Project Examining the Progressive Erasure of Minority Linguistic Identities in India between 1870 and 1931 using the decadal British Census Records

Nanda Rajiv

Index

| Introduction

An old Hindi proverb wittily observes the linguistic diversity in India – *'Kos kos par badle paani, char kos par vaani'* - for every *kos* (approximately 3km) the water changes, and every four *koses*, the language does.[1] Census data from the period of the British Raj holds useful information about the linguistic identities of people in India.

| Limitations of Source

Before beginning to establish a methodology and then look into the data, it is imperative to note the potential limitations that this particular source could pose.
Walter Graham Blackie notes in *The Imperial Gazetteer* (1856) –

The languages of peninsular India are Tamil, Teloogoo, Canarese, and Malayim, with some lesser dialects spoken by the more barbaric tribes on the mountains; Mahralhi and Gujerathi prevail in the North and Northwest parts of the presidency; Oorea in the Northeast, and Hindoostani is the language spoken everywhere by the Mahometans. [II, 257] [2]

There is certainly useful information to be obtained from the above. We gain insights about the spatial distribution of different languages in Southern India. But as is clear from the choice of words, such as the use of the phrases 'the more barbaric tribes on the mountains' and 'lesser dialects', there are Orientalist undertones to these official government documents.

Furthermore, the British classified Indian languages into *vernacular* and *classical* based on their use by the 'common people', or by the upper and educated class respectively.[3] Basing the classification based on an existing hierarchical system of class could potentially lead to the creation of an artificial parallel hierarchy of languages.
For example, despite being a majority language in certain regions, languages in tribal and hilly areas were categorised as minority languages. Often, the census has a minimum number of total speakers of a language in order for it to be formally accounted for. Failing to meet the criterion makes the language be grouped into the 'Others' category, erasing entire linguistic identities.[4] These areas presumably had highly localised and unique linguistic identities, which can never truly be confirmed because of a dearth of data regarding them.
It must be conceded that the purpose of the census was the study the subcontinent at large, in which case this criterion can be justified, but there is no alternate source of official data about these 'minority languages', and hence there is no information accessible to us now in the present-day.

| The Objective and Scope of this Project

After an initial examination of the tables of the decadal census, as well as the general notes on the data collected, an aspect that stands out is the paradox of *which* languages are given precedence. As expected, there are accounts of the dominant Indian languages, as well as accounts of English literacy across various categories. Beyond this, there are records of minority languages that fit a certain criterion of number of speakers. They are put into a new set of tables, and presumably of lesser number of speakers than the first two categories, but enough to make note of, at least at that point. Then, there often are tables of foreign languages, such as French or German. Notably, there is one

---

[1] Lobo, B. (2018, 15 Jul). Losing Our Language. Church of St. Andrew, Bandra. Retrieved December 5, 2020, from https://standrewchurch.in/losing-our-language/
[2] Blackie, Walker Graham. The Imperial Gazetteer: A General Dictionary of Geography, Physical, Political, Statistical and Descriptive. 4 vols. London: Blackie & Son, 1856. Internet Archive online version of a copy in the University of California Library. Web. 21 November 2018.
[3] Grierson, G. A. (1903-1927). Linguistic Survey of India (Vol. 1, Part 1: Introductory). Calcutta: Government of India.
[4] Groff, C. (2017). Language and language-in-education planning in multilingual India: a minoritized language perspective. Language Policy, 16(2), 135-164.

census, 1881, in which a singular Latin speaker in the Madras Province is recorded. [5]
The scientific method of the officials of the British Raj presides again. Until we read general notes written by officers, which debate where languages of tribal areas are worth recording, as they perhaps will die out sooner than the census takers collates their data, or postulate that the Aryan languages are successful because they are the 'languages of a superior civilization'.

Now, it is true that the British Raj is behind us, these notes were known to hold these opinions, and hence this is hardly surprising. What *is* concerning, however, is the potentially numerous linguistic identities that have been likely overlooked as they did not meet a criterion, or were soon to be dead anyway.

It is an important task to look into how so-called minority languages came to be negligible in the records, any trends to how these specific languages which are outside of the dominant vernacular, classical and foreign languages are recorded, and how the narrative of the notes regarding these change as time passes and linguistic identities transform.

