

**SPATIAL DISTRIBUTION OF LANGUAGES AND DIALECTS IN**  
**INDIA**

**Component-I (A) - Personal Details**

<b>Role</b>	<b>Name</b>	<b>Affiliation</b>
Principal Investigator	<b>Prof. Masood Ahsan Siddiqui</b>	Department of Geography, Jamia Millia Islamia, New Delhi
Paper Coordinator, if any		
Content Writer/Author (CW)	<b>Dr. Taruna Bansal</b>	Department of Geography, Jamia Millia Islamia, New Delhi
Content Reviewer (CR)		
Language Editor (LE)		

**Component-I (B) - Description of Module**

<b>Items</b>	<b>Description of Module</b>
Subject Name	Geography
Paper Name	Social Geography
Module Name/Title	Spatial Distribution of Languages and Dialects in India
Module Id	GEOG/SG 25.
Pre-requisites	Distribution of Languages in India, India as a Linguistic Region
Objectives	To study the spatial distribution of Languages and Dialects in India
Keywords	Languages, Dialects, linguistic Regions

## **Component II - e-Text**

### **Spatial Distribution of Languages and Dialects in India**

#### **I. Introduction**

India is often termed as land of 'linguistic pluralism' because numerous languages and dialects are spoken here. But this may not be a correct description as at present the prevailing situation in India is not pluralistic but that of a continuum. One dialect merges into the other almost undetectably; one language replaces the other gradually and along the line of interface between two dialects or languages, there is a zone of transition in which people are bilingual. In other words, languages do not exist in water-tight compartments. Rather there is a state of mutual existence of several dialects or languages in a contiguous space. The nature of linguistic diversity in India was first observed by Sir George Abraham Grierson in his Linguistic Survey of India conducted towards the end of the nineteenth century, where he identified 179 languages and 544 dialects in India. Similarly, the 1961 census recorded 187 languages and it was found that 94 out of the 187 languages were spoken by small populations of 10,000 persons or less. Which means that in 1961, about 97 per cent of the country's population was found affiliated with just 23 languages.

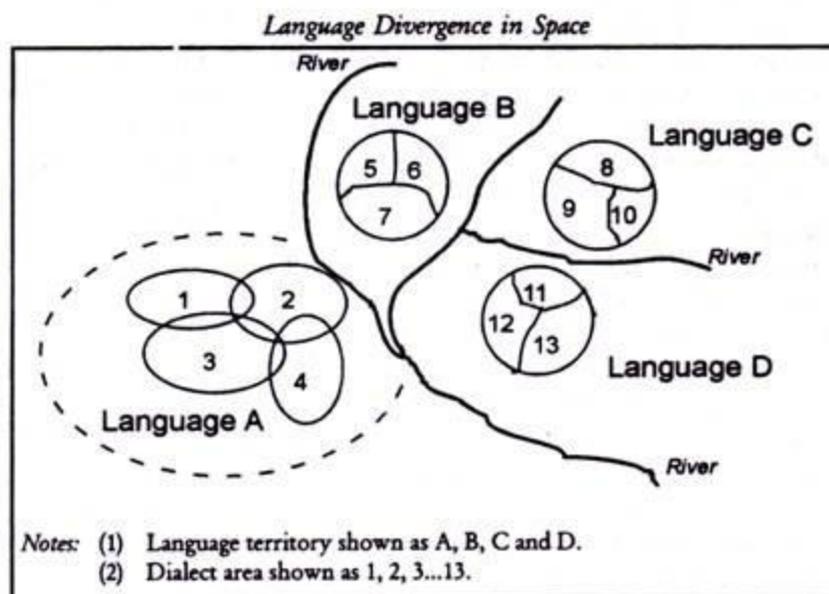
The diversity of dialects and languages is a reality which cannot be ignored. Moreover, it is not the numerical strength of the speakers of a language which is important, what is important is the fact that there are people who claim a certain language as their mother tongue. The development of script is another related progress which contributed to linguistic diversity in India. Different Indian languages were written in different scripts. In other words, with the development of a script, oral communication is supplemented by a more powerful form of written communication.

