

Linguistic Diversity in India: Tracing Roots and Evolution of Major Indian Languages

Vinayak B. Sutar

Lecturer, Dept. of English, KLE's G.I. Bagewadi Arts, Science and Commerce College, Nipani.

DOI: <https://doi.org/10.5281/zenodo.18326906>

ABSTRACT:

India is among the most linguistically diverse countries globally, home to hundreds of languages and dialects shaped by ancient migrations, invasions, trade routes, cultural interactions, and regional geography. Recent estimates place the number of living languages in India between approximately (including dialects and mother tongues, per the People's Linguistic Survey of India), positioning the nation as a remarkable hub of linguistic variety where Indo-Aryan, Dravidian, Austroasiatic, Sino-Tibetan, and other families intersect.

The Indo-Aryan group, part of the broader Indo-European family, dominates northern, central, and western regions. It originated with the arrival of Indo-Aryan speakers in the subcontinent around 1500 BCE, introducing Vedic Sanskrit—the archaic language preserved in sacred texts such as the Rigveda. Panini's seminal grammar, the *Ashtadhyayi*, codified Classical Sanskrit, which later gave way to spoken Middle Indo-Aryan forms: Prakrits (everyday vernaculars) and Apabhramshas. These transitional stages eventually gave rise to contemporary languages like Hindi (rooted in the Khadiboli dialect, now used by roughly 345 million native speakers and over 600 million total including second-language users) and Marathi (evolving from Maharashtri Prakrit, with strong literary traditions under dynasties like the Yadavas and Marathas). These languages show evolutionary patterns such as phonological streamlining (e.g., reduced consonant clusters and prominent retroflex sounds), simplified morphology with postpositions replacing cases, and consistent subject-object-verb word order, alongside borrowings from neighboring families and later colonial influences.

Conversely, Dravidian languages, indigenous to southern India, form a distinct family with proto-Dravidian origins

potentially stretching back 4,500–5,000 years. They evolved largely independently, preserving native grammatical structures and core lexicon despite gradual Sanskrit and Prakrit loans through cultural contact. Kannada, a prominent South Dravidian member with about 44 million native speakers (and additional millions as a second language), exemplifies this resilience. Its earliest records appear in inscriptions like Halmidi (around 450 CE), followed by literary flourishing in texts such as Kavirajamarga, under empires including the Kadambas, Rashtrakutas, and Vijayanagara.

This overview explores the origins, developmental paths, literary and social advancements, and present-day roles of key languages: Sanskrit, Hindi, Marathi, and Kannada, while underscoring mutual influences. Ultimately, India’s linguistic mosaic reflects a dynamic blend of indigenous foundations, migratory inputs, and ongoing exchanges, cementing its status as a living repository or “museum” of global linguistic heritage.

KEYWORDS:

Indigenous, Linguistic, Dialect, Evolution, Cultural Diversity.

.....

Introduction

The Indian linguistic landscape is shaped by ancient traditions, oral literature, documented texts, and socio-political, cultural, and literary forces. This linguistic study categorizes Indian languages into major language families based on historical lineage. The major Indian languages—Sanskrit, Hindi, Marathi, and Kannada—represent diverse linguistic families and different evolutionary paths. Sanskrit, an ancient Indo-Aryan language, is a foundational root for many North Indian languages like Hindi and Marathi, whereas Kannada, in contrast, belongs to Dravidian.

Origin and Evolution of Indian Languages

Antiquity of Sanskrit:

Indo-Aryan is a branch of the Indo-European language family and is spoken by the majority of the Indian population. These languages evolved from Vedic Sanskrit, through Prakrit, to Hindi and Marathi. Vedic Sanskrit was brought to the Indian subcontinent by Aryans around 1500 BCE or earlier.

