


Fundamental Concepts in Sociolinguistics
Professor. Dr Om Prakash
School of Humanities and Social Sciences
Gautam Buddha University, Greater Noida
Linguistic Diversity of India

Welcome to the class. Today, we are going to talk about the linguistic diversity of India. India is a multicultural society, and we have had a continued history of civilizational growth since time immemorial, we have a diverse cultural history, diverse, we have diverse societies, and this diversity is also reflected in languages.


We find variation in the language every few kilometers of this vast yet there is a beautiful harmony in this diversity, which startles scholars in the West. And we have survived ages and centuries with this diversity and a continuum. And, we are going to talk about the linguistic diversity of India.

(Refer to Slide Time: 1:40)

Statement - 9
 Family-wise Grouping of the 121 Scheduled and Non-scheduled Languages – Census of India 2011



Language families	Number of Languages	Persons who returned the languages as their mother tongue	Percentage to total population
1	2	3	4
1. Indo-European			
(a) Indo-Aryan	21	94,50,52,555	78.05
(b) Iranian	1	21,677	0.00
(c) Germanic	1	2,59,678	0.02
2. Dravidian	17	23,78,40,116	19.44
3. Austro-Asiatic	14	1,34,93,080	1.11
4. Tibeto-Burmese	66	1,22,57,382	1.01
5. Semito-Hamitic	1	54,947	0.00
Total	121	1,20,87,77,433	77.83



India is known to have 4 distinct language families. These language families are genealogically classified as we saw earlier.

Out of the language families, minority and endangered ones are Islanders/ Anandamanese. In the Indo-European category, we have Indo-Aryan, Iranian and Germanic. They constitute almost more than 78 per cent of the total population, making it the largest language family according to the 2011 Census. Followed by the Dravidian languages family, which accounts for 19.64 per cent of the total population of India. Then, we have the Austroasiatic language family, and it accounts for, this family accounts for 1.11 per cent of the total population of

India. And finally, Sino-Tibetan or Tibeto-Burmese, account for 1.01 per cent of the total population of India, and the rest of the languages constitute a negligible percentage, so it is not mentioned in the Census.

As Census accounts for 99.85 per cent of the recorded population of India the data can be taken seriously. And when the process was complete, the enumerators were left with a huge number of designated mother tongues that goes up to 19,569. And the further classification, linguistic scrutiny and techniques used to re-cluster these numbers, regroup these numbers, these designated mother tongues they arrive at 1,369 mother tongue, designated mother tongues.

India stands and according to Ethnologue, is the second diverse linguistically diverse country in the world after Papua New Guinea. And diversity has not been an impediment in communication, interaction and socialization that is the beauty of this diversity, we have survived for centuries, interacting, communicating and sharing.

It is only the idea of national languages that is alien to us. The whole theory of one nation, and one language is alien. It has never been the case here in India. India is not a homogeneous society. And Indian ecology is so robust and so patterned, that diversity is not a problem in communicating, sharing and socializing.

(Refer to Slide Time: 12:53)

INDO-EUROPEAN

◦ **INDO-ARYAN**

1. Assamese (S),
2. Bengali(S),
3. Bhill/Bhilodi,
4. Bishnupuriya,
5. Dogri(S)
6. Gujarati(S),
7. Halabi,
8. Hindi(S),
9. Kashmiri(S),
10. Khandeshi,
11. Konkani(S),
12. Lahnda,
13. Maithili(S),
14. Marathi(S),
15. Nepali(S),
16. Oaia(S),
17. Punjabi(S),
18. Sanskrit(S),
19. Shina,
20. Sindhi(S),
21. Urdu(S),

◦ **IRANIAN**

1. Afghani/Kabuli/Pashto

◦ **GERMANIC**

- English.

NPTEL

Let us look at the family-wise distribution of languages. So, under Indo-European, we had three groups, Indo-Aryan, Indo-Iranian and Germanic. And in Indo-Aryan, you have 21

languages. And this listing of these languages has nothing to do with hierarchy. It is simply alphabetical in nature. So, it has nothing to do with the hierarchy, alphabetical in nature.

So, we have Assamese then Bangla, Bhili or we call Bhilodi, Bishnupuriya, Dogri, Gujarati, Halabi, Hindi, Kashmiri then Khandeshi, Konkani, Lahnda, Maithili, Marathi, Nepalis, Odia, Punjabi, Sanskrit, Shina, Sindhi, and Urdu. In Iranian, which is a negligible number of Speakers in the country, we also have responders for Afghani, from Afghani migrants, thus Kabuli and Pashto can also be found along with English a Germanic language.

The total languages, like we have 23 languages, including Iranian and Germanic, of which the Indo-Aryan family accounts for 78.85 per cent of the total population of the country.

