*, 1855.

Brandreth and Damant were not restricted to Assam. Their purpose was wider, and they had to draw limits for each group they were concerned with. Their practice was to put together what looked<sup>20</sup> similar, or not, to a Boro core. They worked mainly with lexicons. Here are their classifications, with their original names for the languages. What they call Hojai is our Dimasa (“Purbutta Kachari” means *Hill Kachari*), and “Tipura” means what we now call Kokborok.

Brandreth 1878	Damant 1880
Kachari or Bodo	
Mech	Mech
Hojai	Hojai, or Purbutta Kachari
	Rabha
Garó	Garó
Pani-Koch	Koch
Deori-Chutia	Chutia
Tipura	? Tippurah

Brandreth neatly grouped Mech and Hojai (Dimasa) with Boro; this is less clear in Damant’s essay. Both were puzzled by the exact status of Tipura, which they both indicated at the end of the list: Brandreth decided that it was a Boro-Garó language (he is right) while Damant only suggested it.

The overall result is convincing, and the only important discoveries still pending during the British period concerned the Lalung/Tiwas and the Morans. The Lalungs were to arrive with a note by E. Stack in 1883. The Morans were to appear in extremis, in the paper by P. R. T. Gurdon in 1904, when only very few speakers were left.

The work by Brandreth and Damant concludes a period. Their comparative “technique”, actually very rough and informal, had been in use for some time: we saw how Nathan Brown already worked in such a way in the 1830s. The advantage in the 1870s was the larger amount of data, which was induced by colonization, and the subsequent curiosity of officers and administrators, enhanced by the possibility of publishing in two prestigious journals.

#### 4. Between the first Census and the Linguistic Survey

##### 4.1. *The 1881 Census*

The first British *Census* of 1881, the results of which were published in 1883, was an outstanding feat of administration and publication. Here are

---

<sup>20</sup> “Looked”, because they worked with printed data, not from a direct study.

the data, for numbers of speakers, concerning Boro-Garo groups, with the names used at that time. These numbers are to be compared with the 1,361,359 Assamese speakers, an incredibly small number (today more than ten times more), and 2,425,878 Bengali speakers. The general population of Assam was calculated as being 4,881,426, giving an average density of 104 hab/square mile; densities in the same period, also per sq. mile, are: Scotland 123, England 484, the United Kingdom 287.

Census 1881	Assam	Bengal	Burma	TOT
Cachari	263 186			263 186
Garo	112 248	24 949		137 197
Hajong	1 246			1 246
Koch		5 631		5 631
Lalung	46 920			46 920
Mech	57 890	11 101		68 991
Rabha	56 499			56 499
Tipperah	3 984	95	11	4 090
TOT	541 973	41 776	11	583 760

*Boro-Garo languages: number of speakers in 1881.*

#### 4.2. Handling categories with care

We have now reached the “Census period”, when labels have to be exclusive and non-ambiguous. This is no longer an “approach” (see Introduction), but a planned and systematic view. Therefore, the labels that have been selected for people and/or language are something like an official identity stamp. They are to stick for a long time.

Yet, census results for languages and for “castes” or “tribes” may differ widely, especially in the case of Koch, which the Report (p. 284) describes as “the remnant of an aboriginal tribe inhabiting the north-east of Bengal”.

	caste	language
Kachari	281 611	263 186
Koch	1 878 804	5 631

The Report commenting the Census is an interesting document. Officers in charge have their own *franc parler*. For the name “Kachari” (p. 291): “Under the term Kachari, 282,566 persons are shown. Of these 281,611 are recorded in Assam and 955 in Bombay. I doubt whether the designation is properly used to describe a caste. It appears to be a territorial designation for the inhabitants of Kachar in the Assam territory

(...)” Such doubts have many causes. An obvious one is the fact that “caste”, be it a “system” or not, was not easy to handle as a category. All the less so when you are attempting, as any good *census* should, to interpret each caste name in the same way throughout India.

As we saw with Hamilton, the British administration had a technique for understanding these complications. The general idea was that most castes were transformed aboriginals, who usually wished to be integrated into the Indian “system”. The price to pay for this integration was to become the shadow of an aboriginal. British officers, especially those working or travelling in the hills, were not happy with the Plains people, while they often admired the Savage and the Primitive in the Mountains - one model of which was “the Naga”.<sup>21</sup> Notwithstanding its obvious shortcomings, this view helped British officers to handle local designations with care: they knew that most names or categories in the North-East were those created by the Assamese or the Bengali clerks, and they handled this with some suspicion. This is one reason why descriptive anthropology was so prolific in North-East India: Western science had to know what was on the other side of the Indian curtain, had to “deconstruct” the Indian or Hindu approach, in order to reconstruct the Primitive reality.

A typical case of the Shadow Aboriginal was the Koch. Buchanan-Hamilton, in his description of the Pani Koch, already described them as a relict population, the last witnesses. He was largely right. In his book, *History of the Koch Kingdom 1515-1615*, published in 1989, D. Nath, in the footsteps of Gait who often agreed with Buchanan, tells how the term “Koch” had had for some time a Barbaric flavour which induced many people to prefer the designation “Rajbangsi” (or Rajvamsi), a term which means “of the royal clan”. In their steady 19<sup>th</sup> century process of integration into Hindu India, these people for some time tried to change their name. However, earlier 18<sup>th</sup> century sources in Assamese called them Koch or Mech, and we saw in Section 2.1. that these two names were already in use in the 13<sup>th</sup> century.

#### 4.3. About “Kachari” and “Chutiyas”

Sidney Endle’s 1884 *Outline of Kachari (Bârâ) Language*, was the first of these Outlines that became quite an institution.<sup>22</sup> In 1895, Anderson’s book of

<sup>21</sup> See the excellent book by Verrier Elwin, *The Nagas in the Nineteenth century*, 1969.