Sanskrit is one of the archaic languages in Indo-European, a hypothetical ancestor language spoken around 6500 BCE to 3500 BCE. It evolved from Indo-European; its predecessors were Old Aryans migrating around 2000–1500 BCE. These migrants originated from Central Asia, entering North West India (modern Northern Pakistan and Northern India). The term ‘Sanskrit’ means refined or perfected, reflecting its polished grammatical structure. This is the earliest form preserved in Vedas, ancient hymns, and texts like the Rigveda (composed around 1500–2000 BCE). It was an oral language used by Aryan tribes for rituals and philosophy, featuring flexible grammar. Around 500 BCE, it was standardized by the grammarian Panini in his work *Ashtadhyayi*, which is a treasure of codified syntax rules and morphology using about 4000 sutras. Panini’s grammar influenced later commentators like Katyayana and Patanjali. It is used in the composition of classical works like Ramayana and Mahabharata and in the works of poets

like Kalidasa in Shakuntala (5th/6th century CE).

Over time, Sanskrit transitioned from vernacular to scholarly, while Prakrit ultimately evolved into other languages like Hindi, Marathi, Bengali, and Gujarati. It also influenced Dravidian languages like Kannada, etc. It profoundly shaped Indian culture, serving as the medium of religious texts of Hinduism, Jainism, and Buddhism. Today it is not a vernacular language but is present in rituals and scholarly education in India.

Early Roots and Evolution of Hindi:

The Hindi language has a rich history. It is a part of the Indo-Aryan branch of the Indo-European family, spanning from 1500 BCE. It traces its roots to Vedic Sanskrit. Over time, it evolved from the Middle Indo-Aryan language, Prakrit—meaning ‘natural’ language. Prakrit was used as a vernacular, and Sanskrit remained a language of elites. In the 6th and 7th centuries, Prakrit was simplified and evolved into a new version called Apabhramsha, a dialect used in the Meerut and Delhi regions. It was believed to be the early version of Hindi or Hindavi, a Delhi dialect. This period saw rich growth in literature. Its dialects were Braj Bhasha, Awadhi, and Khariboli. Later, the same dialect was used as a lingua franca.

During the modern period, efforts were made to standardize this language. Many scholars and reformers pushed it with Sanskrit-derived terms, and the same became Standard Hindi. In modern days, Standard Hindi, based on Khariboli, is spoken by over 500–600 million people as a first or second language. Its journey from Sanskrit to Prakrit to Hindustani reflects India’s complex history, cultural diversity, and synthesis. Over the years, Hindi evolved in many respects; to standardize phonologically, it lost its complex consonant clusters and emerged with retroflex sounds. At the morphological level, it reduced its case endings and developed postpositions. Syntactically, it shifted to an SOV (Subject–Object–Verb) structure.

Linguistic Origin of Marathi:

Marathi, an Indo–Aryan language primarily spoken in Maharashtra, India, and by the Marathi community worldwide, has a history spanning over a millennium. It descends from Maharashtri Prakrit, a standard Prakrit version spoken in the western and Deccan regions of ancient India. During Satavahana rule (2nd century BCE), Prakrit was an official language.

In the course of time, Maharashtri Prakrit developed into Maharashtri Apabhramsha, a corrupt form which bridges Middle Indo–Aryan and early New Indo–Aryan languages. This process involved phonetic, grammatical, and lexical changes. Alongside, it was also influenced by Dravidian languages like Kannada and Telugu. The language emerged around the 10th and 11th centuries, including inscriptions on stones and copper plates found after 1187 CE under the Yadava dynasty of Devagiri (1187–1317 CE). The Yadavas shifted from Kannada to Marathi as the court language. During the early period, literature includes religious texts like Mahanubhava Panth, prose works like Lilacharitra (1275 CE), and the Warkari tradition. The middle period witnessed tremendous literary growth by great Maharashtrian saints like Dnyaneshwar in his Dnyaneshwari (a commentary on the Bhagavad Gita). Namdev, Eknath, and Tukaram, etc., composed Abhangas (devotional poetry), making it a vehicle of spiritual poetry. During the reign of Chhatrapati Shivaji Maharaj (17th century), Marathi was promoted as an administrative language.

