

**Indira Gandhi National Open University
School of Translation Studies and Training**

Translation and Indian Languages



Expert Committee

Prof. Manager Pandey
Former Chairman, Center for Indian Languages, Jawaharlal Nehru University

Prof. Ganga Prasad Vimal
Former Chairman, Center for Indian Languages, Jawaharlal Nehru University

Prof. C.P. Shivdasan
Regional Institute of English Bangalore

Prof. Harish Narang
Former Chairman, Center for English Studies, Jawaharlal Nehru University

Prof. Chaman Lal
Former Chairman Indian Language Center Jawaharlal Nehru University

Prof. Harimohan Sharma
Department of Hindi, Delhi University, Delhi

Prof. Ramabaksh
Chairman, Center for Indian Languages, Jawaharlal Nehru University

Dr. Puran Chand Tandon
Department of Hindi, Delhi University Delhi

Prof. J.M. Parakh
School of Humanities, IGNOU, New Delhi

Prof. Ritarani Paliwal
School of Humanities, IGNOU, New Delhi

Shri Kumar Vikram
National Book Trust India

Faculty (SOTST, IGNOU, New Delhi)

Dr. Rajendra Prasad Pandey

Dr. Deoshankar Navin

Dr. Jagdish Sharma

Dr. Manjit Baruah

Dr. Harish Kumar Sethi

Dr. Jyoti Chawla

Programme Coordinators (Hindi Version)

Prof. Rajendra Prasad Pandey
Prof. Deo Shankar navin

Course Coordinator (Hindi Version)

Dr. Jyoti Chawla
Assistant Professor, SOTST, IGNOU

Course Editor (Hindi Version)

Prof. Harimohan Sharma
Head, Dept. of Hindi, Delhi University, Delhi

Course Writers

Dr. Vimlesh Kanti Varsha, Associate Professor (Retd.), Department of Hindi, P.G. D.A.V. College, University of Delhi	Unit 01
Dr. Manju Mukul, Associate Professor, Department of Hindi, Delhi Delhi	Unit 02
Dr. Bhagwanti Jadvani, Assistant Professor, School of Social Sciences, IGNOU	Unit 03, 04
Dr. Rajkumar Prasad, Assistant Professor Shyam Lal College University of Delhi, Delhi	Unit 05
Dr. Nirja, Assistant Professor, Department of Elementary Education, Mata Sundri College, University of Delhi, Delhi	Unit 06
Dr. Manjit Baruah, School of Translation Studies and Training, IGNOU (Translation Shweta Yadav)	Unit 07
Dr. Harish Kumar Sethi, School of Translation Studies and Training, IGNOU	Unit 08
Mr. Kumar Vikram, Editor, National Book Trust, India (Translated by Vipin Jha)	Unit 09, 10
Prof. Radhavallabh Tripathi, Former Vice-Chancellor, National Sanskrit Institute, Delhi	Unit 11, 12
Prof. Suvas Kumar, Professor (Retd.) Department of Hindi, Hyderabad Central University Hyderabad	Unit 13
Dr. Jyoti Chawla, School of Translation Studies and Training, IGNOU	Unit 14

Course Preparation (English Version)

Programme Coordinator

Prof. Rajendra Prasad Pandey
Professor & Director
SOTST, IGNOU, New Delhi

Course Coordinator

Dr. Jyoti Chawla
Associate Professor
SOTST, IGNOU, New Delhi

Vetting & Editing of English Translation

Dr. Jyoti Chawla
Associate Professor
SOTST, IGNOU, New Delhi

Faculty Members

Prof. Rajendra Prasad Pandey
Dr. Harish Kumar Sethi
Associate Professor
Dr. Jyoti Chawla
Associate Professor
Dr. Chitresh Soni
Assistant Professor

2014 (Hindi Version), February, 2023 (English Version)

@Indira Gandhi National Open University, 2014 ISBN: 978-81-266-6683-6

All rights reserved. No part of this work may be reproduced or transmitted in any form without permission in writing from the copyright holder.