<sup>22</sup> The type was T.J. Keith, *Outline Grammar of the Garo Language*, published in 1874 in Sibsagar, on the Baptist Mission Press. But Endle’s volume was the first to be printed by the Assam Secretariat in Shillong. Several followed: Soppitt 1885 (Kaccha Naga), Needham 1886 (Shaiyang Miri), MacCabe 1887 (Angami Naga), Soppitt 1887 (Rangkhoh-Lushai), Witter 1888 (Lhota Naga), Needham 1889

Kachari texts, *A Collection of Kachári Folk-tales and Rhymes, intended as a Supplement to Rev. S. Endle's Kachári Grammar* built on that, and the same Anderson helped publishing Endle's ethnographic work in 1911, after Endle's death in 1907. The whole enterprise marks a first attempt at a monograph about people and language. Endle was clear about what is to be understood by "Kachari". He explains that we have Plains Kacharis, viz. the Bodos (or Boros), and the Hills Kacharis, viz. the Dimasas. He also knows that there are good historical, ethnological and linguistic reasons to group the Plains and the Hills people under the same label, "Kachari". Although his book, *The Kacharis* is mainly about the Boros, he gives a comprehensive view of the situation. He is also driven by a kind of love for the people described, which was not rare among such authors; this loving attitude does not preclude paternalism.<sup>23</sup>

Between them, these books present a rather modern outlook. We can certainly discuss their content, and disagree with some aspects. But on the whole, they offer a first model of what, during the 20<sup>th</sup> century, will be this kind of description. Anderson's lines in his introduction should be quoted here (Endle 1911: XVI):

Now, the anthropologists rightly caution us against rashly concluding that common speech, where races are in contact, implies a common origin, since everywhere, and especially among people who use an unwritten language, nothing is more common than the borrowing of a neighbouring tongue. But where, as here, we have five absolutely separate communities of semi-savage people, who nowadays are not so much as aware of one another's existence, and yet speak what is to all purposes the same language, it is plain that they must have been united at no very distant date by some common social bond.

Another attitude is illustrated by William B. Brown's book about the Deoris. Deori was the smallest Boro-Garo language, in terms of the number of speakers (c. 4000) at that time, while Boro-Kachari was the by far the biggest. This explains why the Deoris were "discovered" only in the 1840s. They were a rather discreet group, living mostly in Lakhimpur district, at least for those (the Dibongiyas) who still spoke the language.

---

(Singpho), Needham 1894 (Khamti), Brown 1895 (Deori-Chutiya), Hamilton 1900 (Dafla), Phillips 1904 (Garo), Barman 1908 (Dimasa).

<sup>23</sup> Anderson in his introduction (Endle 1911: XIII): "The Bodos (...) are, like most of the aboriginal races of Assam, cheery, good-natured, semi-savage folk; candid, simple, trustful, but incorrigibly disrespectful according to Indian notions of good manners."



2. The two “Chutiya” (i.e. Deuri) spots in the 1911 Endle map.

In the chronicles of Assam, either in the Tai-Ahom or Assamese languages, two kingdoms were important in 15<sup>th</sup> and 16<sup>th</sup> century Upper Assam. These two “peoples” were called Kachari and Chutiya in the Assamese language, and respectively Tumisa (or Timisa) and Tiora in the Tai-Ahom language. It was clear that a link existed between the Kachari-Timisa and the present-day Dimasas; such as to be Edward Gait’s position in his great *History of Assam*. The question was: what about the Chutiya-Tiora? Many were the people throughout Assam who considered themselves “Chutiyas”, especially in Upper Assam, but they were considered to be an Assamese “caste” since they were (and still are) quite indistinguishable from common Assamese people; actually, they were (and still are) one of those traditional groups of Assam that came to form the Assamese people. Maybe part of their ancestors were “tribal” at an earlier period, and probably the present-day Chutiyas are but another illustration of the accretion process that came to form the mainstream population.

William B. Brown did not “discover” the Deoris (a small lexicon was published by Nathan Brown in 1850), but he was the first to describe their language, in his small 1895 monograph, *Outline Grammar of the Deori Chutiya Language*. As the title of his book makes clear, he also followed the tradition (Nathan Brown’s, at least) claiming that the Deori language was actually the Chutiyas’ language, and therefore that he had unearthed, hidden in this small Deori tribe, the lost language of the old and famous Chutiya Kingdom (Brown 1895: III): “the original language of Upper Assam”.

Modern Chutiyas, who would be very pleased to be registered as a schedule tribe, have now and then used Brown’s book (or at least its title) as a political weapon. The Deoris, on the contrary, are not happy with this unfortunate misunderstanding, because they hope their smaller tribe will not be merged into the much larger Chutiya group. In my book about the



Deori language, I showed that the Deoris are right, since the features that have given their language its specific shift<sup>24</sup> show that it was shaped in the north-easternmost recess of Assam, close to the Dibang valley, where indeed according to traditional lore the Deoris came from, whereas the numerous Chutiyas have never been isolated in this small place, but were widespread throughout Upper Assam.

The point I want to illustrate here concerns the connexion W. B. Brown tried to make between linguistics and history. He was fascinated (as many British people were, and sometimes very knowingly) by the Antiquities of Assam. He was not the first, and not the last, to discuss at length the reputation the Deoris had had, until recently, of being responsible for human sacrifices in the temples of Upper Assam. He thought that such people could only have a very old language (a rather meaningless phrase, I am afraid), and therefore be related to a famous kingdom. It is of course quite biased to deduce that what is remote should be old, and what is savage should be remarkable. The Deori people are remarkable, but for quite different reasons.

When languages are classified, there is the temptation to use the device in order to classify people as well. Very often, the idea was to put them on a scale ranging from the hoary Savage to the most refined Civilized. Surprisingly, the Civilized is the one that describes the scale.

---

<sup>24</sup> For instance, the Deuri language has 5 nasalized vowels, a rare feature that also emerged in eastern Tani and some Mishmi languages. Deuri is the only Boro-Garo language that has nasalized vowels, and the contact with Tani and Mishmi could only occur in the Dibang valley region.