In the course of time some of the minor dialect and language groups have lost their identity as they have assimilated with developed languages. A study of historical linguistics reveals that India has gone through different phases of language development and the present linguistic map of India is a product of these historical developments.

## II. Language and Dialect:

Human beings use a system of language for communication which distinguishes them from the rest of the animal kingdom. It was through language that communication between the different members of a human group started in the early stages of social evolution. Language thus facilitated multiple forms of human cooperation. A structured language was the invention of the human mind, and the most effective tool of communication. In the course of time, several speech communities are formed, each occupying a chunk of a geographically contiguous space (Fig. 1).

**Figure 1**



Source: Language Divergence in Space, Social Geography by A. Ahmad, Rawat Publications (2004), p. 278

Each language eventually expands over a territory, homogeneous in terms of its language structure—vocal sounds, words, sentences and conventionalized symbols. When a language is written in a script it lends to stabilize its distinguishing features and promotes communication over long distances between people.

## III. India as a Linguistic Area:

India's unity as a socio-linguistic area is quite impressive in spite of the widely perceived linguistic diversity that exists in India. Several linguists have analysed the basic elements of India as a socio-linguistic area. For example, Khubchandani (1975) described language as an

'autonomous system' and recognized the major characteristics of the speech forms of modern India. He opines that each region in the country is characterized by the plurality of not only cultures but of languages too "with a unique mosaic of verbal experience". He stated that India presents a striking example of the process of diffusion through its modern languages be it grammatical or phonetic. Some of the basic elements of India's linguistic unity may be seen in the nature of language boundaries as they are usually fuzzy. Moreover, there exists fluidity in language identity and complementarity of inter-group and intra-group communication.

Here reference may also be made to the seminal work of Murray B. Emeneau (1956) who analysed the characteristics of India as a region of languages. While tracing the process of development of the Indo-Aryan and the Dravidian languages he evaluated the shared experiences of the different speech communities. Emeneau defined linguistic area as "an area which includes languages belonging to more than one family but sharing traits in common which are found not to belong to other members of (at least) one of the families".

#### **a) Geographic Patterning of Languages in India :**

The geographic patterning of languages in the South Asian sub-continent can perhaps be understood in the context of the space relations the region had with other parts of Asia. The overland connections with West and Central Asia, Tibet, China and other regions of South-east Asia helped the process of infiltration of linguistic influences into the South Asian region. This is evident from the fact that the languages spoken in the peripheral regions of South Asia, such as Baluchistan, Pak-Afghan borderlands, Kashmir, and Ladakh as well as in the hilly parts of Himachal Pradesh have strong affinity with the languages spoken in the regions beyond the Hindu-Kush Himalayas. The remote Himalayan regions in the Northeast became the abode of Tibeto-Chinese (Sino-Tibetan) languages and were strongly influenced by the neighbouring parts of Myanmar, Thailand and Indo-China.

The people in the plains of North India from Sind to Assam are dominated by the different branches of the Indo-European family of languages. The peninsular region continued to retain the Dravidian speech-forms even though the north was completely swayed over by the Indo-European languages. Between the Indo-European and the Dravidian one finds the Austric-speaking tribes nestled in the hills of the mid-Indian region.

The linguistic heterogeneity of India can perhaps be brought to some order when one realizes that these speeches really belong to four language families:

- i. Sino-Tibetan (Tibeto-Burman),
- ii. Austro-Asiatic,
- iii. Dravidian and
- iv. Indo-European.

Their geographical patterning throws some light on the routes through which these language families reached India. In fact, despite the vast heterogeneity, Indian languages experienced parallel trends in linguistic and literary development during the long phases of shared history. This has made India 'a composite region' in terms of linguistic attributes.