Because this will likely involve 'reading against the grain', and looking at what has not been recorded, in order to study a set of identities that have been cast out, this aligns with the general ideology behind subaltern studies.

| Census Questions

Colonial census-takers seem to grapple with enumerating linguistic data. The questions regarding languages for the first three decades vary. In the 1872 'census', which is also referred to as the 'House Register', there were no questions about language at all, only about the ability to read or write. In 1881, the census only asks a respondent what their 'mother tongue' is, which, in 1891 changes into what their 'parent tongue' is as well as what foreign languages they know, if any. Following this, there is a stabilisation in the questions asked from 1901 to 1931 – what language is ordinarily spoken in the household, and whether or not individuals are literate in English.[6]

| Year-wise Observations of Tables and General Remarks with Specific Focus on What are Referred to as "Tribal Languages"

| 1871-72

The House Register, or the 1871-72 census has only a brief overview of languages spoken across the subcontinent. As established in the previous section, there were no official questions about mothertongue, or English proficiency. However, it does contain a record of all dominant languages in India discussed by their spatial distribution. This is quite general, and there are no statistics. The languages mentioned are all dominant – Hindi, Oriya, Tamil, Malayalam, Punjabi, Telugu, etc. are written about. The section of language in this particular report is of a total of only about 700 words, which is mostly about spatial distribution of dominant languages. There is in effect, no mention about languages other than these. There is mention of tribes of hills in Odisha or the forests in the North-East that 'have dialects of their own', or 'languages or their own'.[7] Overall, there is not much information to be gained from this particular census, pertaining to language.

| 1881

[5] Superintendent of Government Printing, Calcutta. (n.d.). Census of 1881: Statistics of Population. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1881/CensusIndia1881GeneralReport.pdf (p. 139)
[6] Census Questions (1872-2011). (n.d.). Retrieved November 29, 2022, from https://censusindia.gov.in/census.website/CENSUS_ques
[7] George Edward Eyre, et.al. Census of British In 1871-72. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1871/CensusBritishIndia1871.pdf (p. 20)

In 1881, there are more extensive tables available. The first few focus on the dominant languages – Tamil, Hindi, Canarese (Kannada), Urdu, Sindhi, Punjabi, etc. Supplemental tables record several other languages as well as their provinces.

Supplemental tables record data of languages that do not have a number of speakers at the same range as those of the main tables. Many of these have only one recorded province in which they are spoken. These are likely highly localised languages that have developed separately from the dominant language of the province, and hence maintain a unique identity which does not allow them to be grouped under them. The provinces to which these languages belong are spatially far distributed. This is of note.

The languages themselves are highly localised and uniquely developed, but this phenomenon of such occurrences is widely distributed across many provinces. To demonstrate this interesting paradox, we can observe the following examples –

Province or State.	Both Sexes	Males	Females.
	Language—Kaikadi		
Hyderabad	5,294	2,694	2,600
	Language—Lepcha.		
Bengal	4,011	2,433	2,178
	Language—Kurumba		
Madras	3,886	1,913	1,974

([8])

These three languages had similar numbers of speakers, in 1881. Kaikadi and Kurumba are classified as former tribal languages, which are recorded as castes now. They respectively have about 20000 speakers as per the 2011 census.[9] The Lepcha language spoken by people in Sikkim has 30000 speakers in the modern day. [10]

This is an interesting phenomenon of localised languages developing uniquely in a small spatial region. But, the phenomenon is wide-spread from Madras to Assam. Provinces that seem to exhibit this trend as per the supplementary tables are Assam, Madras, Bengal, Burmah and the Central Provinces.[8]

[8] Superintendent of Government Printing, Calcutta. (n.d.). Census of 1881: Statistics of Population. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1881/CensusIndia1881GeneralReport.pdf (p. 137)
[9] "Statement 1: Abstract of speakers' strength of languages and mother tongues - 2011". www.censusindia.gov.in. Office of the Registrar General & Census Commissioner, India. Retrieved 2018-07-07.
[10] Plaisier, Heleen (2007). A grammar of Lepcha. Tibetan studies library: Languages of the greater Himalayan region. Vol. 5. BRILL. ISBN 978-90-04-15525-1.

| 1891

In 1891, the general remarks mention an integration of Western languages in 'hill tribes' -