CONTENTS

BLOCK 1 INDIAN LANGUAGE FAMILY

- UNIT 1 INDO-ARYAN LANGUAGE FAMILY
- UNIT 2 DRAVID LANGUAGE FAMILY
- UNIT 3 AUSTRO-ASIATIC LANGUAGE FAMILY
- UNIT 4 CHINESE-TIBETAN AND OTHER LANGUAGE FAMILIES

BLOCK 2 MODERN INDIAN LANGUAGES AND MUTUAL TRANSLATION

- UNIT 5 INTRODUCTION TO MODERN INDIAN LANGUAGES WITH REFERENCE TO TRANSLATION
- UNIT 6 COMMON AND DISTINCT FEATURES OF MODERN INDIAN LANGUAGES
- UNIT 7 MODERN INDIAN LANGUAGES AND METHODS OF TRANSLATION

BLOCK 3 DIMENSIONS OF TRANSLATION AND INDIAN MULTILINGUALISM

- UNIT 8 INDIAN MULTILINGUALISM AND TRANSLATION
- UNIT 9 NATION BUILDING AND DIMENSIONS OF TRANSLATION-1
- UNIT 10 NATION- BUILDING AND DIMENSIONS OF TRANSLATION-2

BLOCK 4 TRANSLATION OF INDIAN CLASSICAL LITERATURE INTO MODERN INDIAN LANGUAGES

- UNIT 11 INDIAN CLASSICAL LITERATURE: AN INTRODUCTION
- UNIT 12 TRANSLATION OF ANCIENT INDIAN CLASSICS INTO MODERN INDIAN LANGUAGES
- UNIT 13 TRANSLATION OF MEDIEVAL INDIAN CLASSICS INTO MODERN INDIAN LANGUAGES
- UNIT 14 TRANSLATION OF MODERN INDIAN CLASSICS INTO MODERN INDIAN LANGUAGES

Course Introduction

This course titled *Translation and Indian Languages* (MTT014) is designed for the translation of M.A. (Translation Studies) is the fifth course of M.A (Translation Studies). As the title of the course suggests, various points related to translation and Indian languages have been considered within this course. Under this course, considering all the major Indian language families in detail, first of all the subject has been entered. After that, explaining the concept of modern Indian languages, the role of translation in their development has been studied in detail. At the time of conceptualizing the present course, attention was paid to this. It has been done so that all aspects of translation in Indian languages and Indian linguistic scenario can be discussed in detail. Material is available here on all the complete points including Indian multilingualism, role of translation in nation- building, Indian classics composed in different times.

There are total four blocks in this course, whose names are as follows.

- Indian language family
- Modern Indian Languages and Mutual Translation
- Dimensions of Translation and Indian Multilingualism
- Translation of Indian Literature into Modern Indian Languages

These four blocks contain 14 units in total.

The first block is **Indian Language Family**. There are four units under this section which basically focus upon four major Indian language families.. These language families are the Indo-Aryan language family, the Dravidian language family, the Austro-Asiatic language family and the Chinese-Tibetan language family. These four language families have been discussed in detail in this section.

The second block is **Modern Indian Languages and Mutual Translation**. There are three units in the present block, in which the introduction of modern Indian languages in the context of translation has been given in the first unit. What were the important factors in the rise of modern languages in India, the time and circumstances of the rise of different languages and the important literature written in them, etc. have been discussed in detail.

Third, **Dimensions of translation and Indian multilingualism**: There are three units under this section, in which the first unit is Indian multilingualism and translation, under which Indian multilingualism and its various dimensions have been discussed in detail. The second and third units are titled Nation-Building and Dimensions of Translation- 1 and 2. In the first part nation- building has been discussed in the European context as a prelude and in the second part an in- depth study of this unit has been done in the Indian context.