#### **b) Classification of Indian Languages:**

India is a country characterized with continuity and vastness with great linguistic variations. History shows that over the period people from different periodic line came and settled here resulting in the mixing of cultures. This further led to the amalgamation of their languages as well as dialects. As a result, in present times, people from different cultures, races and classes live in different area of the nation speaking various languages and dialects. These languages can be broadly divided into the following four families. The classification is based upon the number of people speaking languages and dialects in each family (Table 1 and Figure 2):

1. Indo-European Family (Aryan),
2. Dravidian Family (Dravida),
3. Austric Family (Nishada), and
4. Sino Tibetan Family (Kirata).

The most important family in India is the Aryan family, both numerically and culturally. Nearly 73 per cent of the population in India either speaks language or dialect belonging to this family. Nearly 20 percent of the population speaks languages of the Dravidian family. A very small percentage of the population speaks the languages and dialects belonging to Austric and the Sino-Tibetan family.

**Table 1**

**Broad Classification of Modern Indian Languages**

Family	Sub-Family	Branch/Group	Speech Area
AUSTRIC (Nishada)	Austro-Asiatic	Mon-Khmer	Meghalaya, Nicobar Islands
		Munda	West Bengal, Bihar, Orissa, Assam, Madhya Pradesh, Maharashtra
	Austronesian		Outside India
DRAVIDIAN (Dravida)		South Dravidian	Tamil Nadu, Karnataka, Kerala
		Central Dravidian	Andhra Pradesh, Madhya Pradesh, Orissa, Maharashtra
		North Dravidian	Bihar, Orissa, West Bengal, Madhya Pradesh
SINO-TIBETAN (Kirata)	Tibeto-Myanmari	Tibeto-Himalayan	Jammu and Kashmir, Himachal Pradesh, Sikkim,
		North Assam	Arunachal Pradesh
		Assam-Myanmari	Assam, Nagaland, Manipur, Mizoram, Tripura, Meghalaya
	Siamese-Chinese		Outside India
INDO-EUROPEAN (Arya)	Indo-Aryan	Iranian	Outside India
		Dardic	Jammu and Kashmir
		Indo-Aryan	Jammu and Kashmir, Punjab, Himachal Pradesh, Uttar Pradesh, Rajasthan, Haryana, Madhya Pradesh, Bihar, Orissa, West Bengal, Assam, Gujarat, Maharashtra, Goa

Source: Broad Classification of Modern Indian Languages, Social Geography by A. Ahmad, Rawat Publications (2004), p. 282

**1. The Aryan Family:**

The most important of all the families of languages is the Aryan family as it is being spoken by nearly three fourths of the Indian population. This family can be clearly segregated into two branches:

- (i) The Dardic Languages and
- (ii) The Indo-Aryan Languages

**(i) The Dardic Languages:**

Under the Dardic branch of languages are those languages which at present few spoken among the mountain communities in Kashmir. Outside India, these languages are spoken by people residing on the frontiers between Pakistan and Afghanistan. This branch can be further classified into three sub-branches:

- (a) Shina (includes Kashmiri, Shina proper and Kohistani)
- (b) Chatran or Khowar or Chitrali
- (c) Nuristani or Kafuistan dialects

Shina and Kashmiri along with some dialects allied to Kashmiri are spoken in Kashmir and it appears that Kashmiri is the base of Dardic Aryan dialect. In view of most of the scholars, Dardic is mere a branch of Indo-Aryan as these Dardic dialects are largely on the edge of extinction. The only recognised national language of the Indian Union is Kashmiri (spoken by almost 20 lakh people).

