



Received: 08/March/2025

IJRAW: 2025; 4(4):113-116

Accepted: 16/April/2025

Language Recognition of Indian Sub-Continent

^{*1}Bhumika Patel and ²Dr. Shweta Gohel

^{*1}Lecturer, GMB Polytechnic, Rajula, Gujarat, India.

²Dean, School of Arts and Humanities, Junagadh, Gujarat, India.

Abstract

South Asia, encompassing Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka, hosts several hundred languages. Prominent among these are Hindi-Urdu and Bengali, the third and sixth most spoken languages globally, respectively. The region's languages primarily include Indo-Iranic and Dravidian languages, along with Austroasiatic and Tibeto-Burman families. English serves as the international lingua franca, significantly influencing South Asian languages since the colonial era. A notable result of this influence is Hinglish, an English-influenced variant of Hindustani, widely spoken in urban areas.

Keywords: Language-recognition, Indian Sub-continent, South Asia.

Introduction

Languages of South Asia:

The South Asian region, having nations as mentioned above, is characterized by an immense linguistic diversity, with several hundred distinct languages spoken across these countries. Among these languages, Hindi-Urdu, often collectively referred to as Hindustani, holds a particularly prominent status, being the third most spoken language globally. Similarly, Bengali ranks sixth in terms of the number of speakers worldwide.

The majority of languages spoken throughout South Asia belong primarily to the Indo-Iranic and Dravidian language families. However, the linguistic landscape is even more varied, also including languages from other families such as Austroasiatic and Tibeto-Burman.

English plays a vital role in the region and is widely regarded as the international lingua franca among South Asian countries. Its influence has been especially strong since the colonial period, during which it became deeply embedded in the sociolinguistic fabric of the subcontinent. As a result, many South Asian languages have incorporated elements of English vocabulary and expressions. This influence is particularly evident in Hindustani, where a hybridized version known as "Hinglish" has emerged. Hinglish blends Hindi (and Urdu) with English and is especially prevalent in urban centers, reflecting both cultural shifts and the continuing evolution of language use in modern South Asia.

Geographic Distribution of Languages

Geo-linguistically, the Indian subcontinent predominantly features Indo-Aryan, Dravidian, and Munda language groups,

collectively known as Indic languages. The region also includes language isolates like Burushaski, Kusunda, Nihali, and Vedda. The linguistic influence of the subcontinent extends to neighbouring Southern, East, and Southeast Asian regions, forming an extended linguistic area known as the Indo sphere. The sprachbund of Indic and neighbouring languages is referred to as South Asian languages, which also encompass Eastern-Iranic, Nuristani, Central, and Western Tibeto-Burman languages.

The Iranian Plateau to the west of the subcontinent hosts Iranic languages, including Pashto in Pashtunistan and Balochi in Balochistan. Kafiri languages are spoken in northern pockets of the plateau and the subcontinent. The Trans-Himalayan Tibeto-Burman languages and Khasi-Palaungic languages of the Austroasiatic family are found in the Himalayan and Indo-Burman Ranges, mainly on the Tibetan Plateau and in Burma. Andamanese languages are spoken on the Andaman Islands.

Distribution by Country

Afghanistan:

- **Official Languages:** Pashto and Dari (Farsi), both Iranic languages.

Dari

- Afghan standardized register of Persian.
- Considered the lingua franca of Afghanistan.
- Used in Afghan literature.

Tajik

- Spoken by people near Tajikistan.

- Officially regarded as the same as Dari.

Pashto

- Widely spoken by Pashtun people.
- Mainly reside in the south, near the Pakistani-Afghan border.

Turkic Languages

- Uzbek and Turkmen spoken near Uzbekistan and Turkmenistan.

Bangladesh

- **National Language:** Standard Bengali (based on the Rarhi dialect).
- **Majority Language:** Eastern Bengali.
- **Native Languages:** Sylheti and Chittagonian.
- **Ethnic Minority Languages:** Tibeto-Burman, Dravidian, and Austro-Asiatic languages.

Bhutan

- **National Language:** Dzongkha.
- **Other Languages Spoken:** Brokpa, Dzala, Chali, Chocangacakha, Dakpa, Khengkha, Nepali, Gongduk, Nyenkha, Lhokpu, Takpa, Tshangla.
- **Language Families:** Mostly Tibetic languages./Nepali (Indo-Aryan language).