## 5. The Linguistic Survey, 1903

### 5.1. The LSI on Boro-Garo



3. "Languages & dialects of the Bârâ or Bodo group of the Tibeto-Burman family"<sup>25</sup>

The famous *Linguistic Survey of India (LSI)*, or rather the volume we are concerned with, III-2, published in 1903, benefited from most of the previous publications, and from systematic enquiries in the field. The work was carefully planned, carried out, and published. At first sight, it is a pure product of the engineering age which also produced the Surveys. On second thoughts, it is something rather different. Although, regarding most points, it is now outdated, it does deserve its ongoing reputation, would it not be for the infelicitous consequence that some scholars still uncritically copy it. The Konow-Anderson classification is:

Bârâ, Bodo, or Plains Kachari  
 Mes or Mech  
 Lalung  
 Dimasa or Hills Kachari

<sup>25</sup> *Linguistic Survey of India*, III-2. This extract, with the north-western part of the map, illustrates the westernmost extension of Boro dialects.

Hojai  
 Garo  
   Achik or standard dialect  
   Abeng  
   Atong, Kuchu or Ating  
   Garo of Cooch Behar and Jalpaiguri  
   Other dialects  
   Koch dialects  
   Tintekia of Goalpara  
 Rabha  
 Tipura  
 Chutiya  
 Moran

The number of speakers is added:

	Assam	Bengal	Total
True Bârâ (Kachâri & Mech)	247 520	25 011	272 531
Râbhâ	31 370	0	31 370
Lâlung	40 160	0	40 160
Dimâ-sâ (or Hill Kachâris)	18 681	0	18 681
Gâro (or Mânde)	120 780	28 313	149 093
Tipurâ	300	105 550	105 850
Chutiya	304	0	304
	459 115	158 874	617 989

Konow wrote (p. 2):

The philological interest of this group of languages consists largely in the fact that they are agglutinative tongues which have learned inflexion by coming into contact with the speech of Aryan peoples. Thus, a Boro living in Darrang can talk, not only Assamese and a rich idiomatic Boro, made picturesque and vivid by the use of polysyllabic agglutinative verbs, but also an Aryanised Boro which freely borrows the linguistic artifices of Aryan tongues, such as the use of the relative clause, of the passive voice, of adverbs, etc, and which almost wholly abjures the characteristic agglutinative verb that does the work of these more analytic devices of language.

Were I an old-fashioned guru with disciples studying Boro-Garo languages, I would first order them to learn the above quotation, and to ponder each sentence. After an abridged presentation of the grammatical features, his introduction gives a small comparative lexicon, in order to help the reader to grasp the consistency of the group of languages. The volume also contains an excellent map — I am sorry to say that since then

no other linguistic map of this quality, for Boro-Garo languages, has ever been published.

The *LSI* lexicons give 241 entries for 15 languages or dialects. Some of them are borrowed from previous (and duly acknowledged) publications, others result from specific questionnaires.

	water	fire	sun	moon
Bârâ or Plains Kachârî (Darrang)	dui	at	san	noka-buri
Mech (Jalpaiguri)	doi	wat	san	nokhafor
Lâlung (of Nowgong)	di	sara	sala	sanai
Dîmāsâ or Hills Kachârî (Cachar)	di	wai	shâin	dâi
Dîmāsâ or Hills Kachârî (Hojai of Nowgong)	dii	wai	sheng	deng
Gârô (standard, and Kamrup)	chi	wa'al	sal	ja
Gârô (Abeng, of Garo Hills)	chi	wal	sal	jajong
Gârô (Jalpaiguri)	chika	oar	rasan	rangret
Gârô (Âtông, of Garo Hills)	tai	wal	rangsan	changae
Gârô (Rugâ, of Garo Hills)	ti	wala	rasan	rarek
Kôch of Dacca	chi	al	sal	chând
Konch (Williamson)	ti	war	rashan	narek
Tipurâ (of Dacca)	tui	hor	sal	tal
Deuri Chutiyâ (Lakhimpur)	ji	nye	sa	ya
Deuri-Chutiyâ (Sibsagar)	ji	nye	sa	ya

For Rabha, which does have a short description (pp. 102-105), no satisfactory lexicon could be compiled in time.

The method is excellent. Local languages (what we call *parler* in French) are provided, with indications of location, and they are grouped according to eight more comprehensive and standard categories: Boro, Mech, Lalung, Dimasa, Garo, Koch, Tipura and Deuri.<sup>26</sup> Even the average reader may remark that the same label “Dimasa” groups two distinct *parlers*, one in Cachar (probably not far from Haflong) and a more northern one. The southern one does not have the /əi/ diphthong alluded to above, when describing the 1873 *JASB* paper by Butler, and people there pronounce /di/ for water, for which reason they are called *Dimasa* and not /dəimasa/: the first syllable here means “water, river”. This southern dialect is influenced by the local Bengali dialect and before final /ŋ/, here

<sup>26</sup> Note that the *LSI* writes “Deuri”, more exact than the previous Deori, since the Assamese orthography corresponds to /deuri/, not /deori/.



So far as the vocabulary of the specimens goes, Dimā-sā, Hojai and Tipurā are nearer the standard dialect than the others, and Chutiya is least like Bodo. But many words run through the whole group, and in some cases afford interesting phonetic changes. I give some instances. [the chart follows]

And after the chart, we find only these two lines:

The words “give”, “seize’ and “cloth’ seem to show that Bodo is a degenerate member of the group and has softened its sounds.

Here are the 3 degenerate words:

	give	seize	cloth
Bârâ	hŭ	hom	Hi
Rābhā	rā	rim	nen
Lālung	as	ram	Re
Dimā-sā	ri	rim	Ri
Hojai	ri	rem	rei
Gāro	-	-	-
Tipurā	ru	rom	Ri
Chutiya	re	-	-

We understand what he means: instead of the common /r/ sound, Boro has a /h/: a specific change.

But this is the only definite change which is described. Any other important question (Why do these languages have to be grouped together? Why do some form a closer unit? Why are others such as “Chutiya” (Deuri) less close to the supposed Boro center? And why choose Boro as a centre?) is not even touched upon. We are supposed to look at the comparative lexicon and conclude.