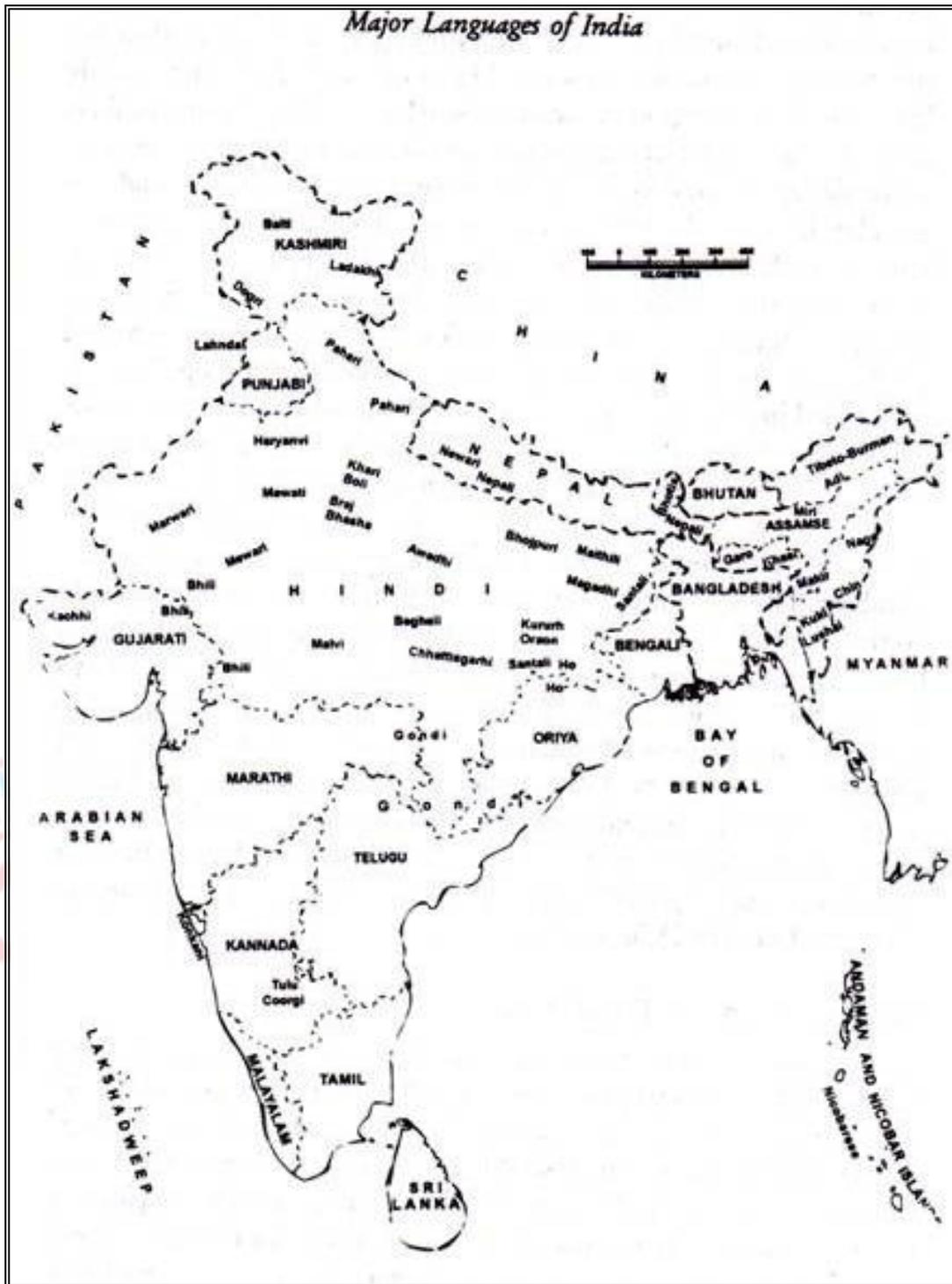
***(ii) The Indo-Aryan Languages:***

The Indo-Aryan branch of the Aryan family includes Hindi, Bengali, Punjabi, Rajasthani, Gujarati, Sindhi, Kachchi, Marathi, Oriya, Sanskrit, Assamese and Urdu. On the basis of the per cent of the population speaking the languages under this branch, they have been classified into the following categories:

- (a) Northern Aryan Languages: Languages of this group belong to the dialects spoken by the hilly people in North India. They include Nepali, Central Pahari and Western Pahari Aryan languages.
- (b) North-Western Aryan Languages: Khanda, Kachchi and Sindhi are the well-known Aryan languages which are spoken by the people living in the north-western part of the country.
- (c) Southern Aryan Languages: Marathi and Konkani are the languages included in the Southern group of Aryan languages.
- (d) Eastern Aryan Languages: The region of these languages lies in the eastern parts of the country, Bihari, Oriya, Bengali and Assamese languages constitute this group of Aryan languages.
- (e) East Central Aryan Languages: Avadh, Bundelkhand and Chhattisgarh regions include these languages as the languages of the people living there. Avadhi, Bugheli and Chhattisgarhi are their languages.
- (f) Central Aryan Languages: The central region of India is the region of Central Aryan languages. The major languages of this region are Hindi, Punjabi, Rajasthani and Alawari.

Among the languages in the Aryan family, Hindi is the most important language, spoken by a huge number of people in the country. In every state of India people speaking Hindi can be found as even the illiterates can speak and communicate in Hindi.

**Figure 2**



Source: Major Languages of India, Social Geography by A. Ahmad, Rawat Publications (2004), p. 307

## 2. The Dravidian Family:

As Dravidians entered India much before the Aryans, the languages in the Dravidian family are considered older than the languages of the Aryan family. In present times, the languages under this family are well-knit among themselves unlike the families of the Aryan, the Austric and the Sino- Tibetan. Another significant characteristic of this family is that it does not have any relation outside the Indian subcontinent. The languages of the Dravidian family can be categorised into several group but usually are branched in to two sub-branches as:

(i) The North Dravidian Languages:

(ii) South Dravidian Languages:

### *(i) The North Dravidian Languages:*

This branch mainly includes Telugu and a number of other dialects like Gondi, Oraon, Malpahariya or Maler, Kandh or Kui, Kolami and Parji. Numerically the most important of all the languages/dialects in the Dravidian family is Telugu which has a very rich literature as its vocabulary to a great extent is influenced by Sanskrit. This language is peculiar as unlike other languages and dialects under this tree, it has spread to South Africa, Myanmar and Indo-China.

### *(ii) South Dravidian Languages:*

Tamil, Kannada and Malayalam along with the dialects of Tulu, Kota, Kurgi (or Kedagu) and Toda are part of this branch. Tamil is spoken mainly in Tamil Nadu and even by a large population in Sri Lanka. The significant characteristic of Tamil is that to a great extent it has been able to preserve the old Dravidian character in its original form.

Malayalam is spoken by people residing in Kerala and Lakshadweep. Although, it has its origin in the old Tamil language and dates back about 1,500 years it could carve out its own path during the early 10<sup>th</sup> century and established itself as an independent language by the 15th century. The influence of Sanskrit has been more on Malayalam, than any other language of India.

Kannada, is spoken in Karnataka and has witnessed three stages of development:

- (a) Old Kannada up to 13th century
- (b) Medieval Kannada up to 16th century and
- (c) Hosa Kannada, the language of the present day.

## **2. The Austric Family:**

The Austric family of languages in India basically belong to the Austro-Asiatic sub-family of the world. This family can be divided into two branches:

(i) Munda or Kol Languages

(ii) Mon-Khmer Languages:

### ***(i) Kol or Munda Languages:***

This branch of languages is the prime branch among the Austric family consisting of fourteen tribal languages. The Kherwari spoken in the Eastern parts of the country (Chota Nagpur, Orissa, Chhattisgarh and West Bengal) is the major sub-branch. It further includes Bhumij, Korwa, Santhali, Mundari, Ho, Birhor and Kurku). The languages of Ho, Mundari and Santhali are notable as they have orally preserved their literature through mythological romantic stories as well as songs.

### ***(ii) Mon-Khmer Languages:***

This branch has two further sub-branches— Nicobari and Khasi. The tribal people of the Nicobar Islands speak Nicobari language. While, the Khasi tribal people of Meghalaya speak Khasi language.