The last block of this course is **Translation of Indian Classical Literature into Modern Indian Languages**. In this block, a general introduction to Indian classical literature has been given, under which major Indian classical compositions have been discussed. Rest three units focus on ancient, medieval and modern literature and their translation into modern Indian languages respectively. Overall, this entire course covers Indian languages, their introduction, classical literature written in them and their interaction with each other.

BLOCK 1

INDIAN LANGUAGE FAMILY

UNIT 1

INDO-ARYAN LANGUAGE FAMILIES

UNIT 2

DRAVID LANGUAGE FAMILY

UNIT 3

AUSTRO-ASIATIC LANGUAGE FAMILY

UNIT 4

CHINESE-TIBETAN AND OTHER LANGUAGE FAMILIES

ignou
THE PEOPLE'S
UNIVERSITY

Introduction to Block-1

The first block of **Translation and Indian Languages** (M.T.T.014) is Indian Language Families. Indian language family refers to those language families whose languages are spoken in India. Mainly their number is four, under which many languages come. Translation acts as an important bridge in a multilingual country like India.

There are total four units in this block:

Unit 1. The title is **Indo-Aryan Language Family**. The Indo-Aryan language family is not a separate language family but a part of the larger Indo-European language family. In this unit, while explaining and classifying this, attention has been focused on the Indo-Aryan language family.

Unit 2 is titled **Dravidian Language Family**. Languages of this family are spoken in the southern part of India. Unit discusses the origin of its nomenclature, place of Dravidian languages and characteristics of these languages in detail.

Unit 3 is titled **Austro-Asiatic Language Family**. The Austro-Asiatic language family is mainly divided into two parts - the Austronesian language family and the Austroasiatic language family. Languages spoken in India and around India come under the Austroasiatic language family. Therefore, it has been discussed in detail in the present unit.

Unit 4 is titled **Chinese-Tibetan and Other Language Families**. The Chinese-Tibetan language family, also known as Sino-Tibetan or Tibetan Chinese, is one of the largest language families in the world. It has the largest number of speakers. There are two main groups of this language family - Chinese and Tibeto-Burmese. Chinese refers to various Chinese languages and Tibeto-Burman languages are mainly spoken in Tibet, Myanmar and North-East India. In these languages, only the languages of the Tibetan group are spoken in India. In addition, other language families have also been briefly discussed in this unit.

UNIT 1 INDO-ARYAN LANGUAGE FAMILY

Unit Structure

- 1.0 Objectives
- 1.1 Introduction
- 1.2 Family classification of 12 Indian languages
 - 1.2.1 Erani Sub-Branch
 - 1.2.2 Darad Sub-Branch
 - 1.2.3 Indians Arya Sub-Branch
- 1.3 Historical development of 13 Indian Aryan languages
- 1.4 Explanation of 14 Indian languages and Hindi Interrelationships
- 1.5. Let Us Sum Up
- 1.6 Exercise
- 1.7 Suggested Readings

1.0 Objectives

After reading this unit you will:

- Learn about different language families:
- You will be able to know which language families are spoken in India,
- understand the historical development sequence of the Indo-Aryan languages:
- Will be able to discuss the similarities and differences between Indian Aryan languages, and will be able to explain the inter-relationship between Indian languages and Hindi.