### 5.2. Comparative practice in Europe: the professional context

The comparative practice at that time was theoretically different. You were to follow the method (borrowed from Natural History) of common innovations. The languages in a group were supposed to “descend” from a common ancestor and inherit all its characters. Sometimes, a character changed (an innovation occurred), and all languages within the group showing this change descended from this specific ancestor-language, which then formed the specific branch where this change occurs. “Reconstructing” the history of the language group involved (and still does so to some extent) tracing back over the history of specific changes

through a kind of genealogical tree with “mothers” and “sisters”.<sup>28</sup> For instance, the “r- to h-” change is specific of “Boro” as a whole, which means that all *parlers* that exhibit this change are considered Boro.

Ideally, one should find features that exist in all Boro-Garo languages, and only in Boro-Garo languages; they would technically define the group. Such features do exist. For instance, all Boro-Garo languages have something like /aŋ/ for the pronoun “I, me”; and neighbouring languages do not have this specific pronoun, they mostly have words like *ŋa* or *ka*. However, neither Anderson nor Konow mentions such important features.

Of course, there is another possibility. Suppose one language, say Boro, developed such a change from *ŋa* to *aŋ*; then, this innovation was borrowed by the other Boro-Garo languages. In that case, the “*ŋa* to *aŋ* change” is not as old as expected: it was widespread among speakers of languages that were already distinct. Thus, it cannot be considered as the direct witness of an older common language from which all Boro-Garo languages ultimately descended, but only as an indirect witness of communication and exchanges between speakers of these languages, at some period. Such borrowings do happen, even with pronouns. For instance, most Khasi languages have *ŋa* for “I, me”, although they are, given all the other features, very different languages from their neighbouring Tibeto-Burmese languages. This fact supports the “diffusion” theory, rather than the “inheritance” theory — even if in that case, the borrowed pronoun is not from a Boro-Garo language.

The debate between these two theories, diffusion vs. inheritance, was at its peak in European universities, at the time of Konow and Grierson, at the end of 19<sup>th</sup> century. In the *LSI* nothing transpires.

However, it is not quite fair to describe the professional arena in this way, as if two contrasting theories were opposing each other, one more dependant on biological metaphors (family, sister, descent, inheritance), with the other more socially oriented, taking into account contact, gift and exchange. It gives the wrong feeling that you were to side either with physical or with social anthropology. If such ideas were indeed published and supported, especially in Germany, the first one most notably by August Schleicher (1821-1868), the second one by Johannes Schmidt (1843-1901), yet Max Müller (1823-1900) in England harboured a critical view of Schleicher’s *Ursprache*. Müller explained, in one essay about Phonetic Laws,<sup>29</sup> that the very idea of an Original Language is wrong because the diversity of dialects always precedes any classical or standard language

---

<sup>28</sup> Linguists spoke of “sister language” (not “brother”) because *die Sprache* is feminine in German, the language in which the professional terminology was developed.

<sup>29</sup> In *Contributions to the Science of Mythology*, 1897.

that could only rely on them. Even in Germany, a strong reaction blew up against the transposition of Natural Sciences and the reign of “phonetic laws”, for instance in the work of Hugo Schuchardt (1842-1927).



5. “To milk”: distribution (Brun-Trigaud *et al.* 2005: 30).

The French situation is interesting in this respect, because dialectology was flourishing during the very period when the *LSI* was written. French dialectology underwent impressive developments, partly for political reasons after the unpleasant defeat by Germany in 1871. The idea of publishing an Atlas of the French *parlers* was considered a national feat, and planned for the International Paris Exhibition, to be held in 1900. The extraordinary enterprise of the *Atlas linguistique de la France* was achieved by two men, professor Jules Gilliéron and his assistant, Edmond Edmont, a retired grocer and a gifted amateur linguist. Between 1897 and 1901, thanks to the railways, a bicycle, and his own feet, Edmont visited 639 spots where he investigated the actual pronunciation of hundreds of words, which he all wrote down with a special system he had been trained in by Gilliéron, who also drew up maps where the data could be efficiently compared. In 1911, Edmont also visited Corsica. The upshot of the scientific findings fell like a bolt of lightning on the (almost) peaceful



theoretical landscape: the “laws” and predictions of “German science” were not working so well, because the real words in the villages displayed a much more variegated picture than was expected.

One expected vast areas, each with a characteristic pronunciation and, after a more or less definite border, another dominant area, something similar to states controlling countries. But in reality it was rather different. One main reason is the “lexical replacement”. Suppose you want to know how “to milk (the cow)” is said all over France. You would perhaps expect two main areas, northern (with a Germanic influence) and southern (closer to the Latin lexicon). What you actually get can be represented, after due simplification, by map 5 above.

Indeed, one observes two main zones, one from the Latin *mulgere* (giving *molzer* and *moude*), another from the Latin *trahere* (giving *traire* and *tirer*), but the extension of each is rather unexpected. The *mulgere* zone is strikingly divided into two, the south and the north, and the *trahere* zone goes as far as the Basque country in the SW. Moreover, intruders appear right in the middle: the small zones with *ajouter* and *aria*. Finally, small spots pop up in many places. There are detailed explanations in most cases, but our main point here is that words do extend their influence: some words do gnaw away at the influence of others, and cross expected borders. More embarrassing: some words do that, and others do not, with the logical result that, depending on which word you study, you will get different borders, different maps. Of course, when we write “words do this or that”, we only try to depict the behaviour of speakers.

This is vividly illustrated by the comparative table given by Anderson in the Boro-Garo section of the *LSI*. Let us select 3 different “words” (that is: meanings), “to die”, “swine”, “good”. When drawing maps from this chart, you would have to draw three different border lines.