(Refer to Slide Time: 14:59)

DRAVIDIAN

NPTEL

1. Coorgi/Kodagu,
2. Gondi,
3. Jatapu,
4. Kannada(S),
5. Khond/Kondh,
6. Kisan,
7. Kolami,
8. Konda,
9. Koya,
10. Kui,
11. Kurukh/Oraon,
12. Malayalam(S),
13. Malto,
14. Parji,
15. Tamil(S),
16. Telugu(S),
17. Tulu.

The slide features a list of 17 languages from the Dravidian family. A speaker with a beard and glasses, wearing a blue shirt, is visible in the bottom right corner of the slide frame. The NPTEL logo is in the top right corner.

Then we move to the Dravidian language family, which accounts for almost 20 per cent of the population with 17 languages listed under this family in the 2011 Census. And these languages are Corgi or Kodagu. Gondi, Jatapu, Kannada, Khond, Kisan, Kolami, Konda, Koya, Kui, Kurukh, Malayalam, Malto, Parji, Tamil, Telugu and Tulu.

And there is a total of 70 dialects as designated in the 2011 Census accounts for almost 19 per cent plus of the total population of the country.

(Refer to Slide Time: 16:14)

AUSTRO-ASIATIC

NPTEL

1. Bhumij,
2. Gadaba,
3. Ho,
4. Juang,
5. Kharia,
6. Khasi,
7. Koda/Kora,
8. Korku,
9. Korwa,
10. Munda,
11. Mundari,
12. Nicobarese,
13. Santali(S),
14. Savara.

The slide features a list of 14 languages from the Austroasiatic family. A speaker with a beard and glasses, wearing a blue shirt, is visible in the bottom right corner of the slide frame. The NPTEL logo is in the top right corner.

Moving on to the third language family, which is the Bhumij so Austroasiatic family. And languages like Bhumij, Gadaba, Ho, Juang, Kharia, Khasi, Koda, Korku, Korwa, Munda,

Mundari, Nicobarese, Santalis and Savara. Under this, we have 14 languages which account for 1.11 per cent of the total population of the country.

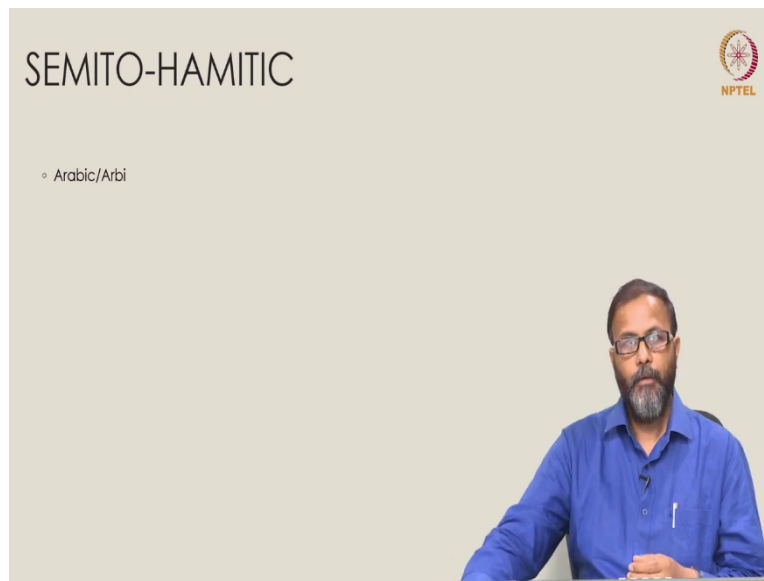
(Refer to Slide Time: 16:51)

TIBETO-BURMESE

1. Adi,	19. Khezha,	37. Maram,	55. Tamang
2. Anal,	20. Khiemlungan,	38. Maring,	56. Tangkhul,
3. Angami,	21. Kinnauri,	39. Niri/Mishing,	57. Tangsa,
4. Ao,	22. Koch,	40. Mishmi,	58. Thado,
5. Balti,	23. Kom,	41. Mogh,	59. Tibetan,
6. Bhotia,	24. Konyak,	42. Monpa,	60. Tripuri,
7. Bodo (S),	25. Kuki,	43. Nisi/Dalla,	61. Vaiphei,
8. Chakesang,	26. Ladakhi,	44. Nocte,	62. Wancho,
9. Chakri/Chokri,	27. Lahauli,	45. Palte,	63. Yimchungre,
10. Chang,	28. Lakher,	46. Pawl,	64. Zeliang,
11. Deori,	29. Lalung,	47. Phom,	65. Zemi,
12. Dimas,	30. Lepcha,	48. Pochury,	66. Zou.
13. Gangte,	31. Uiangmei,	49. Rabha,	
14. Garo,	32. Limbu,	50. Rai,	
15. Halam,	33. Locha,	51. Rengma,	
16. Hmar,	34. Lushai/Mizo,	52. Sangtam,	
17. Kabui,	35. Manipuri(S),	53. Sema,	
18. Karbi/Mikir,	36. Mao,	54. Sherpa,	