1.1 Introduction

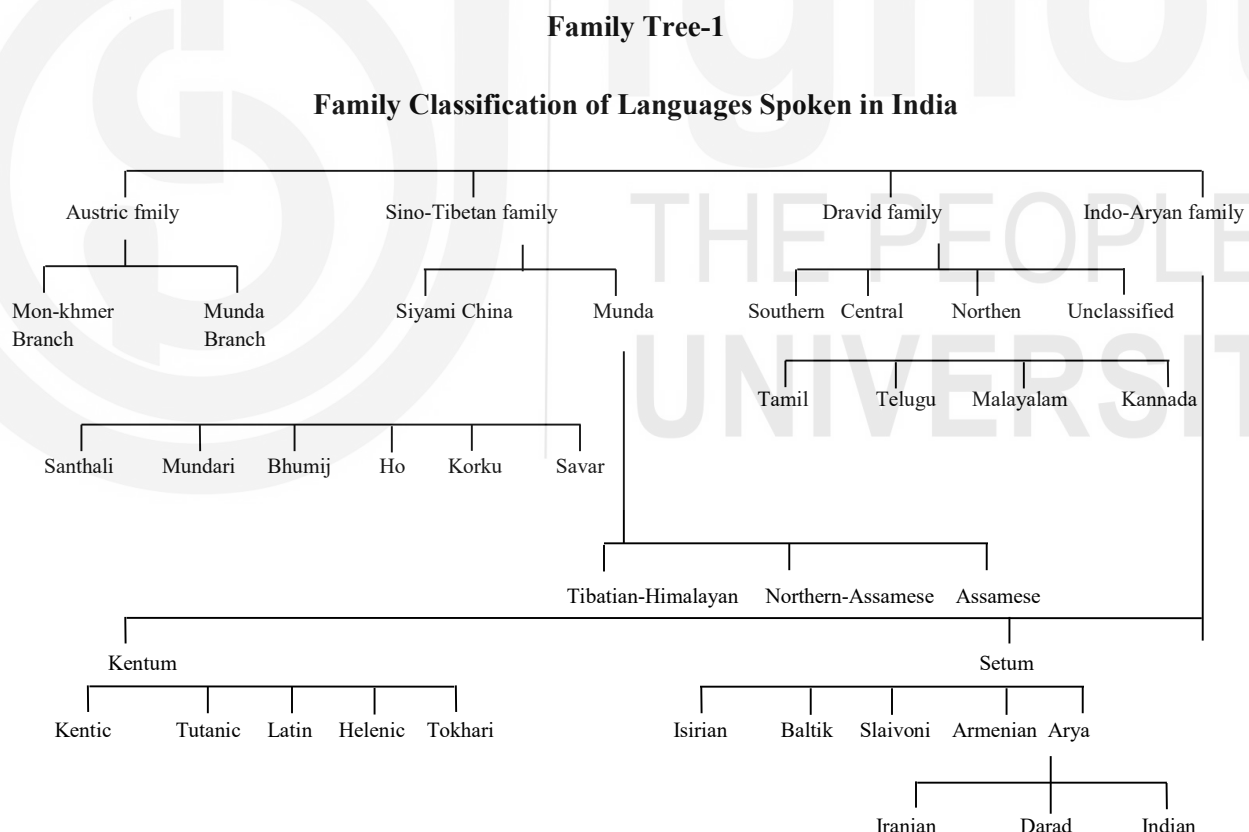
Language is related to individual, society and country. Therefore each language has its own characteristics. On the basis of these, it establishes similarity with some languages and on the other hand with many languages. Inequality. On the basis of this, they are divided into language families. The basis of this classification is completely scientific. In fact, all the languages of the world have been divided into

language families on the basis of their properties and similarity. There are eleven language families in the world, Indo-European, Dravidian, Austric, Sino-Tibetan, Semitic, Hermetic, Ural Altaic, Malayo-Polynesian, Antu, and Central African American. Out of these four language families, the languages spoken in India are Dravidian, Austric, Indo-European and Sino-Tibetan. Indo-European family of these is the largest. Most of the languages spoken in India belong to this language family. We will discuss them in detail in the next section.

1.2 Family Classification of Indian Languages

India is a vast linguistic area and there is a commonality among the diversity of linguistic and social units of the country which keeps everyone connected. 1652 languages (Indian census 1961) of four language families are spoken in India.

There are four language families (a) Austric language family (b) Sino-Tibetan language family (c) Dravidian language family and (d) Indo-European language family. The position of Indo-Aryan languages in the language families of India can be easily understood through the following diagram. See Vash tree 1.