	“to die”	“swine”	“be good”
Bârâ	Thoi	oma	ham
Rābhā	Si	bak	nem
Lālung	Thi	oa	
Dimā-sā	Ti	hono	ham
Hojai	Thei	han	
Gāro	Si	wak	nam
Tipurā	Thui	wak	
Chutiyā	Si		

For “to die”, all Boro-Garo (BG) languages have a comparable word, which can reasonably be ascribed to a common older etymon, say BG \*thai. For “swine”, contrary to all appearances it is also true that most

languages have a common etymon, say BG \*hwak, sometimes reduced to /o/, then augmented with a suffix *-na* or *-ma*; but Deuri does not have this, except if it can be supposed in *meja*. For “good”, according to Anderson’s chart, the etymon BG \*nham is only found in 4 languages. Research shows that it also existed in “Tipura” (Kokborok) *ham*, but it is not attested in Lalung-Tiwa (which has *ku-mun* instead), nor in Deuri (which has *ču*). One possible explanation for the restricted spreading of this “root” is borrowing: Tai languages have a word *naam* that might be the source.

Scholars such as George Grierson, Sten Konow and James Anderson, though unfamiliar with contemporary developments in French dialectology, knew the difficulties met when using too a rigorous view of “phonetic laws”. Their methodological indifference when forging classifications and establishing linguistic groups was only partial or apparent. Moreover, they wanted to provide a linguistic Survey, not volumes of discussions.

Nevertheless, once again, comparison was a rather clumsy exercise. The reader is provided with lists of words, then lists of “closely related” languages. But the gap between the cause and the consequence is still quite wide.

### 5.3. *The consequences of the LSI*

Until the 1950s, the picture portrayed in the *LSI*, as far as Boro-Garo languages are concerned, was left practically untouched. All Indian or foreign scholars involved in languages in India used it. Nothing was ever done on the same scale and, I believe, is ever likely to be done.<sup>30</sup>

The political consequences, which are real, are more difficult to assess. The British power in India developed a system of schedule, basically a list of “tribes”, to mark them out from “the average Indian”. Among the criteria for “being tribal”, the use of a specific, non Indo-Aryan language was all important. This system was completely adopted by independent India after 1947, even if some modifications were made to the list.

Therefore, the definitions and labels provided by the *Linguistic Survey of India*, a quite official venture, became the Book and the Law about finding who is tribal and who is not. Nowadays, many revival movements, and efforts to teach the old language to youngsters, have no other obvious aim than assessing or confirming the scheduled status of the group. This tendency is certainly not limited to India.

---

<sup>30</sup> This is not to diminish the importance of the “New Linguistic Survey of India” enterprise launched by our Indian colleagues. But circumstances are different and, in North-East India, investigations in many places would now be more difficult than in the 1900s, as is perceptible from the gaps in the recent 2001 Census.

## 6. Boro-Garo after the LSI

### 6.1. The LSI imposes its labels

The older labels were not always maintained in the later Census, after the LSI became influential. The label “Minor Bodo” was merged in “Kachari”, the label “Mech” as well, while “Dimasa” was differentiated, also because it had a different geographical asset. The fantastic drop in the number of Lalung is probably due to the fact that Meghalaya is not Assam. A similar remark is relevant to “Tipura”. Part of these changes appears in the statistics, although comparisons are made very difficult by the shift in district borders.

	1881	1891	1901	1911	1921	1931
Kachari	263	199	239	283	260	283
Min. Bodo d.		4				
Mech	58	91				
Dimasa			20	17	11	15
Moran			78	24		
Tipura	4	122	112	136		
Koch		8		10	5	1
Hajong	1					
Rabha	56		20	29	22	27
Lalung	47	40	16	12	10	9
Garo	112	145	186	193	173	192
Chutia			2	3	4	4
TOT	542	609	596	682	485	531
Assamese	1 361	1 436	1 539	1 534	1 726	1 995

**Number of speakers (in thousands) according to LSI**

Such numbers do not give the real linguistic situation in Assam, if only because they do not represent the importance of bilingualism (taken into account in the modern Census). Other problems come into play. A major one is the ongoing numerical minorisation of the “local” tribal languages. The enormous increase in population, especially in Upper Assam, was the result of causes which are not shown above: the importation of a massive labour force for the tea gardens, an important cash crop; a massive influx from its overpopulated neighbour, Bengal.

### 6.2. On classification: Shafer and Burling

The first sharp divergence from the LSI can be seen in Robert Shafer’s classification, proposed in his 1953 paper:

LSI	Shafer
Bodo + Mech	→ Dimasa
Lalung	Lalung
Dimasa + Hojai	incl. Moran and Bodo-Mets
Garó (incl. Atong and Koch)	Garó
	Koch, incl. Atong and Rabha
Rabha	→ Koch
Tipura	Tipura
Chutiya	Tsutuya
Moran	→Dimasa

Shafer reviewed the whole of Tibeto-Burmese languages, mostly from a Himalayan point of view because data from China hardly existed at that time, and those from Burma were also rare, except for Jingpho which Shafer calls “Kachin”. Shafer was neither a British officer on duty, nor responsible for public opinion and therefore quite free to busy himself only with linguistics.

The main points in Shafer are the following: (1) Boro-Mech, Dimasa and P. R. T. Gurdon’s Moran data are one language. (2) Koch is distinct from Garó but identical enough to Rabha. He puts Atong with Koch. I cannot vouch for Atong, but as far as his two main points are concerned, I believe he was right. Shafer briefly commented his classification. I quote his paper because it has become rare. This first extract shows how he used the *LSI*:

Atsik, Awe, and Abeng differ from each other only slightly. Dacca, however, is a slightly aberrant dialect of Garó. I have used “Koch’ to designate a branch of Barish which was not clearly differentiated by Sten Konow in the *LSI*. The Cooch Behar text in the *LSI* is the same as the Jalpaiguri dialect. Konow placed Cooch Behar and Atong texts under Garó, where they do not belong. Konow listed as a source on Koch the *Essay on the Koch, Bodo, and Dhimal Tribes* of Brian Houghton [Hodgson] of which the Koch is worthless except as an Indic dialect, and he omitted Hodgson’s Garó, which is Jalpaiguri (Koch), in the latter’s article “On the Aborigines of North-Eastern India”. Tipura is phonetically similar to Koch, but in vocabulary is probably more like Garó and Dimasa. The classification of Lalung is based on very meagre material, but is believed to be approximately correct. (Shafer 1953)

Schafer was also the first scholar to specify the generic features of the “Boro-Garó” group, which he called “Barish”. He thought that he could group his “Barish” with his “Nagish” in a “Baric” larger unit. TB is for Tibeto-Burmese, ST for Sino-Tibetan.