India

Language Families

- Indo-Aryan, Dravidian, Austroasiatic (Munda), Tibeto-Burman.
- Some Himalayan languages remain unclassified.
- 461 living languages listed by SIL Ethnologue.

Most Widespread Language

- Hindustani.
- "Hindi" includes a broad variety of Hindi languages, with native speakers accounting for 39% of the population.

Second Most Spoken Language

- Bengali, found in Bangladesh and Indian states of West Bengal, Tripura, and Assam.

International Mother Language Day

- Created by UNESCO to commemorate the Bengali language.

Notable Languages

- Odia, Telugu, Punjabi, Marathi, Tamil, Urdu, Sindhi, Kannada, Pashto, Malayalam, Maithili, Meitei (Manipuri), Konkani, Tulu.
- Thirteen languages account for over 95% of the population; all are "scheduled languages of the Constitution."

Scheduled Languages Spoken by Less Than 1%

- Santali, Meitei (Manipuri), Bodo, Dogri.

Largest Non-Scheduled Languages

- Bhili, followed by Gondi, Tulu, and Kurukh.

Maldives

- **National Language:** Divehi, spoken by 95% of the population.

- **Religious Language:** Arabic.

- **Language of Instruction and International Use:** English.

Nepal

- **Language Families:** Indo-Aryan and Sino-Tibetan.
- **Official Language:** Nepali (formerly Gorkhali), spoken by the majority.
- **Indo-Aryan Languages in Nepal:** Maithili, Bhojpur, Tharu (southern Nepal in the Terai region).
- **Sino-Tibetan Languages in Nepal:** Tamang, Newari, Magar, Gurung, Kiranti, Sherpa (central and northern regions).

Pakistan

- **Major Language Families:** Indo-Iranian (western regions: Iranic languages; eastern regions: Indo-Aryan languages).
- **Other Language Families:** Dravidian (Brahui in Central Balochistan), Sino-Tibetan (Balti, Purgi in the north-east), Nuristani (Kamkata-vari in the north-west), Language Isolate (Burushaski in the north), Turkic (Kyrgyz, Uzbeks, Turkmen).
- **National Language:** Urdu, a Persianized register of Hindustani.
- **Major Native Languages:** Punjabi, Saraiki, Sindhi, Baluchi, Pashto.
- **Other Languages:** Over 70 languages including Shina, Balti, Gujarati, Bengali.

Sri Lanka

- **Official Languages:** Sinhala and Tamil.
- **Link Language:** English.
- **Language Families:** Tamil (South-Dravidian), Sinhala (Insular Indic, related to Dhivehi of Maldives).
- **Indigenous Language:** Vedda, predating Indo-Aryan and Dravidian arrivals.

Language Family of South Asia Post Partition:

Ethno Linguistically, South Asia's languages reflect the region's complex history and geography, forming a diverse patchwork of language families, phyla, and isolates. In India, languages belong to several families, with the major ones being Indo-Aryan, spoken by 78.05% of the population, and Dravidian, spoken by 19.64%. These two families are the most significant in terms of the number of speakers, highlighting the linguistic richness and diversity of the Indian subcontinent.

South Asia is one of the most linguistically diverse regions in the world, with its languages belonging to a variety of distinct language families. The most dominant among them is the Indo-Iranian branch of the larger Indo-European family, which includes the Indo-Aryan languages spoken by the majority population in countries such as India, Pakistan, Bangladesh, Nepal, and Sri Lanka. Prominent Indo-Aryan languages include Hindi, Urdu, Bengali, Punjabi, Marathi, Gujarati, Sindhi, Sinhala, and Nepali. Another subgroup within the Indo-Iranian family is the Iranian branch, which includes languages such as Pashto, Balochi, and Dari, spoken primarily in Pakistan and Afghanistan.

The Dravidian language family is primarily spoken in southern India and parts of Sri Lanka. It includes major regional languages like Tamil, Telugu, Kannada, and Malayalam, along with several smaller languages such as

Tulu. These languages have a long literary tradition and are culturally significant in their respective regions.

