

May 21, 2020

Counts and Consequences

Citizenship and Language Identity

By: G. N. Devy

Language diversity has always been a given for Indians. The Constitution endorses this diversity as ‘a noble intention’. But as in 2011, the 2021 Census too may see hundreds of languages removed from official records in the service of Hindi-Hindu nationalism.

The unprecedented situation arising from the coronavirus pandemic may appear to be a complete break from the past, a kind of a threshold marking a new epoch. However, one very important strand of continuity in India between the pre-coronavirus era and the current corona times is the diminishing of citizenship rights, a kind of tragic finale of a two-century-old struggle for freedom.

The Citizenship Amendment Act (CAA) was passed by Parliament in the second week of December last year. Well until the announcement of the lockdown in the third week of March, massive agitations against the CAA, the National Register of Citizens, and the National Population Register rocked the country. During those four months, activists, analysts, and the media commented on every aspect of the clash between the current regime’s idea of nationalism and the idea of citizenship inscribed in the Constitution.

There is one aspect of the issue though that did not find much articulation during the agitation. It was the position of the minor ethnic-linguistic-religious groups. That apparently minor strand occupied a relatively minor place during the CAA debate. But in the post-coronavirus epoch, it is likely to acquire greater significance.

When the amended Citizenship Act declared that migrants from Bangladesh, Pakistan, and Afghanistan would be given Indian citizenship if they were minorities in those countries—such as Hindus, Christians, Sikhs, Buddhists, Jain, or Parsis—the new law left out other minorities who belonged to neither of these religions nor to the broad spectrum of Islam (such as Shia or Ahmedia). These were the [Mro](#), [Meitei](#), [Tripura](#), [Marma](#), [Tanchangya](#), [Barua](#), [Khasi](#), [Santhals](#), [Chakma](#), [Garo](#), [Oraons](#), [Mundas](#), [Marmas](#), [Tripperas](#), and [Tanchangya](#), all listed in Bangladesh as minorities. What marks out all these communities is that, apart from being theologically different, they are primarily linguistic communities.

Each of them speaks a distinct language. The names of their languages and of the community are identical. These minorities are smaller in population in Bangladesh than the other minorities that the CAA had declared as ‘welcome’ communities, but all of them face multiple vulnerabilities. Besides, all of them have been known in all historical periods to have migrated between what are now Bangladesh and India. Why they were left out is an enigma, unless we turn our attention to the regime’s handling of another aspect of the citizenship-language issue in India.

|| Though inspired by Europe’s ideas of nationalism, the Indian struggle for independence did not get bogged down by any linguistic chauvinism.

Since Independence, Indians have accepted language diversity as a social norm and an inalienable feature of the Indian nation. India’s Constitution endorses language diversity, at least as a “noble intention,” if not as a non-negotiable right of the republic. As citizens, we waste no opportunity to state our tolerance towards different languages. When nationalism emerged in Europe during the nineteenth century, conformity to a ‘national language’ was seen as an attribute of citizenship. Though inspired by Europe’s ideas of nationalism, the Indian struggle for independence did not get bogged down by any linguistic chauvinism. In social and personal spaces, formal and informal, numerous languages have continued to co-exist in India, both in the ancient and medieval past and during the seven decades since Independence. Yet, there are some grey areas in India’s knowledge about its linguistic diversity, that provide militant nationalism with triggers for justifying actions leading to mass extinction of individuals and their thoughts and cultures. The grey areas are: one, the country does not know how many languages its people really speak; and two, there is still a mix of myth and history surrounding the emergence and evolution of these languages.

In ancient and medieval sources, there is an impressive range of writings on the etymology, grammar, semantics, and philosophy of language. In works by scholars such as Panini, Matanga, Dandin, Rajashekhara, Al-Biruni, and Amir Khushro, we get details of various

languages spoken in the Indian subcontinent. However, no exact inventory of all of the languages was ever available. The first such enumeration was by the Irish scholar George Abraham Grierson, who worked in the colonial administration. He slogged for a good three decades to prepare the first and fairly definitive survey and in the first two decades of the twentieth century produced 11 volumes in 19 parts that offered a good descriptive record of 189 languages and several hundreds of what he classified as ‘dialects’. His description was based on the edifice of modern linguistics which, in turn, was quickened a century before him by William Jones’ chance observation about the striking similarity between the stock of words in Greek, Latin, Sanskrit, and Persian.

The map of India has substantially changed since Grierson’s time. After Independence, Indian authorities should have carried out a fresh survey. That exercise, however, remained desired but never accomplished, until a group of volunteers decided to make an attempt and compiled the *People’s Linguistic Survey of India* (2013-20). Invariably, in the absence of a definitive survey of languages, India depended more on the census count of languages for its language data requirements. The Census, carried out every decade, is scheduled to be carried out next year, in 2021.