Although most Baric stems are found in TB or ST languages, I have classified Baric as a separate division because it has some very common stems that have not been found in most of the ST languages. Thus most TB languages have a

word for “sun” corresponding phonetically to old Bodish *nyi-*, but the Baric languages have \**sal* “sun”. Most TB languages have a word \**mei* “fire”, but Baric has \**war*. Now, most of such stems are also found in Kachin, and the question arises whether Baric and Kachin should not be included in a single division of the ST family. This cannot be answered definitely at present; but since Kachinish has a much larger number of comparisons with Burmish and Kukish than with Baric, I have tentatively placed it in the Burmic division. The determination of the position of Kachin depends on whether Baric once extended into territory now occupied by the Kachins and Kachinish borrowed those words from Baric, or whether Baric and Kachinish both borrowed from a substratum language, or whether these unusual words were newly coined and replaced the old ones. One can only say that the extent to which such unusual words are used becomes less as one goes from west to east, from the Garo Hills to the Kachin country, so their place of origin seems to have been in the west.

These ideas will be made more explicit still by Robbins Burling in his 1983 *Language* paper, “The Sal languages”, where Burling develops Shafer’s idea of a specific link between Boro-Garo and Jingpho, starting with this same example of *sal* for “sun”, but discards the link with Kuki and substitutes Northern Naga.

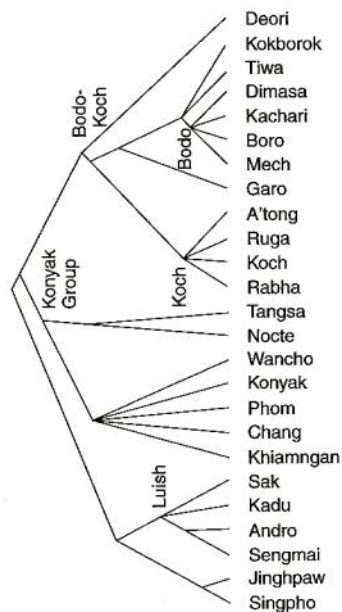
Since 1956, when as a young anthropologist, he did his PhD fieldwork in Rengsangri, a Garo village, Burling has been working on Boro-Garo languages and the Boro-Garo grouping.<sup>31</sup> At that time, he compared his Garo data with the Boro lexicon, and in 1959 he published, in *Language*, a paper entitled “Proto-Bodo”. In this paper, he more or less starts with Shafer’s work, but then takes a different direction, perhaps because he considers that Garo has a major role to play: although he still refers to the whole group by the name “Bodo”, the long privilege of Boro as a core language for the group, dies out.

After his return to North-East India, Burling published his 1983 paper and many other ones. Among the more important ones for us are his 2004 *Mandi Grammar* which elaborates on his earlier *Garo Grammar*, the 2006 *Comparative Bodo-Garo* book he wrote with U. V. Joseph, and his 2003 paper about the languages of North-East India, which sums up his views about Boro-Garo classification. As the 2006 book shows, Burling concentrated on western Boro-Garo, because he had easier access to languages in and around Meghalaya, and because he thought — not without good reason — that Garo was more a “core language” than Boro.

This is clearly shown in his 2003 classification below.

---

<sup>31</sup> From his experience in Garo country, he published his PhD work (*Rengsangri, Family and Kinship in a Garo village*, 1963) and a small but good *Garo Grammar*, 1961.



It looks like a rather complicated stemma, where the length of branching suggests a degree in relatedness. He has 4 (or 3) main groups under his “Bodo-Konyak-Jinghpaw” : (1) Bodo-Koch, (2) “Konyak group”, (3) “Luish” and (4) Jinghpaw, the last two being somewhat closer to each other than to the others.

His “Bodo-Koch” includes (1a) Deori, (1b) “Bodo”, from Kokborok to Mech, (1c) Garo, (1d) “Koch”, from A’tong to Rabha. Deori belongs here but is farther from the norm; Garo (Burling’s favourite language) holds a central position between “Boro” and “Koch”.

Burling tends to use *group*, he is embarrassed by the metaphor with *branch*: “I do not use “branch”, as Shafer did, to imply a specific level in a taxonomic hierarchy, but only as an informal way to label one part of a larger group.” Then he adds:

The data given in Burling (1959) suggests that Garo is closer to Bodo than to Koch, which is why I prefer to call the larger group Bodo-Koch, but most published classifications imply that Garo is closer to Koch.

The most important points for us in these developments will be commented in section 7.

Burling's paper was published in a state-of-the-art volume, and gives a panoramic view; this perhaps explains why no technical reason (except the citation of his 1959 paper) is given for his classification. Reasons will be made more explicit in his 2006 book written with Joseph. The basic idea, as with Jacquesson's 2005 essay on the same topic, is that comparative phonology provides a sound basis. However, lexicon plays an important role in Burling's classification, since he accepts Shafer's ideas about the "sal" vocabulary as specific to a whole western area or branch of Tibeto-Burmese languages. According to Burling, "sal" words are not borrowings: they are as many witnesses of an older common lore between Boro-Garo and Northern Naga ("Konyak") languages, and they are the reason to propose a higher "Bodo-Konyak-Jingpho" branching.

## 7. Grouping languages

### 7.1. *"Les mots et les choses"*

Although the Boro-Garo group of languages has certainly, been the most distinctly identified in North-East India, and the one that has been studied most closely for longest, the "Boro-Garo" label is still undergoing transformations.