## **3. The Sino-Tibetan Family:**

The Sino-Tibetan family can be grouped into three main branches:

(i) The Tibeto-Himalayan

(ii) The North-Assam

(iii) The Assam-Myanmari or Burmese

### ***(i) The Tibeto-Himalayan Languages:***

This branch of the Sino-Tibetan family is further branched into

a) the Himalayan

b) the Bhutia

(a) The Himalayan Group:

Chamba, Lahauli, Kannaure and Lepcha are included in the Himalayan group with Kannaure being the most widely spoken language.

(b) The Bhutia Group:

Lahauli, Sherpa, Tibetan, Balti, Ladakhi and Sikkim Bhutia form the Bhutia group with Ladakhi having the largest number of speakers. It is followed by Sikkim Bhutia and the Tibetan language.

*(ii) North Assam Languages:*

This branch of North Assam languages is also called the Arunachal branch. Aka, Daflta, Abor, Miri, Mishnil and Mishing are included in this branch with Miri having the largest number of speakers.

*(iii) The Assam Myanmar Languages:*

Boro or Bodo, Naga, Cochin, Kukichin and Myanmar languages are included in this branch. The largest language spoken of this group is Naga.

Besides the above mentioned three branches, this family also has some other significant languages like Tripuri, Mikir, Manipuri, Garo and Lusai (Mizo).

**c) Linguistic Regions:**

During the British period, a Linguistic Survey was conducted between 1903 and 1928 and the results showed that there were 179 languages and 544 dialects spoken in India. The survey was conducted under the leadership of Sir George A. Grierson who classified the 179 languages and 544 dialects into different language families along the historical lines. During the 1961 Census for the first time the most comprehensive information was collected on languages. The census figures show that in India 187 languages were spoken. Out of these 187 about 94 languages were spoken by not even 10,000 persons each and only 23 languages (in total) are spoken by approximately 77 per cent of the population of India.

The Eighth Schedule of the Constitution of India has specified 22 languages in addition to English. These are (1) Assamese, (2) Bengali, (3) Gujarati, (4) Hindi, (5) Kannada, (6) Kashmiri, (7) Konkani, (8) Malayalam, (9) Manipuri, (10) Marathi, (11) Nepali, (12) Oriya, (13) Punjabi, (14) Sanskrit, (15) Sindhi, (16) Tamil, (17) Telugu, (18) Urdu (19) Bodo, (20) Santhali, (21) Maithili and (22) Dogri. Nepali, Konkani and Manipuri were added by a Parliamentary Act on 20th August, 1992, while in 2004 Bodo, Dogri, Maithili and Santhali were added.