Thanks to the classification protocol, the linguistic citizenship of 60 lakh citizens was just axed, rendered as ‘not worth consideration.’

One can imagine the outcome of the next language count going by what was released in June 2018 as the ‘Language Data of 2011 Census’. During the last census, the citizens of India provided 19,569 names of ‘mother tongues’. This data was described in technical terms as ‘raw returns’. Based on previously available linguistic and sociological information, the authorities decided that 18,200 of these reported mother tongues did not match “logically” with any known information. A total of only 1,369 names, or ‘labels’ as they are technically called, were picked up as being “names of languages.” The excluded ‘raw returns’ represented the responses of nearly 60 lakh citizens. Thanks to the classification protocol, the linguistic citizenship of these 60 lakh citizens was just axed, rendered as not worth consideration.

In addition to the 1,369 mother tongue names shortlisted in scrutiny, there were another 1,474 mother tongue names placed under the generic label ‘Others’. The enumerated people had languages of their own, but the classification system could not identify the languages they spoke.

The fortunate 1,369 were further grouped together under a total of 121 ‘group labels’. These were presented to the country as ‘Languages’. Of these, 22 were the languages included in the Eighth Schedule of the Constitution: the ‘Scheduled Languages’. The remaining 99 were described as the ‘non-scheduled’ languages.

The 2011 language census recorded that a total of only 2,59,678 Indians speak English as their mother tongue. Numerically accurate, sociologically disastrous!

Strangely, most of the groupings of languages were ‘forced’. For instance, under the heading ‘Hindi’, there were nearly 50 other languages. Bhojpuri, spoken by more than 5 crore people, and with its own cinema, theatre, literature, vocabulary, and style, was shown as Hindi. A population of nearly 3 crore from Rajasthan had their own independent languages, but they were shown to have Hindi as their mother tongue. The Pawri of tribal communities in Maharashtra and Madhya Pradesh had been hitched to Hindi. So was Kumauni, of Uttarakhand. The report stated that 528,347,193 people spoke Hindi as their mother tongue, which is simply not the case. A similarly inflated figure was given for Sanskrit as a mother tongue by counting the returns given for this language against the question about the ‘second language’. As against this, the use of English was not seen through the prism of a second language. Counting of English was restricted to the mother tongue category, in effect, bringing down the number of people reported as speaking English. Considering the widespread use of English in education, law, administration, media, and healthcare, a very significant number of Indians use English as a utility language. But the 2011 language census recorded that only 259,678 Indians speak English as their mother tongue. Numerically accurate, sociologically disastrous!

No matter if language is the main means of expression for humans, hundreds of the languages of Adivasis and nomadic communities in India are likely to be wiped out from official records in the service of Hindu-Hindi nationalism.

A few months from now, we shall see the Census gathering and analysing language data once again. Hindi, a language that I love greatly, will once again be shown as having gained tremendously in number of speakers. Sanskrit will be shown as being spoken by

more people than those speaking English or Urdu.

Who, it may be asked, will be the communities that stand to lose in the process? In most instances, it is the Adivasis and the nomadic communities, whose numbers are relatively small. However, the total number of such communities whose languages will go unreported in the Census is large enough to cause concern. The speakers of these communities will be subsumed within the statistical figures of speakers of other languages, mostly Hindi, since the larger sections of Adivasi communities are located along the margins of the Hindi speaking area. No matter if the Constitution accepts the right to expression as a fundamental right, no matter if language is the main means of expression for humans: hundreds of the languages of Adivasis and nomadic communities in India are likely to be wiped out from official records in the service of Hindu-Hindi nationalism.

The Census of India is a massive exercise. To dismiss its findings out of hand would not be fair. Besides, one can contest the figures and the hyper-nationalistic arguments by presenting alternate figures collected through smaller but specific data collection exercises. But there is an area that hyper-nationalistic language wisdom imposes on the country and which has so far not received a befitting rejoinder.

As mentioned earlier, a chance remark of William Jones during the last quarter of the eighteenth century excited linguists all over the world, giving rise to the discipline of Historical Linguistics. Scholars of this discipline joined their intellectual might to map out language genealogies and the idea of distributing languages of the world in terms of ‘families’ received universal acceptance. One such proposed family was the Indo-European. It was proposed that one could trace the past as far back as an imagined ‘proto Indo-European’, for whose existence there was neither linguistic nor archeological evidence. Beginning with this proto-Indo-European, linguists traced the historical evolution of languages, comparing words from different languages, matching them syllable by syllable. Their enormous toil, spread over two centuries, has shown that the Indo-European family evolved into two major branches, Indo-Iranian and Indo-Aryan. The Indo-Aryan, in turn, has interacted with extant languages in India to produce the modern languages of Pakistan, India, Nepal, Bangladesh, and Sri Lanka.

|| The answer to the question as to which came first in India, the Indo-Aryan or the Dravidian, has remained in a kind of animated suspension.