The Indo-European family is the most important language family in the world. Its expansion is from India to Europe. The languages of this family are divided into two classes Ketum and Satam. Of these, the Indo-Iranian or Aryan branch is under the seventh class. Sub-branches of this Indo-Iranian branch are the

Iranian, Dard and Indo-Aryan languages. Of these, the languages of the Dard and Indo-Aryan sub-branches are spoken in India.

1.2.1 Iranian sub-branch

The main language of the Iranian sub-branch 'Iranian' language is very ancient and its oldest available literature is available in the form of 'Avesta'. Inscriptions from the 6th century BCE of the Harmani emperors are also examples of the ancient form of the Iranian language. There are two forms of the Iranian language - Eastern and Western. Eastern Iranian is the language of Pamiri dialects and Avesta, while Western Iranian includes Median and Old Persian.

The western part of ancient Iran was called 'Persia' and the language there was almost contemporary with both forms of Persian Iranian, Old Persian and Avesta. Some scholars are in favor of considering 'Avesta' as more ancient. The earliest form of Old Persian can be found in the cuneiform inscriptions of the Achaemenian kings. The developed form of ancient Persian is called 'Pahlavi' and the literature of Pahlavi begins to be found from the third century. There are two forms of Pahlavi- 1. Hujvareh 2. Parsi or Pajand. The influence of the Semitic family can be seen on Huzvareh. Its script is also Semitic while it has not affected Parsi or Pajand. This is the language of the Parsis who settled in India and Pajand has considerable influence on Gujarati. Avesta and ancient Persian Sanskrit is near Sanskrit while medieval Persian is near to Apbhansh. Firdausi's 'Shahnama' is an important work in modern Persian literature. There is no abundance of Arabic words in its language, but the abundance of Arabic words started increasing in later Persian. Here again, the use of Iranian words is increasing in Persian by replacing Arabic words due to the feeling of nationalism. Vargista, Pashto (Padto), Dewari, Balochi, Kurdi etc. language forms are of Iranian only and all of them are rich because of their folk literature.

1.2.2 Dard sub-branch

The main language of this sub-branch is Kashmiri. Kashmiri is spoken over a wide area of about 10,000 square miles of Jammu and Kashmir. It is the official language of Jammu and Kashmir under Article 8 of the Indian Constitution. It is mainly spoken in Anantnag, Srinagar, Daramulla and Doda. According to the Indian census 2001, it is spoken by 55,27,698 (0.54 percent) people as a mother tongue, and in terms of numbers, it ranks fifteenth among Indian languages. Kashmiri native speakers call it by the name 'Kashur' but it has been named Kashmiri along with other Indian languages in the Indian Constitution. Although the name Kashmiri was used by Amir Khusro in the thirteenth century, but the language name continued to be used for it in the Kashmir province till the seventeenth century. Thus the name Kashmiri came into vogue much later, but literary works were being done in this language much earlier. In the thirteenth century, Shitikantha 'Mahanay Prakash' and Lalded enriched Kashmiri with their works. Nundarishi, Habba Khatoon, Arnimal, Habibullah, Naushhari, Sahib Kaul, Prakash Ram, Ramzan Waka Bariullah etc. are the main medieval poets of Kashmiri. Kashmiri had its own Sharada script till 600 years ago, but from 14th century when Persian became the official language of Kashmir, the Persian script started being used for Kashmiri.