"*[A Western language] meets in this stage the lagging Hill dialects; and what with the opening out of the country by roads and railways, the administration of the forests on the modern economic system, and the reluctant advance of the village schoolmaster, even the wildest of those tribes are being brought within touch of the outside world. Means of livelihood are being disclosed to them of which a few years ago they never dreamed, and each step forward is accompanied by the acquisition of something fresh in the way of vocabulary, even if the adventurer does not pick up an entirely new dialect."* [11]

From this particular census, records of languages are organised by their origin. We see a table of 'Indic-Aryan' languages, such as Hindi, Bengali, Marathi, etc., of 'Dravidian' languages, like Telugu, Malayalam, Tulu, Tamil, Kanarese, etc., 'Thibeto-Burman' (Kathe (Manipuri), Miri-Abor), Semitic (Arabic), Eranic-Aryan, etc. Under each of these, the last category records accounts of 'Minor Languages' of this origin. The numbers of this category are mostly 1 speaker per 10,000 people.

There are two categories that are kept separate alongside the broad origins of languages – the 'Gipsy Dialects' and 'Khasi Dialect', having 15 and 7 speakers per 10,000 respectively. The general remarks for these two groups observe that they are 'statistically insignificant'.

D. Gipsy Dialects - -	15
E. Khási Dialects -	7

Khasi, the language spoken between the Khasia and Jaintia Hills in Assam, is not being affiliated with any surrounding languages of either Aryan or Tonic families.
The report also speaks about the presence of Welsh missionaries in these regions. Khasi which was previously unwritten, is now adopted the Roman characters, because of the 'study and enterprise' of these missionaries. The number of speakers in 1891 is 178,000[12], and now has just over 1,000,000 speakers as of 2011.

Gipsy dialects encompass the languages spoken by wandering tribes in the plains of India, having 400,000 speakers in 1891. This category is not found in 2011 census.

It is of note that these two languages were kept from being integrated into some larger category, or a 'Minor Language' under one of the other broad classifications. It is acknowledged in the general remarks about the Gipsy dialects, that "*It is out of the question to distribute these languages amongst those having fixed dialects, as their character changes with the locality most favoured by the tribe using them, and, whilst retaining a backbone peculiar to itself, freely assimilates the local vocabulary and pronunciation.*" [12]

It is, however, of note, that all of the languages spoken by wandering tribes in the plains of India are put into one single category of Gipsy dialects. This offers no information about any unique characteristics of each of these, or of what areas these different tribes within this category are in.

When it comes to Southern Indian tribes, the tables reflect no data. In the general remarks, there is paragraph speaking of the Toda and Kota tribes near the Nilgiri hills. The general remarks concede, *"The [Toda] tribe has received a great deal of attention, ethnological, phrenological, social, and linguistic, mainly because it resides within an easy walk of a favourite hill resort."* This raises the question of, how many other tribes whose unique linguistic identity will never be recorded, by virtue of them *not* being within walking distance of some colonial hill resort and are just 'Minor Languages'.

---

[11] Eyre and Spottiswood. Census of 1891. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1891/CensusIndia1891GeneralReport.pdf (p. 136)
[12] (p. 138)

| 1901-1931

The 1901 Census retains the trend of categorising languages spoken in the subcontinent based on their origin. In fact, as was discussed in the questions section, the format of the language section from census in 1901-1931 remain the same by-and-large. The tables look the same in terms of categorisation.

General remarks note from this year, as well as similar statements from the previous census (1891) note that the languages of the eastern frontier are falling into disuse. The context of this statement is to debate whether or not it was worth sending officials to continue to methodically record each of the languages, given that once their work is collated and corroborated, the languages would have already fallen greatly in terms of use.[13]

Several factors could get into play for this, but none of these hypotheses can be confirmed to be true, given that there is no data available. Some of the factors include great increase in missionary enterprises, especially among indigenous people, the growth of cities, leading to migration away from indigenous lands and integration into other linguistic identities. Western languages are increasingly being adopted, or unique identities are assimilating into dominant languages for ease of communication in larger settlements.