Then we come to the most diverse family. Though the total number accounts for only 1.01 per cent of the total population of India, it is the most diverse family, which has 66 distinct languages recorded under this family in the Census 2011. And the languages listed in this family in Census 2011 are Adi, Anal, Angami, Ao, Balti, Bhotia, Bodo, Mishing, Mishmi, Mogh, Monpa, Ladakhi, Lahauli, Lakher, Lepcha, Limbu, Koch, Kom, Konyak, Garo, Hmar, Kabui, Karbi, Deori, Rai, Rabha Ringma, Sangtom, Sema, Sherpa, Tamang, Tangkhul, Thoda, Thado, Tibetan, Tripuri, like, Zeliang, Zemi, Zou, Tamang, Tangkhul, Maring, Maram, Kinnauri, Lepcha, Lakher, Lahauli, Ladakhi, Lushai or Mizo, Manipuri, Mao, and several others account for 1.01 per cent of the total population of India.

(Refer to Slide Time: 19:10)



And the Census also records the SEMITO-HAMITIC family, which has one language Arabic. So, if you go by the distribution of speakers that we talked about earlier also, Indo-European families have 94,50,52,555 speakers. Say 94,50,52,555 speakers account for 78.05 per cent of the total population, distributed among 21 different languages. Hence, the majority of the largest languages spoken in the country are grouped under the Indo-Aryan family.

Then, with Iranian speakers like Pashto, and Kabuli speakers, we have 21,677 speakers. Then Germanic speakers like English, we have 2,59,678 as first language speakers and the percentage is negligible approximately 0.02 per cent of the total population of the country. Then, the second largest family Dravidian language family has 17 languages recorded under it, and it accounts for 23,78,40,116 speakers and accounts for 19.64 per cent of the total population of the country.

Then Austro-Asiatic family with 14 languages has 1, 34,93,080 speakers and accounts for 1.11 per cent of the total population. Tibeto-Burmese, the most diverse family with 66 languages has 1,22, 57,382 speakers accounts for 1.01 per cent of the total population. Semito-Hamitic family with one language Arabic has 54,947 speakers percentage is negligible.

Now, if you look at the distribution and concentration of these families on the map of India, though the speakers are scattered, the Census survey depicts the distribution of these languages 121 languages pan India, across states and union territories.

For example, is not the case that, Dravidian speakers are only in the southern part of the country, they are also in the north, the southeast the east in the northeast. Since numbers are distributed pan India broadly, we can only recognize the concentration areas of these languages where the speakers are largely located.

The majority of speakers are located on the map of India. However, it is not absolute, and the speakers of all these languages and all these language families are mixed and scattered across India. The Indo-European language concentration can be seen from Kashmiri at the top to Marathi, Bangla and Gujarati. If you look at the southern part of the country, you have Dravidian speakers with 17 languages concentrated around states like Kerala, Karnataka, Tamil Nadu, Andhra, and Telangana.

If you look at the Tibeto-Burman language concentration with the most diverse language family with 66 varieties recorded in the Census 2011 you have the northeastern part of the country. We call them the land of 7 sisters and 1 brother. They are the states like Sikkim, Assam, Tripura, Manipur, Nagaland, Mizo, Mizoram, and Arunachal.

If you look at Austro-Asiatic languages, they are scattered all over, but the majority of them are in the central part of India, and they cover states like Maharashtra, Chhattisgarh, Jharkhand, part of Orissa, part of Bengal. So, there is the spread of this family. So, if you look at the Indian map, you have a country with multiple languages and multiple societies, multiple cultures. And their diversity finds meaning in the etymological in the demographic sense of the word.

And we have a beautiful balance of linguistic ecological balance, and linguistic harmony. There are some undercurrents and conflicts as well, but these demands are local, and they have been handled effectively.

The Census also manifests the same idea that these speakers are scattered all over. They are freely mixing, sharing and living together in perfect harmony. And we need to celebrate this diversity, which is unique, and the fact that India is the largest diverse linguistically diverse country after Papua New Guinea.

Of the 450 languages as per the Ethnologue records, one-fourth of the languages are spoken alone in the Indian subcontinent. Hence language formed the basis for state bifurcation post-independence. And for more details, language-wise details, dialect wise details, you can

refer to the Census data 2011. In our next video, we will talk about the scheduled and non-scheduled languages of India. Thank you.