1.2.3 Indo-Aryan sub-branch

The languages of the Indo-Aryan sub-branch are spoken in most parts of India (except South India). Assamese, Odia, Bengali, Marathi, Punjabi, Gujarati, Sindhi, Hindi and Urdu are the languages of this sub-branch in the scheduled Languages by the Constitution of India. The following are the main Indo-Aryan languages spoken in India:

Odia: Odia is the official language of Orissa state. It is spoken and understood in 60,136 square miles of Orissa region, as well as in Medinipur of Bengal: Bihar, Singhbhum, Sarai Kela, Kharsua, Pugnjar, Raigad, Bastar of Madhya Pradesh and some parts of Andhra Pradesh. It has been developed from Ardhamagadi Apabhramsa and its script is circular Devanagari. According to the 2001 Indian census, there are 3,301,7446 (3.21 percent) Odia mother tongue speakers in India, which is the tenth largest number in India. Odia's literature, both ancient and modern, is rich. Among the modern poets, Gop Bandhu Das, Neelkanth Das, Padmacharan Pattanayak, Laxmikant Mahapatra and Kuntala Kumari Sawant etc. are well known names. Kalindicharan Panigrahi, Harihar Mohapatra, Harishchandra Mukerjee and Anand Shankar Rai are also notable writers of Odia.

Bangla: Magadhi Apabhramsa Prasut of the Aryan language family, Modern Aryan language Bangla is the main spoken language in West Bengal. It has been accepted as the official language of West Bengal under Article 8 of the Indian Constitution. In terms of population, Bangla language speakers have the second place among the modern Indian languages. Bengali is the mother tongue of approximately 8,33,69,769 (8.11 percent) Indians. There are five main dialects of Bangla:

1. North-Eastern Bangla, 2. Northern Bangla, 3. Eastern Bangla, 4. Western Bangla, 5. South-Western Bangla. Of these, the first three are spoken in Bangladesh and the remaining two in the state of West Bengal in the Republic of India. The earliest literary form of Bengali language is found in the Charyapadas of the Sahajiya sect of Buddhism. Long after the Charyapadas, another poetic form of Bengali language came to us based on Krishna's rasa gathas. It is found in the poetry of Jayadeva. Chaitanya Mahaprabhu Krishna Kirtan presents form of fifteenth century Bangla language. The Ramayana written by Krittivas, Geet Govind written by Jaidev and the works of Chandidas are unique treasures of Bengali literature. Today all literature genres are found in Bangla language. The names of Rabindranath Tagore, Sharat Chandra, Vimal Mitra, Tarashankar Padhyay Shankar etc. are particularly notable among the writers of the modern era.

Assamese: Assamese, spoken in a vast area of 85,000 square miles in the eastern part of India, dating back to the Middle Ages, is the official language of the state of Assam. Asamia is spoken mainly in the districts of the Brahmaputra region. According to the Indian Census 2001, 131,565,151 (1.25 percent) people in India called it their mother tongue. The dialects of Assamiya are Jharwa, Nayang etc. and Assamese is also spoken in Nepal and Arunachal Pradesh apart from Assam. Its script is a form of Devanagari script which has evolved from Bangla. We find the oldest literary form of Assamese language in couplets of Charyapad. Assamese is a language rich in literature. Navkal Harihar Kakati, Virendra Bhattacharya, Mahendra Bara Heemen Var and Virendra Bargohai are noteworthy among budding poets.

Sindhi: Sindhi, a Brachad Apbhransh of the Udichi branch of the Indian languages, is accepted as a mother tongue by 25, 35, 450 (0.25 percent) people in India. Its speakers are also in Madhya Pradesh, Maharashtra, Punjab. It is basically the language of Sindh Pradesh.

Punjabi: Punjabi is the main spoken language of the state of Punjab. The number of its speakers is 29102477 (2.83 percent) and in terms of population, its place is eleventh among the popular languages in India. The main script of Punjabi is Gurmukhi, which is a form of Devanagari. Guru Nanak, Guru Arjun, Bhai Gurudas are devotee poets of Punjabi. There are 29 dialects of Punjabi in which Bhatiyali, Jagri Gurmukhi, Kangri and Rathli are the main ones.

Gujarati : Derived from Shaurseni Apbhransh Gujarati is spoken by 46,091,617 (4.48 per cent) people in the 710,072-square-mile state of Gujarat, from Kutch and Marwara in the north to Mumbai's Thane district in the south and the Arabian Sea in the west and Malwa and Khandesh in the east. In terms of population in India, the place of Gujarati is seventh. It is the official language of Gujarat. Its script is headline free Devanagari script.