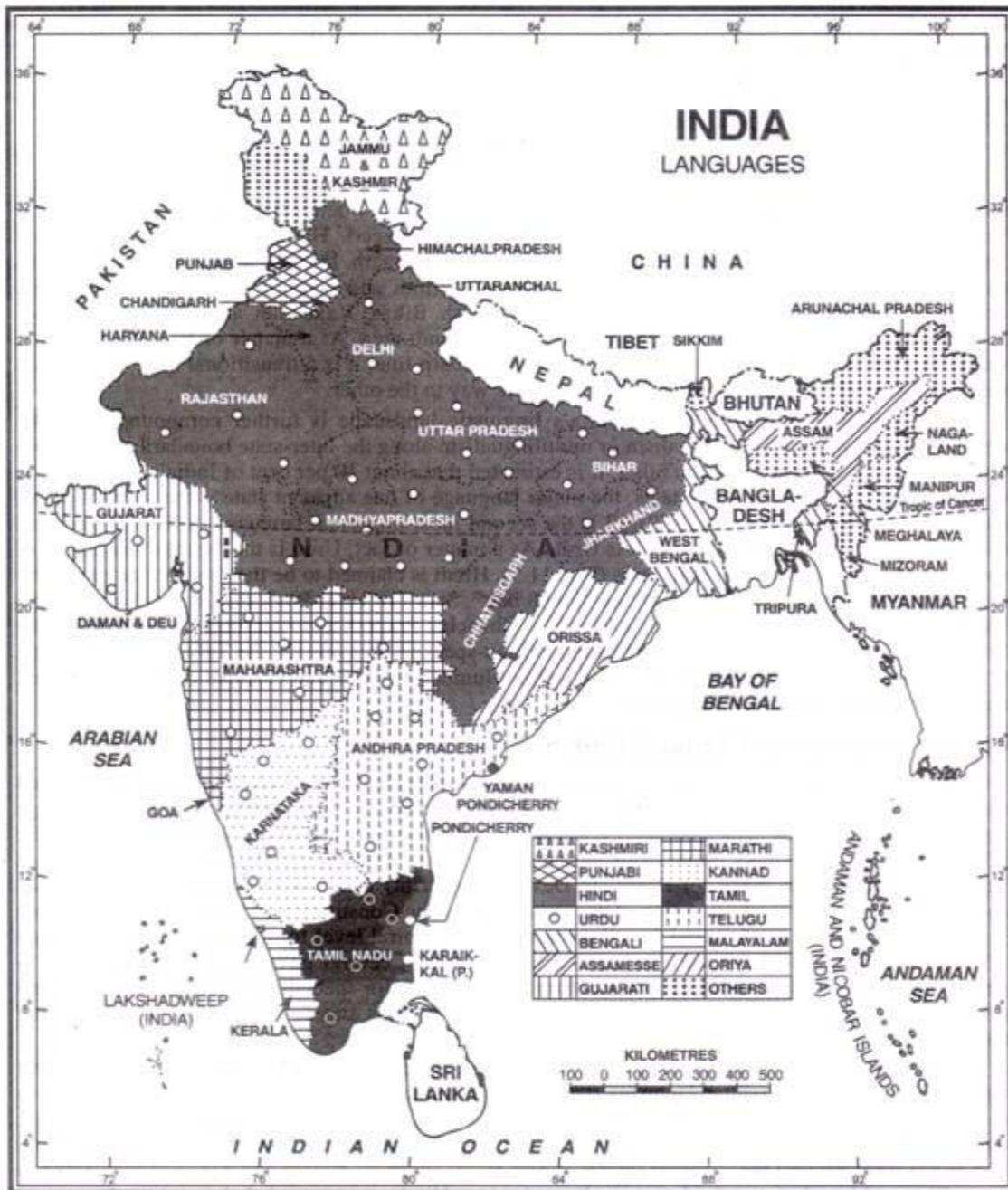
Interestingly, after Independence the formation of states in India has been done on the basis of languages. This added a political dimension to the spatial pattern of languages in the country. The languages and dialects of the tribal population of the north-eastern, eastern and central parts of the country are not part of the spatial pattern as observed in other parts of the country. This is due to two reasons, first the number of tribal languages and dialects varies and secondly the people speaking these languages or dialects live in unspecified areas or regions. In all, there are 12 major languages which are widely spoken in India and based on their spatial distribution twelve linguistic states can be demarcated as depicted in Table 2 and Fig. 3.

**Table 2**  
**Linguistic Regions of India**

<b>Linguistic Region</b>	<b>Political State/ U.T.</b>
1. Kashmiri	Valley of Kashmir
2. Punjabi	Punjab and adjoining parts of Haryana
3. Hindi	U.P., Haryana, H.P., M.P., Chhattisgarh, Jharkhand, Uttaranchal, Delhi, Bihar and Rajasthan
4. Bengali	West Bengal and Parts of Tripura
5. Assamese	Assam and other north-eastern states
6. Oriya	Odisha
7. Gujarati	Gujarat
8. Marathi	Maharashtra, Goa
9. Kannada	Karnataka
10. Telugu	Andhra Pradesh
11. Tamil	Tamil Nadu, Pondicherry
12. Malayalam	Kerala, Lakshadweep

Source -<http://www.yourarticlelibrary.com/language/indian-languages-classification-of-indian-languages/19813/>

**Figure 3**  
**Linguistic Regions of India**



Source -<http://www.yourarticlelibrary.com/language/indian-languages-classification-of-indian-languages/19813/>



A Gateway to All Post Graduate Courses