In the modern period (19th to 20th century), many scholars tried to standardize it, producing many grammar books, dictionaries, and translations. Reformers like Tilak, Gokhale, and Phule modernized it by using it as a language of newspapers. Today, it incorporates English loan words and continues to evolve in mass media, education, and digital contexts.

Antiquity of Kannada:

Kannada is a Dravidian language spoken in the state of

Karnataka, India, with around 44 million native speakers and additional millions using it as a second language. It is one of the classical Indian languages, second only to Tamil. It belongs to the South Dravidian branch, which includes Tamil, Telugu, Kannada, Malayalam, and Tulu, etc. It descends from Proto-Dravidian languages believed to have been spoken around 4500–5000 years ago. It evolved within the Tamil and Kannada subgroup, sharing core grammar and vocabulary with Tamil-related languages, but it developed independently. Sanskrit and Prakrit influenced it, but it retained strong native Dravidian roots using purely native words and structures. Some scholars claim its origin existed in oral form by the 3rd century BCE, mentioned in Ashokan edicts in place names like Isila.

Proto-Old Kannada:

The earliest written evidence appeared in the Halmidi inscriptions (450 CE), and the oldest full-length Kannada inscriptions were found in Karnataka. It shows Kannada as the old administrative language of the Kadambas, who ruled Banavasi around 345 CE. The script evolved from the Southern variety of the Ashokan Brahmi script, developed into the Kadamba script in the 5th century, which became the modern Kannada script.

Halegannada:

Literary Kannada flourished under empires like the Rashtrakutas, 9th-century Western Gangas, Chalukyas, and Hoysalas. The earliest major literary work is Kavirajamarga (850 CE), a treatise on poetics by King Nripatunga. Jain and Hindu influences dominated with poets like Pampa, Ranna, and Ponna, the “Three Jewels” of the 10th century, producing epics. This period saw maturity in grammar, vocabulary, and style like Champu.

Nadugannada:

Grammatical and phonological changes occurred, influenced by Vachana literature of the 12th century, and by the devotional

poetry of Basavanna and others during the Lingayat movement. This saw an increase in Sanskrit borrowings and regional variations under the Vijayanagara Empire. During the modern period (17th century to present day), Kannada standardized and developed with contributions from writers like Kuvempu, Shivaram Karanth, etc. It preserved classical roots and evolved as the official language of Karnataka. Kannada's journey reflects resilience through dynasty patronages, religious movements, and cultural exchanges, making it a vibrant classical language with deep historical roots in South India.

Conclusion

India's linguistic diversity stands as one of the most profound examples of cultural and historical richness in the world, with an estimated 450 to 800 distinct languages spoken across the subcontinent, making it the fourth most linguistically diverse country globally. This feature reflects millennia of migrations, invasions, trade, and geographical isolation fostering a unique convergence of language families.

Tracing the historical evolution reveals indigenous roots of the Indo-Aryan and Dravidian families fragmenting into branches like Tamil, Telugu, Kannada, Malayalam, and Tulu around the 3rd millennium BCE. In conclusion, India's linguistic evolution underscores a dynamic interplay of indigenous roots, migratory influxes, and cultural exchanges, transforming India into a museum of languages, where diversity serves as a cornerstone of national identity.

References:

1. Devy, G.N. India, A Linguistic Civilization.
2. Sarangi, A. (ed). Language and Politics in India. Oxford University Press.
3. Sen, S. Kumar. Bhashar Itibritta – A Comprehensive Documentation of Indian Languages.
4. Grierson, G.A. (1927). Linguistic Survey of India, Central Publication Branch.
5. Chandramouli, S.C. (2014). Linguistic Survey of India: A Perspective.
6. Kulkarni-Joshi, S. (2019). Linguistic History and Language Diversity in India.
7. Springer Link Article (2019). Linguistic History of India.
8. Journal of Scientometric Research (2024). Documenting Tribal and Indigenous Languages in India.

Funding:

This study was not funded by any grant.

Conflict of interest:

The Authors have no conflict of interest to declare that they are relevant to the content of this article.

About the License:

© The Authors 2024. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.