This scholarship also gave rise to study of Tamil and other languages spoken in south India. Some decades after the Indo-Aryan hypothesis was proposed, Alexander Campbell suggested that there could be exceptions in the subcontinent. His work, and that of Francis Ellis, established the non-Sanskrit origin of Tamil and nearly 70 other languages. These Dravidian languages were found to have existed not just in the south but also in distant locations in central and north India, and sites [as far as in Baluchistan](#). The answer to the question as to which came first in India, the Indo-Aryan or the Dravidian, has remained in a kind of animated suspension. To add to the mystery and mythmaking, came the archeological discovery of a culture spread over a large area stretching from the northern bank of Narmada all the way through present-day Pakistan and beyond, described as the Harappan civilisation. Until now, there have not been enough material samples available to help experts decipher the script used by the people who had formed this civilisation.

There have been claims that the ancestors of the people in the south, who spoke ancient Tamil, had clear links with the people of the Harappan civilisation. It has also been argued that Tamil predates Sanskrit in India by several centuries if not by several millennia. The claims are there and have been made with ample justification, but they have not been conclusively proven with archeological evidence. Besides, not knowing anything at all about the language used by the Harappan civilisation during the years 2500 BCE to 1900 BCE leaves everything open about the origin of Indian languages prior to the arrival of the Indo-Aryan, sometime after 1800 BCE. This lack of clarity provides a fertile ground for mythmaking, cultural bravado, and for proposing pseudo-science as history. The unfortunate historical episode of the rise of fascism in Europe under Adolf Hitler’s leadership nine decades before our time, his use of the twisted and falsified ‘history’ generated by Historical Linguistics before his time, his giving currency to terms like ‘Aryan’ and ‘Aryan race’, have all been drilled into the minds of the followers of the Hindutva ideology.

In various historical phases, various languages emerge. They merge with other languages, decline, perish, take on other forms, and continue to exist as substrata of new languages. There is nothing in this process that should evoke either pride or shame. But for any brand of militant nationalism, if pride is a valuable political commodity, so is ‘shame’ arising out of a real or imagined humiliation and calling for historical revenge.

As an Indian I cherish the great language diversity that we have, and I am grateful to our ancestors in India for leaving for us such a wonderful composite linguistic heritage.

The Hindutva ideology of nationalism thrives upon the lack of scientific clarity about the exact genesis of languages in India and the exact nature of exchanges and mutual influence between Tamil, Sanskrit, Pali, and several other Prakrits. It deliberately encourages misconception about a language like Urdu, which took birth in India, partly in Gulbarga in the south and partly in what is today Uttar Pradesh in the north. It requires the silencing of Adivasi languages, their assimilation within a larger language like Hindi, and their dissolution in the Hindu identity. These psychological and political requirements of Hindutva shall still be in place beyond the coronavirus. It is to be expected that the next census will bring to the country language data that will further marginalise the languages of the Adivasis.

I love Sanskrit. It is a majestic language. I also like Hindi, a great language, as much I like my two mother tongues, Gujarati and Marathi. But I like the languages of the Adivasis as well. And I adore English as I do Urdu. As an Indian I cherish the great language diversity that we have, and I am grateful to our ancestors in India for leaving for us such a wonderful composite linguistic heritage. The Constitution of India has accepted language diversity as a noble idea. But if the Census knowingly diminishes the diversity and if the regime consciously tries to set the speakers of one language against the speakers of another language, one has to ask if there is a way to resist these actions.

In my opinion, knowledge is that way. To know as much as we can, using scientific methods and proper evidence, about the past of Indian languages, and to know the exact spread and scope of language diversity in our time, becomes our obligation to India's democracy.

Fortunately, new research is available: in genetics (David Reich, *Who We are & How We Got Here*, 2018), linguistics (David Anthony, *The Horse, The Wheel and Language*, 2007), scripts (Andrew Robinson, *Lost Languages: The Enigma of the World's Undeciphered Scripts*, 2009), and in migration studies (Tanuja Kothiyal, *Nomadic Narratives*, 2016). Ultimately, any shade of fascism can be fought by knowledge alone. Every militant variety of nationalism can be resisted with logic and argument. In historical terms, citizenship rights are a given. In existential terms, it is a matter of bringing facts to bear upon the discourse, again and again. Counting has its consequences and citizens need to stay invested in counting, forever.