Another important but less widespread family is the Austroasiatic family, particularly its Munda branch, which is spoken by indigenous tribal populations in central and eastern India, as well as small communities in Bangladesh and Nepal. Examples of Munda languages include Santali, Mundari, and Ho. The Khasi language of north eastern India also belongs to the Austroasiatic family but falls under a different branch.

In the Himalayan and north eastern regions of South Asia, the Tibeto-Burman languages (part of the Sino-Tibetan family) are spoken. These include languages such as Tibetan, Bodo, Meitei (Manipuri), Lepcha, and Dzongkha, the national language of Bhutan. These languages are typically spoken by communities residing in the hills and mountainous areas, extending into Nepal, north eastern India, and parts of Bhutan and Myanmar.

Aside from these major families, South Asia also has languages from smaller or isolated families. The Andamanese languages, spoken by indigenous groups in the Andaman Islands, represent some of the world's most endangered linguistic traditions. The Tai-Kadai family, though mostly found in Southeast Asia, includes extinct languages like Ahom once spoken in Assam. The Nuristani languages, spoken in eastern Afghanistan, form a small and unique subgroup within the Indo-Iranian family.

Additionally, English serves as a major link language or lingua franca across the region, particularly in urban centers, education, administration, and international communication. While it is not native to the region, its influence has been significant since colonial times. Classical languages such as Sanskrit and Pali are no longer spoken in everyday life but continue to hold religious, cultural, and scholarly importance. Arabic is also used among Muslim communities, mainly in religious contexts. This wide array of language families and the interaction between them contribute to the rich linguistic heritage of South Asia.

i). Indo-European Language Family in India:

- **Largest Language Family:**
 - Indo-Aryan, a branch of the Indo-Iranian family.
 - Indo-Iranian is the easternmost subfamily of the Indo-European language family.
- **Prevalence:**
 - Indo-Aryan languages account for about 1,035 million speakers, over 76.5% of India's population (2018 estimate).
- **Most Widely Spoken Indo-Aryan Languages:**
 - Hindi, Bengali, Marathi, Urdu, Gujarati, Punjabi, Kashmiri, Marwari, Sindhi, Assamese (Asamiya), Maithili, Odia.
- **Other Indo-European Languages:**
 - English is prominent as a lingua franca.

ii). Dravidian Language Family:

- **Size and Distribution:**
 - Second largest language family in India.
 - Accounts for approximately 277 million speakers (20.5% of the population, 2018 estimate).
 - Spoken mainly in southern India, parts of eastern and central India, and northeastern Sri Lanka.
 - Also found in parts of Pakistan, Nepal, and Bangladesh.

- **Major Dravidian Languages:** Telugu, Tamil, Kannada, Malayalam.
- **Speakers Beyond Mainstream Population:** Also spoken by small scheduled tribes like the Oraon and Gond.
- **Languages Exclusively Spoken Outside India:**
 - Brahui in Balochistan, Pakistan.
 - Dhangar (a dialect of Kurukh) in Nepal.

iii). Austroasiatic Language Family:

- **Prevalence:** Smaller language family with about 10 million speakers (approximately 3% of the population).
- **Geographic Distribution:** Native to Southeast Asia, with migration to India.
- **Austroasiatic Languages in India:**
- **Mainland India:** Khasi and Munda languages, including Bhumij and Santali.
- **Nicobar Islands:** Also part of the Austroasiatic family.
- **Status:** Except for Khasi and Santali, most Austroasiatic languages in India are endangered.

iv). Sino-Tibetan Language Family:

- **Size and Scope:**
 - Comprises more than 400 languages.
 - Second largest language family by number of native speakers, with around 1.4 billion speakers.
- **Major Languages:**
 - **Sinitic Languages:** Predominant, with approximately 1.3 billion speakers.
 - **Burmese:** About 33 million speakers.
 - **Tibetic Languages:** Around 6 million speakers.
- **Geographic Distribution:**
 - Spoken across the Himalayas, Southeast Asian Massif, and the eastern edge of the Tibetan Plateau.
 - Many languages are spoken in remote mountain areas with small, poorly documented speech communities.
- **Linguistic Research**
 - Several subgroups have been reconstructed, but a comprehensive proto-language for the entire family is still under research.
 - Traditionally divided into Sinitic (Chinese languages) and Tibeto-Burman branches, though a common origin for non-Sinitic languages has not been established.
- **Controversial Inclusions**
 - Kra-Dai and Hmong-Mien languages are included by some Chinese linguists but are generally excluded by the international community.
 - Various links to other language families have been proposed but lack broad acceptance.

v). Tibeto-Burman Language Family:

- **Representation in India:**
 - Well-represented across the Indian subcontinent.

