Language and Identity in South Asia

MixCache.com

Table of Contents

Introduction

This book begins with a simple observation: a map does not speak, but people do. Across India, Pakistan, Bangladesh, Sri Lanka, and Nepal, spoken and written words have never been mere vehicles of meaning; they are badges of belonging, engines of mobilization, and blueprints for statecraft. The sounds of everyday life—market haggles, courtroom arguments, lullabies, film songs—reveal how identities are made and remade. By following these sounds into schools, parliaments, printing houses, and digital platforms, this book investigates how language formation, script reforms, and the politics of linguistic nationalism have shaped the subcontinent's borders, education systems, and cultural identities.

Languages here are not timeless essences but evolving projects. They are assembled through grammars and dictionaries, consecrated in school curricula, broadcast over radio waves, and encoded into fonts and software. Scripts are not only technical choices; they are visible flags that can unite, divide, or reimagine communities. When activists rally around a script or a standard, they are not simply tidying orthography. They are voting on who counts as a legitimate speaker, which pasts deserve reverence, and what futures are possible.

The story, however, begins before modern nationalism. South Asia's linguistic worlds emerged from layered histories of Sanskrit and Prakrit learning, Persianate administration, and vibrant vernacular literatures. Colonial rule reframed these worlds through surveys, censuses, and print capitalism, often freezing fluid speech continua into named and ranked "languages." The twentieth century then turned classification into contestation: mass politics, partition, and postcolonial state-building pushed linguistic identities to the center of public life. Movements over which language deserved official status, which script marked the community, and what tongue would teach the child were not academic squabbles—they were struggles over dignity, opportunity, and power.

This book traces that arc across emblematic cases. It follows the entwined standardizations of Hindi and Urdu, the making of Bengali as a political cause culminating in a martyrdom remembered across generations, the transformation of Tamil into a language of modern protest and cinema, and the legal and social aftermaths of "Sinhala Only" in Sri Lanka. It examines the reorganization of India's internal boundaries around linguistic lines, the careers of Punjabi across Gurmukhi and Shahmukhi scripts, and the place of Nepali within a multilingual nation. It also centers voices too often muffled in the grand narrative—Adivasi and Indigenous communities, linguistic minorities, and speakers whose gender, caste, or class position alters how and whether they are heard.

Methodologically, the chapters move between archives and streets: from early grammars, missionary tracts, and lexicons to textbooks, court judgments, film dialogue, protest posters, and the everyday paperwork of administration. They treat language as a social infrastructure built out of materials—type metal and wood blocks,

dictation slates and school primers, radio transmitters and studio microphones, dot-matrix printers and Unicode code points. Attention to these materials shows how technical reforms—new spellings, standardized scripts, character encodings—can quietly reorder who can read, write, publish, and participate.

Three themes anchor the inquiry. First, borders speak: political lines—from the Radcliffe demarcation to interstate boundaries—rearrange linguistic traffic but never fully contain it. Second, classrooms make citizens: the medium of instruction, often debated in policy and fought over in the streets, forges pathways of mobility and exclusion. Third, voice is stratified: language movements articulate solidarity and pride, yet they can reproduce hierarchies of gender, caste, class, and religion unless challenged by counterpublics and coalitions. Throughout, we ask how law and policy, markets and media, and technologies old and new shape the terms on which communities imagine themselves through language.

The chapters unfold roughly chronologically while pairing historical depth with contemporary resonance. Early chapters map the transformation from precolonial multilingual ecologies to colonial classifications; middle chapters analyze mass movements, constitutional settlements, and the redrawing of maps; later chapters turn to media industries, courts, and digital infrastructures that are refashioning scripts and standards today. Case studies are comparative by design: each illuminates local textures while speaking to broader patterns across the region.

Language and Identity in South Asia argues that linguistic life is not a backdrop to politics but one of its primary stages. By following how words become institutions and how scripts become symbols, the book offers tools for understanding not only the subcontinent’s past but also its multilingual futures. It invites readers—scholars, students, policymakers, and anyone who has ever argued about a spelling or switched tongues mid-sentence—to hear in South Asia’s languages both the echoes of history and the possibilities of a more capacious citizenship.

CHAPTER ONE: Languages Before the Nation: From Cosmopolis to Region

Before the modern nation-state carved up South Asia into distinct linguistic territories, the subcontinent was a vibrant, often fluid, tapestry of languages. It was a world where linguistic boundaries were porous, and multilingualism was not an exception but the norm. People moved between tongues with an ease that would perhaps startle a contemporary observer, driven by trade, religious pilgrimage, imperial administration, and simply the ebb and flow of everyday life. This pre-national era,

spanning millennia, saw the rise and fall of various lingua francas, the flourishing of diverse literary traditions, and the constant intermingling of speech communities, all contributing to a complex linguistic ecology that defied rigid categorization.

At the heart of this ancient linguistic landscape lay the profound influence of Sanskrit. Emerging in northwest South Asia during the late Bronze Age, Sanskrit quickly ascended to become the sacred language of Hinduism, classical Hindu philosophy, and a liturgical language for Buddhism and Jainism. It wasn't merely a language of religious texts like the Vedas, Upanishads, Mahabharata, and Ramayana, but also a sophisticated medium for literature, education, and high culture. For centuries, Sanskrit served as a lingua franca across ancient and medieval South Asia, and its reach extended far beyond, impacting languages in Southeast and Central Asia as Hindu and Buddhist cultures spread. The concept of a "Sanskrit Cosmopolis," coined by historian Sheldon Pollock, describes how Sanskrit diffused not through conquest of populations, but as a language of prestige and complex thought, adopted voluntarily by local elites across a vast region.