It has been transformed in three ways. This will act as a brief conclusion to this paper. The first one is its inner meaning: which languages are concerned, and how should the inner history of the group be understood. The second one is the inscription of this group in the wider grouping of Tibeto-Burmese languages: what about these Northern-Naga (or "Konyak") languages that Burling wants to graft on them, and what is the aim of the discussions about "Baric" or "Barish"? These questions are supposed to have been "technically" answered: with convincing evidence and clear reasoning.

One important aspect should be pointed out. In linguistic discussions held over recent years, the amalgamation between people and language has stopped. If local scholars, for "nationalistic" reasons often want to parade their language as evidence that they are the true people, the oldest ones etc., very few linguists and very few anthropologists would now induce "race" from "language" or vice-versa. This precaution is all the more necessary in North-East India, where matrilinearity (a favourite keyword of ancient classificatory anthropology) is widespread among people of very different languages; and where the court language of old, Ahom, completely disappeared when the speakers gradually shifted to Assamese.

The third and final aspect is about the label and its implications.

### 7.2. “Boro-Garo”, the label

All through this paper, I have favoured the label “Boro-Garo” against several other ones that have been offered. There are a number of reasons for this, which will also illustrate, as examples, a wider concern about “what is at stake, when forging critical entities?”

I think it unwise to extend the use of a well-defined name by introducing “Northern Naga” inside, as one sometimes does. Whatever we think of the relationship between Northern Naga languages and actual Boro-Garo, it is clear that this relationship is looser than within Boro-Garo proper.<sup>32</sup> Therefore, if one uses “Boro-Garo” for any extended group, one blurs the focus and the result. The consequence is that we should keep “Boro-Garo” for what *LSI* calls “Bârâ or Bodo”, and find other names for extended groups.

Burling explains (quoted above) the technique of the *dvandva* (a two-member label), by pointing out that two of the most distant members in a group should be used for coining the name of the group (that is why he prefers “Bodo-Koch”). Historically, as for instance in “Tibeto-Burmese”, it is not true that the two languages involved are linguistically the most distant; they were only the most well known when the label was created. The same principle of notoriousness is satisfactorily applied to “Bodo-Garo” or “Boro-Garo”. Suppose we discover (I wonder by which ratio) that the most distant languages within Boro-Garo are Atong and Deuri, the resulting label (Atong-Deuri instead of Boro-Garo) would be no better.

The *dvandva* technique is better than, for instance, calling “Bodo” or “Boro” the whole group, whatever its extension. Using “Boro” for the whole group would lead, in indexes or out-of-context literature, to severe misunderstandings (most people would not realize whether the group or the eponymous language is being referred to). Secondly, by using one name only, as a kind of symbol or metonymy for the whole, we run the risk of mixing up extensions of different sizes, depending on reference or fantasy.

Now, Boro or Bodo? Anderson explained in the *LSI* that the term Bodo received so far is phonetically inadequate, that the true pronunciation is closer to Boro, actually open /o/ (hence his “â”) and retroflex /r/: he writes it consistently with a dot under the “r”. In present-day literature we find a tendency to write “Boro” for the language, and “Bodo” for the

---

<sup>32</sup> This was exactly Burling’s position in 1959, note 2: “It is true that these “Naga” languages show enough specific lexical correspondences with the Bodo group to make this an attractive conjecture; but even if this should prove to be justified, it would not disturb the coherence of the older Bodo group.” In his 2006 book written with U. V. Joseph, Burling takes into account only the Boro-Garo group *stricto sensu*.



people. This might be interesting, were it only consistent in the press or, more modestly, in professional literature. Unless some agreement is reached, the difference in orthography looks superficial and, since Boro is better (on phonetic grounds) and received in common usage, I think we should write Boro. And therefore “Boro-Garo” instead of “Bodo-Garo”. This is what Joseph & Burling did in their 2006 book.

### *7.3. Playing with names*

Linguists who are interested in the history of their discipline, more specifically when they review (with the usual sympathetic smile) the history of their own field of research, are often amused by the constant shifts in names. These shifts may occur among scientists as well as among the populations studied: on both sides there is a constant shift in domains, in names and in their content.

From the groups under scrutiny, we somehow expect this behaviour: people do change identities, and names are powerful tools for that purpose. My personal example in the small field of Boro-Garo studies (Jacquesson 2006) is found in the constant improvements in dialectal grain, when the numerous lexicons gathered and published by British officers and/or scholars can be organized, and compared with new data collections. We then realize the subtle shifts and shades that unite Dimasa, Moran, Boro, Kokborok and Bru-Riang. And we understand better not only the linguistic area and the work ahead, but also why considering Boro a “core language” was after all not so misleading, at one level: it is true that, historically, communications between the people who spoke these languages are all important in order to understand the history of the region.

On the other hand, debates such as the position of “Koch” are stimulating: is it closer to Garo or, as I suggest, a historical unit with Rabha — which actually tends to isolate Garo in a specific position instead of considering it as the new “core language”... These tactical moves explain why linguists also play with names, behind their spectacles. Putting this here, moving that forward, using this label here may indeed seem like a game. The very label “Boro-Garo”, or its many transformations in form and content, illustrate how coinage can serve as probing hypotheses. Rather often, a new coinage picks up momentum and enjoys a large distribution, while the puzzled scholar still wonders if he or she was right or wrong. Or, conversely, while the scholar basks in the sun of scientific success, the concept he or she coined takes on quite another meaning and comes back to him or her as a nasty snake.