- Relationships among languages are complex and not easily discernible, described as a "patch of leaves on the forest floor" rather than a "family tree."
- **Notable Scholar:** Suniti Kumar Chatterjee, a Padma Vibhushan awardee, highlighted Meitei (Manipuri) as particularly significant in literature compared to other Tibeto-Burman languages like Newari.

- **Geographic Distribution in India:**

- Spoken across the Himalayas and north eastern regions, including: Arunachal Pradesh, Assam (hills and autonomous councils), Himachal Pradesh, Ladakh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura, West Bengal

- **Sino-Tibetan Languages in India:**

- **Constitutionally Recognized Official Languages:** Meitei (Manipuri), Bodo
- **Non-Scheduled Languages:** Karbi, Lepcha and Various Tibetic, West Himalayish, Tani, Brahmaputran, Angami-Pochuri, Tangkhul, Zeme, Kukish sub-linguistic branches, among others.

vi). Tai-Kadai Language Family:

- **Historical Context:**

- The Ahom language, a South western Tai language, was once dominant in the Ahom Kingdom in modern-day Assam but has since been replaced by Assamese.

- **Current Distribution:**

- Small Tai communities and languages are present in Assam and Arunachal Pradesh.
- Notable Tai languages in these regions include Tai Phake, Tai Aiton, and Tai Khamti.

- **Linguistic Similarities:**

- These Tai languages are similar to: Shan (Shan State, Myanmar), Dai (Yunnan, China), Lao (Laos), Thai (Thailand), Zhuang (Guangxi, China)

vii). Great Andamanese Language Family:

- **Great Andamanese Languages:**

- **Extinct Languages:** Several extinct languages within this group.
- **Endangered Language:** Aka-Jeru (highly endangered).

- **Ongan Language Family:**

- **Extant Languages:** Önge and Jarawa.
- **Extinct Language:** Jangil.

- **Sentinelese:** Likely related to the Great Andamanese and Ongan languages.

viii). Language Isolates in South Asia:

- **Indian Mainland:**

- **Nihali:** Considered a language isolate with ambiguous status. It has been classified as a

distinct Austroasiatic language, a dialect of Korku, or even a "thieves' argot."

- **South Asia (Other Regions):**

- **Burushaski:** Spoken in Gilgit-Baltistan (administered by Pakistan).
- **Kusunda:** Found in western Nepal.
- **Vedda:** Spoken in Sri Lanka.
- **Great Andamanese:** The status as a language family is debated, with some considering it a language isolate.

- **Bantu Language:**

- **Sidi:** A Bantu language spoken in Gujarat by the Siddi until the mid-20th century.

References

1. "50th Report of the Commissioner for Linguistic Minorities in India (July 2012 to June 2013)" (PDF). Commissioner for Linguistic Minorities, Ministry of Minority Affairs, Government of India. Archived from the original (PDF) on 26 December 2014. Retrieved 17 September 2016.
2. "Constitution of India". Archived from the original on 2 April 2012. Retrieved 21 March 2012.
3. "Official Language - The Union-Profile - Know India: National Portal of India". Archive.india.gov.in. Retrieved 28 December 2017.
4. "Official Language Act | Government of India, Ministry of Electronics and Information Technology". meity.gov.in. Retrieved 24 January 2017.
5. Salzmänn, Zdenek; Stanlaw, James; Adachi, Nobuko (8 July 2014). *Language, Culture, and Society: An Introduction to Linguistic Anthropology*. Westview Press. ISBN 9780813349558-via Google Books.
6. "Indo-Aryan languages". Encyclopædia Britannica. Retrieved 10 December 2014.
7. "Hindi languages". Encyclopædia Britannica. Retrieved 10 December 2014.