Apart from the usual records of dominant languages in India, there is more deliberation on the part of the British about the reason for the dying of certain languages. They propose that –

"*The reasons for the success of Aryan languages in northern India are not far to seek. They are the languages of a superior civilization, while the tribal dialects with which they compete are unwritten and have no literature of their own. It is noteworthy that the Aryan languages do not seem to wage war amongst themselves. It is nowhere reported that that one such language is spreading at the expense of another."*[14]

With the benefit of hindsight, it is clear to see that this is a colonial point of view. However, there is merit to the question of why these languages are wiping each other out. The increased connectivity provided by the railway system, the deforestation and plantation growing implemented in agriculture, and resulting altered migration patterns could have come into play in this phenomenon. Furthermore, the existence of written script of certain languages and not of others, is point to note, which also factors in.

In the 1921 Census report, under the general remarks, firstly notes a rise in bilingualism. This is an important trend that coincides with the displacement of tribal languages, stemming from a need for a common medium of conversation. In this same line, the logistical benefits of a potential *lingua franca* for India are discussed. It states, *"The combined speakers of Eastern and Western Hindi considerably exceed in number the strength of any other language in India, and if we add […] Bihari and Rajasthani, […] we get well over 100 million speakers of tongues which have some considerable affinities and cover a very large area […]"* [15]

In the pursuit of ease of communication, tribal languages get displaced, and there is an increasing dominance of already numerically larger languages. This sets up a circle of a dominant language being the effective *lingua franca,* and continuing to remain dominant. This not only plays into linguistic tensions between North and South India in the present-day when debating a national language, but also more acutely affecting minor languages that have been completely wiped out as a result of this.

---

[13] Eyre and Spottiswood. Census of 1891. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1891/CensusIndia1891GeneralReport.pdf

[14] Office of the Superintendent of Government Printing, India. Census of 1911; Volume 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1911/1911%20-%20India%20-%20Vol%20I.pdf

[15] Superintendent Government Printing, India. Census of 1921; Part 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1921/CensusIndia1921IndiaReport.pdf

In 1931, the census general remarks note changes in classification for that year. Most of these involve language in the north-eastern states. Examples of the changes are Khoirao and Marami have been merged in Memi and Chang, since there is a decrease in the number of speakers, and Maring is removed from Naga-Kuki subgroup and integrated into Old-Kuki subgroup. These minor changes reflect major changes on the ground for these languages. [16]

Furthermore, in 1931, this table is also published –

SURVIVAL OF TRIBAL LANGUAGES. 351

Language groups.	No. of languages spoken.	No. of speakers as mother-tongue and subsidiary 1931 (bilinguals shown twice).	No. of speakers 1921.	No. of speakers mother-tongue only, 1931	Difference between columns 4 and 5 (increase in 1931+ decrease in 1931—).	Difference between columns 4 and 3 (increase + decrease —).
1	2	3	4	5	6	7
A.—Language of India and Burma.	225	366,480,537	315,525,177	349,887,527	+34,362,350	+50,905,360
(i) *Austric Languages*—						
1. Indonesian languages ..	2	6,542	5,561	6,542	+981	+981
2. Mon-Khmer languages ..	10	734,204	549,917	726,578	+176,661	+184,287
3. Munda languages ..	7	4,710,685	3,973,873	4,609,588	+635,715	+736,812
(ii) *Tibeto-Chinese Languages*—						
1. Tibeto-Burman languages	128	14,167,611	11,959,011	12,982,840	+1,023,829	+2,208,600
2. Tai-Chinese languages ..	11	1,150,220	926,335	1,027,656	+101,321	+223,885
3. Man and Karen languages*	17	1,351,291	1,114,817	1,342,278	+227,661	+236,674
(iii) *Dravidian Languages*—						
1. Dravida languages ..	7	47,032,874	37,285,594	41,454,593	+4,168,999	+9,747,280
2. Intermediate languages ..	5	3,661,277	3,056,598	3,609,418	+552,820	+604,679
3. Andhra language ..	1	28,195,824	23,601,492	26,373,727	+2,772,235	+4,594,332
4. N. W. language	1	231,581	184,368	207,049	+22,681	+47,213
(iv) *Indo-European Languages*—						
1. Eranian languages ..	3	2,457,134	1,981,675	2,270,466	+288,791	+475,459
2. Dardic languages ..	5	1,543,031	1,304,319	1,522,936	+218,617	+238,712
3. Indo-Aryan languages ..	19	261,105,909	229,560,555	253,699,403	+24,138,848	+31,545,354
(v) *Unclassed Languages*—						
1. Andamanese	2	466	580	466	—114	—114
2. Burushaski	1	26,076	..	26,076	+26,076	+26,076
3. Gipsy dialects† ..	6	25,999	15,018	25,999	+10,981	+10,981
4. Languages not returned and unspecified.	..	29,813‡	5,664	1,912	—3,752	+24,149
B. Languages of other Asiatic countries and Africa.	17	305,386	211,894	302,324	+90,430	+93,492
C. Languages of Europe ..	20	452,099	319,112	339,706	+20,594	+132,987