**Abbreviations & References:**

- BSLP. Bulletin de la Société de Linguistique de Paris, Paris.*  
*JASB. Journal of the Asiatic Society of Bengal, Calcutta.*  
*JRAS. Journal of the Royal Asiatic Society, London.*  
*LTBA. Linguistics of the Tibeto-Burman Area, Berkeley.*
- Anderson, J.D. 1895. *A Collection of Kachári Folk-tales and Rhymes, intended as a Supplement to Rev. S. Endle's Kachári Grammar*, Shillong.
- Bradley, D. 1994. "Tibeto-Burman". In *Atlas of the World's languages*, edited by C. Moseley & R.E.Asher, pp. : 168-181. London: Routledge.
- Brandreth, E. J. 1878. "On the non-Aryan languages of India". *JRAS* n.s. 10: 1-32.
- Bright, W. (ed.). 1992. *International Encyclopaedia of Linguistics*. Oxford University Press.
- Brown, N. 1837a. "Alphabets of the Tai language". *JASB* 6: 17-20.
- Brown, N. 1837b. "Comparison of Indo-Chinese Languages". *JASB* 6: 1023-1038.
- Brown, N. 1850. "Aborigines of the North East Frontier". *JASB* 19: 309-316.
- Brown, W. B. 1895. *Outline Grammar of the Deori-Chutiya Language...* Shillong: Assam Secretariat.
- Brun-Trigaud, G., Y. Le Berre, Le Dû, J. 2005. *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont*. Paris: CTHS.
- Burling, R. 1959. "Proto-Bodo". *Language*. 35 (3): 433-453.
- Burling, R. 1983. "The Sal languages". *LTBA* 7 (2): 1-32.
- Burling, R. 2003. "Northeastern India". In *The Sino-Tibetan Languages*, edited by G. Thurgood & R. LaPolla, pp. 169-191. London: Routledge.
- Burling, R. 2004. *The Language of the Modhupur Mandi (Garo)*. Vol. 1: Grammar. Delhi: Bibliophile South Asia/Promilla.
- Butler, J. 1873. "Rough Comparative Vocabulary of some of the dialects spoken in the Naga Hills district". *JASB* 42 (1): Appendix I-XXIX.
- Campbell, G. 1866. "The Ethnology of India". *JASB* 35 (2) special number: 1-278.
- Census of India for 1881. Assam. Ccutta, 1883.*
- Damant, G. H. 1880. "Notes on the Locality and Population of the Tribes dwelling between the Brahmaputra and Ningthi Rivers". *JRAS* n.s. 12: 228-258.
- Deloche, J. 1984. *Les aventures de Jean-Baptiste Chevalier dans l'Inde orientale, 1752-1765*. Paris: Ecole Française d'Extrême-Orient.
- Elwin, V. 1969. *The Nagas in the Nineteenth century*. Oxford University Press.
- Endle, S. 1884. *Outline Grammar of the Kachari (Bara) Language*. Shillong: Secretariat Press.

- Endle, S. 1911. *The Kacharis (Bodo), with an introduction by J. D. Anderson, and editorial notes by P. R. T. Gurdon*. London: MacMillan & Co.
- Fisher, T. 1840. "Memoir of Sylhet, Kachar, and the adjacent districts". *JASB* 9: 808-842.
- Gait, E. 1926. *A History of Assam* (2<sup>nd</sup> ed.). Calcutta: Thaker, Spink & Co.
- Gurdon, P.R.T. 1904. "The Morans". *JASB* 73: 36-48.
- Hale, A. 1982. *Research on Tibeto-Burman Languages*. The Hague: Mouton.
- Hamilton (Buchanan), F. 1940. *An Account of Assam*. Edited by S. K. Bhuyan. Guwahati: DHAS (repub. 1963, 1987).
- Hodgson, B. H. 1847. *On the Aborigines of India. Essay the First: on the Kocch, Bodo, and Dhimal Tribes*. Calcutta: J. Thomas.
- Jacquesson, F. 2005. *Le deuri, langue tibéto-birmane d'Assam*, Leuven/Paris: Peeters.
- Jacquesson, F. 2006. "La reconstruction linguistique du passé : le cas des langues Boro-garo". *BSLP* 101 (1): 273-303.
- Joseph, U. V. & Burling, R. 2006. *Comparative phonology of the Boro-Garo languages*. Mysore: Central Institute of Indian Languages.
- Kaye, G. R. & Johnston, E. H. 1937. *Catalogue of Manuscripts in European Languages, vol.II : Minor collections and miscellaneous manuscripts, section 1 - nos 1-538*. London: British Library, India Office.
- Lewin T.H. 1869. "The Hill tracts of Chittagong and the dwellers therein". In *Selections from the Records of the Government of Bengal*, n°43. Calcutta.
- Montgomery Martin, R. 1838. *The history, antiquities, topography and statistics of Eastern India, comprising the districts of Behar, Shahabad, Bhagulpoor, Goruckpoor, Dinajepoor, Puraniya, Ronggopoor, and Assam*. London: W.H. Allen.
- Nath, D. 1989. *History of the Koch Kingdom 1515-1615*. Delhi, Mittal.
- Phayre, A. P. 1841. "An Account of Arakan". *JASB* 10 (2): 679-712.
- Phayre, A. P. 1844. "On the history of Arakan". *JASB* 13 (1): 23-52.
- Robinson, W. 1841. *Descriptive account of Assam, with sketch of local Geography and a concise history of the tea plant in Assam*. (2 vols) Calcutta: s.n.
- Robinson, W. 1849. "Notes on the languages spoken by the various tribes inhabiting the valley of Assam and its mountain confines". *JRASB* 18 (1): 183-237 & 310-349.
- Schleicher, A. 1866. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen. Kurzer Abriss einer Laut- und Formenlehre der indogermanischen Ursprache, des Altindischen, Alteranischen, Altgriechischen, Altitalischen, Altkeltischen, Altslawischen, Litauischen und Altdeutschen*. Weimar: H. Böhlau.
- Shafer, R. 1953. "Classification of the Northernmost Naga languages". *Journal of the Bihar Research Society* 39 (3): 225-264.

Sirāj, M. Ibn. 1881. *Tabāḳāt i-Nāṣirī: a general history of the Muhammadan dynasties of Asia, including Hindustan; from A. H. 194 (810 A.D.) to A.H. 658 (1260 A.D.) and the irruption of the infidel Mughals into Islam*. Translated from original Persian manuscripts by H. G. Raverty. (3 vol. 1881-1897) London: Gilbert & Rivington.