Alongside Sanskrit, a family of languages known as Prakrits also flourished. These were the more vernacular, colloquial dialects of ancient India, often simpler and more accessible than the highly formalized Sanskrit. While Sanskrit was refined and codified by grammarians, notably Pāṇini in the 4th century BCE, Prakrits continued to evolve naturally, reflecting the diverse speech of common people. Early Buddhist and Jain literature, for instance, often utilized Prakrit languages like Pali and Ardhamagadhi, making religious teachings accessible to a wider audience. The relationship between Sanskrit and Prakrit was not one of simple parent and child; rather, they influenced each other, with Prakrits even contributing to the vocabulary and features of post-Vedic Sanskrit. Many modern Indo-Aryan languages, including Hindi, Bengali, Marathi, and Punjabi, are rooted in these Prakrit languages.

Beyond the Indo-Aryan family, South Asia was and remains a mosaic of other distinct language families. The Dravidian languages, predominantly spoken in southern India and Sri Lanka, represent a linguistic heritage unique in their phonology, grammar, and structure. Major Dravidian languages include Tamil, Telugu, Kannada, and Malayalam. While some linguistic similarities developed between Dravidian and Indo-Aryan languages over centuries of interaction, their origins are entirely separate. There are also Tibeto-Burman languages, spoken in the Himalayan and Indo-Burman ranges, and Munda languages, part of the Austroasiatic family, which are considered some of the oldest languages in the region. The subcontinent even boasts rare language isolates like Burushaski, Kusunda, Nihali, and Vedda, whose origins remain a subject of ongoing study and linguistic mystery.

The advent of various Turkic and Afghan dynasties from the 11th century onwards introduced another powerful linguistic force: Persian. Persian rapidly became the language of administration, court, literature, and social status across much of the

Indian subcontinent, largely replacing Sanskrit in these domains. This wasn't merely a language imposed by rulers; Persian was embraced by many local Indian dynasties, both Muslim and non-Muslim, including the Sikh Empire. It acted as a crucial lingua franca, connecting diverse people and integrating the Indian subcontinent into the broader Persianate world. Evidence of its widespread use can be seen in the fact that in the 17th century, a Marathi ruler like Shivaji would communicate with a Rajasthani general of the Mughal army using Persian.

The influence of Persian was profound, leaving a lasting impact on Indo-Aryan languages, playing a formative role in the emergence of Hindustani, and significantly influencing Punjabi, Sindhi, Bengali, Gujarati, and Kashmiri. The vocabulary of many regional tongues was reshaped by Persian, much as English would later do. This linguistic legacy is particularly evident in the development of Urdu, which is essentially an amalgamation of a Khariboli linguistic base with substantial Persian elements, written in the Perso-Arabic alphabet. For several centuries, Persian held sway as a language of high culture and intellectual exchange, with a vibrant tradition of Persian literature flourishing in India, often surpassing that of Iran itself in lexicography.

The pre-colonial era also witnessed the rich development of vernacular literatures in a multitude of regional languages. While Sanskrit and Persian often held the prestige of court and scholarship, languages like Tamil, Kannada, Bengali, Gujarati, and Marathi saw their own literary traditions flourish, often rooted in devotional poetry and folk narratives. The Tamil Sangam literature, Vachana Sahitya in Kannada, and the Charyapadas in Bengali are just a few examples of the vibrant regional literary output that predated modern linguistic categorization. These literatures were not isolated but often drew themes and forms from Sanskrit, Pali, and Prakrit works, demonstrating a continuous process of cultural and linguistic exchange.

Multilingualism was not simply a feature of elite circles but was woven into the fabric of everyday life. People routinely code-switched and blended languages for practical reasons, a phenomenon still observable today in hybrid forms like "Hinglish." The fluidity of these linguistic interactions meant that speech communities were rarely homogenous or strictly bounded. Dialects and regional variations often formed continua, gradually shifting from one form to another across geographical space, rather than adhering to sharply defined linguistic borders. This nuanced reality would later pose a significant challenge to colonial attempts to categorize and enumerate "languages."

The linguistic landscape of South Asia, therefore, was a dynamic, layered system. It was a place where "language" as a fixed, bounded entity was less of a reality and more of a practical abstraction. The notion of a language as a "badge of belonging" was nascent, evolving more along lines of religious or social community than a modern ethno-linguistic identity. Different languages served different social functions: a courtier might use Persian for administration, Sanskrit for religious ritual, and a local

vernacular for everyday conversation. This functional multilingualism created a rich linguistic ecology where languages coexisted, borrowed from one another, and collectively contributed to the subcontinent's diverse cultural heritage.

This complex pre-national linguistic world, with its overlapping spheres of Sanskrit, Persian, and burgeoning vernacular traditions, would eventually be confronted by the rigid classificatory impulses of colonial administration. The fluid continua of speech, the functional multilingualism, and the locally defined identities would be subjected to surveys, censuses, and educational policies that sought to impose order and, in doing so, inadvertently laid the groundwork for future linguistic nationalisms. The transition from a "cosmopolis" where languages traversed vast cultural spaces, to a "region" where they would increasingly be tied to defined territorial and political identities, was a gradual but ultimately transformative process. This shift, however, was not without its echoes of the past, as the deep historical layers of linguistic interaction continued to resonate within the newly emerging linguistic consciousness.