* The correct classification of these two languages is doubtful. Przyluski treats them as Tai (Meillet and Cohen, *Langues du Monde* 380), but it is probable they have Austric affinities.

† These dialects are drawn from various Indian languages and contain such diverse elements that they cannot fairly be allotted to one family rather than another.

‡ Includes Hill and aboriginal subsidiary languages (27,841).

Titled 'Survival of Tribal Languages'.[16]

The Census Superintendent of the United Provinces writes, "*These gypsy languages … are dying in this province. These wandering tribes are taking to a more settled manner of living. […] Majority of the present members of these tribes have never learnt these dialects, but speak […] Hindustani*.", which his counterpart in Mysore State also corroborates.

He also speaks about a continuous fall in those speaking all tribal languages. Some subsets of languages have greatly reduced, including Munda languages having effectively disappeared, though elements of them are mixed into the jargon of the region. Some tribes exhibit transforming linguistic identities due these changes, such as the Nahals, who initially spoke an indigenous Munda language, which was adopted to be half-Munda, half-Dravidian, and in 1931 was becoming an Aryan language.[17] This transformation has occurred within a span of a hundred years, or about four generations which is a relatively short period of time for a language to transform multiple times.

[16] Officer of the High Commissioner for India. Census of India, 1931; Part 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1931/CensusIndia1931IndiaReport.pdf (p, 358 – 359)

[17] (p. 360)

In 1931, there is a whole section discussing the crisis that tribal languages face. From 1871-72 where there was less than a page of total content speaking about Language, to 1931 where there is a dedicated section within the topic of Language discussing the survival of tribal languages, there seems to be marked improvement in the statistical method for this. This poses a paradox – the 'scientific' method that improves the depth of data it collects, also leads to the erasure or negligence of identities that are beyond the minimum criteria determined by them.

Ultimately, the bevy of factors, such as increased inter-connectivity, changing occupation and migration patterns, displacement from place of origin and Westernisation due to missionaries and education, have led to a huge decrease and in some cases extinction, of distinct identities of minor languages, especially tribal languages.

| What Is Not Seen

Census data and general remarks paint a great picture of how the use and prevalence of different minor languages changes and decreases over decades. There are many tribes that have a significant number of people, and have a compact presence in the areas they belong, and they are recorded to some degree. However, a question is raised of what we cannot see. Even with these relatively more dominant groups, it is clear to see ways in which they could be unaccounted for, or their linguistic identity erased. But what about those which have already ceased to exist, integrated into some larger supergroup?

Most of the limited data that the censuses provide regard tribes in the north-east and their languages. Tribes in the Western Ghats, in parts of the rainforests of Kerala, for example, have distinct identities separate from Malayalam, and are often a complex amalgamation of linguistic factors in the area, but also strongly developed on their own separate from the rest.

South India, especially, continues in the present-day to have a strong hegemony established of the four main languages corresponding to the four states. Although other languages remain in use, such as Tulu and Konkani in coastal areas, this hegemony creates a convenient framework in which languages can simply be clubbed into, losing the nuance and complexity of the rich culture that exists. This is only from the perspective of the *recording* of data. Language gives a glimpse of how *any* measure of identity can be crumbled down to decline.

So far, this study of the decline or erasure of minority linguistic identities has been largely confined to tribes, especially in the North-East, but these two are not necessarily interchangeable. There are linguistic minorities outside of tribal people, perhaps such as outlying castes not integrated into society, which have developed cultures disconnected from the rest.

| Additional Notes on Collecting Linguistic Data in General

Apart from this, the method of data collecting that is employed could potentially lead to a loss of information in translation from the conversation between the census-taker and respondent, and the data in tables we obtain.

An interesting note contained in one of the general remarks, for 1891 observes,

"In accordance with the general tendency noted in the introductory portion of this chapter, the first impulse, in many cases, is to return the name of the caste as that of the language. For example, the potter gives 'potterish,' the tanner 'tannerish,' or the weaver 'weaverish,' as his mother tongue, especially if he be either a member of a large caste or a stranger to the locality where he is being enumerated. In other instances, the name of the District or State will be returned." [13]

There was no vocabulary that existed for Indians, prior to the census, to *name* their language. This is an interesting instance, of a disconnect between the data that is available and the data that is required.

| Conclusion

Circling back to the main purpose of this project, which is to trace patterns of decline or erasure of minority linguistic identities with 70 years of census data from British India, there are key observations that can be made. First, on the ground level, there are a variety of factors at play, which are somewhat mapped by the general remarks. There are changes in patterns of migration and occupation, which create the need for a common tongue, which tends to be a dominant language. There is an influenza epidemic between 1911 and 1921, which affects some of these groups in particular. Some of these groups wage war amongst themselves, and fight for dominance. There are also Western Christian missionary groups that engage with these minority groups to a greater degree, bringing in Western influence, initially in form scripts for these tongues, or increased use of European languages. Apart from this, there is also displacement of indigenous peoples due to deforestation and growing of plantations by the British Raj, destroying the livelihood of these peoples.

At the level of data collection as part of the census, there is the erasure of distinct identities of minority groups that do not meet criteria such as minimum number of speakers, etc., and are clubbed with ‘Others’ or ‘Minor Languages’ categories, or added under existing dominant languages. This reduces the depth of vision, although it served the purpose of the census which was to get a wide-ranging understanding of the subcontinent. There is potentially a challenge to this argument, stating that before the British, there was no such census data whatsoever. However, a problem arises when there is data about *some* groups, excluding others. Now, there is *disparately* no data available from that period about minor groups. Then, the narrative changes, creating cycles of dominance of the measurably major groups.

Overall, there is insight into linguistic identities to be gained from what *is* contained in these official census documents, as well as insight to be gained from what is *not.*

| References

1. George Edward Eyre, et.al. Census of British In 1871-72. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1871/CensusBritishIndia1871.pdf
2. Superintendent of Government Printing, Calcutta. (n.d.). Census of 1881: Statistics of Population. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1881/CensusIndia1881GeneralReport.pdf
3. Eyre and Spottiswood. Census of 1891. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1891/CensusIndia1891GeneralReport.pdf
4. Office of the Superintendent of Government Printing, India. Census of 1901. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1901/CensusIndia1901GeneralReport.pdf
5. Office of the Superintendent of Government Printing, India. Census of 1911; Volume 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1911/1911%20-%20India%20-%20Vol%20I.pdf
6. Superintendent Government Printing, India. Census of 1921; Part 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1921/CensusIndia1921IndiaReport.pdf
7. Officer of the High Commissioner for India. Census of India, 1931; Part 1. Retrieved from http://piketty.pse.ens.fr/files/ideologie/data/CensusIndia/CensusIndia1931/CensusIndia1931IndiaReport.pdf
8. Lobo, B. (2018, 15 Jul). Losing Our Language. Church of St. Andrew, Bandra. Retrieved December 5, 2020, from https://standrewchurch.in/losing-our-language/
9. Blackie, Walker Graham. The Imperial Gazetteer: A General Dictionary of Geography, Physical, Political, Statistical and Descriptive. 4 vols. London: Blackie & Son, 1856. Internet Archive online version of a copy in the University of California Library. Web. 21 November 2018.
10. Grierson, G. A. (1903-1927). Linguistic Survey of India (Vol. 1, Part 1: Introductory). Calcutta: Government of India.
11. Groff, C. (2017). Language and language-in-education planning in multilingual India: a minoritized language perspective. Language Policy, 16(2), 135-164.
12. Census Questions (1872-2011). (n.d.). Retrieved November 29, 2022, from https://censusindia.gov.in/census.website/CENSUS_ques
13. "Statement 1: Abstract of speakers' strength of languages and mother tongues - 2011". www.censusindia.gov.in. Office of the Registrar General & Census Commissioner, India. Retrieved 2018-07-07.
14. Plaisier, Heleen (2007). A grammar of Lepcha. Tibetan studies library: Languages of the greater Himalayan region. Vol. 5. BRILL. ISBN 978-90-04-15525-1.