Marathi: Derived from Maharashtrian Apabhransh Marathi is the main spoken language of Maharashtra and according to the Indian Constitution, it has also been accepted as the official language of the state of Maharashtra. The number of its speakers is 7,19,36,894 (6.9 percent), which comes in fourth place in terms of population among all Indian languages. It has about 65 dialects. Its major dialects are Halbi, Konkani, Kamari Katia, Kaskari, Kothli, Marathi, Kshatriya Marathi, Parkhari and Vat. The script of Marathi language is Devanagari. Marathi is very rich from the literary point of view. Gyandev, Namdev, Dnyaneshwar, Eknath, Mukteshwar, Tukaram, Ramdas are famous Marathi saint-poets.

Hindi: Hindi is one of the main languages of the Modern Aryan language family. It has been accepted as the official language of the Indian Union in the Indian Constitution. Hindi is the language of a wide area of India. It is spoken mainly in Uttar Pradesh, Uttarakhand, Madhya Pradesh, Uttisgarh, Bihar, Jharkhand, Rajasthan, Himachal Pradesh, Haryana, Delhi and Andaman and Nicobar. According to the Indian census 2001, 41.03 percent of the population of India treats Hindi as mother tongue. Hindi is mainly written in Devanagari script.

Urdu and Sanskrit languages are also mentioned in the eighth schedule of Indian languages. According to the Indian census, the number of people who use Urdu as their mother tongue is 5,15,56,111 (5.91 percent). The number of people who use Sanskrit language as their mother tongue is 14,135. In fact, the use of Sanskrit as a colloquial language is negligible. Urdu occupies a prominent place among the scheduled languages. From political point of view, it is also the national language of Pakistan, but from linguistic point of view, Urdu is a literary style dominated by Arabic-Persian words of Hindi. When the Mughals came to India, the language used between the Indian merchants and the Mughal military buyers in the military camps developed, it was a vertical dialect from the point of view of language structure and its vocabulary also had Persian-Arabic words. Since when the use of the word Urdu started in the sense of language, this topic still remains controversial. Most scholars agree that the word Urdu was used for the language only in the late eighteenth century.

1.3 HISTORICAL DEVELOPMENT OF INDIAN ARYAN LANGUAGES

The origin of the modern Indian Aryan languages is Sanskrit, the ancient language of the Aryans. From the point of view of historical development, the branch of Indian Aryan language can be divided into three periods.

1. Ancient Aryan Language Period (2000 BC 500 BC)
2. Medieval Aryan Language Period (500 BC 1000 AD)
3. Modern Aryan language period (1000 BC till date).

Ancient Aryan language period and Medieval Aryan language period can be divided into three sub-periods and the modern Arya language period has been considered in two sub-periods, the duration of each period is 1500 years and the duration of sub-period is about 500 years. The second period of modern Aryan language will be till 2000 AD.

1. Ancient Aryan Language Period (2000 BC 500 BC)
 - (a) Vedic Language Period (2000 BC 1500 BC)
 - (b) Brahmin period (1500 BC to 1000 BC)
 - (c) Literary Sanskrit Period (1000 BC 500 BC)
2. Medieval Aryan Language Period (500 BC 1000 AD)
 - (a) Pali Language Period (500 BC - 001 AD)
 - (b) Prakrit Language Period (001 AD - 500 AD)
 - (c) Apbhramsa Language Period (500 AD 1000 AD)
3. Modern Aryan Language Period (1000 AD - 2010 AD)
 - (a) Forms influenced by ancient Apabhramsa (1000 AD - 1500 AD)
 - (b) Middle Ages (1500 AD to present)

Literary forms of many regional languages developed in the Middle Ages of the modern Aryan language period. It was in this era that Awadhi, Braj, and Maithili gained literary recognition.

Any language can have two forms:

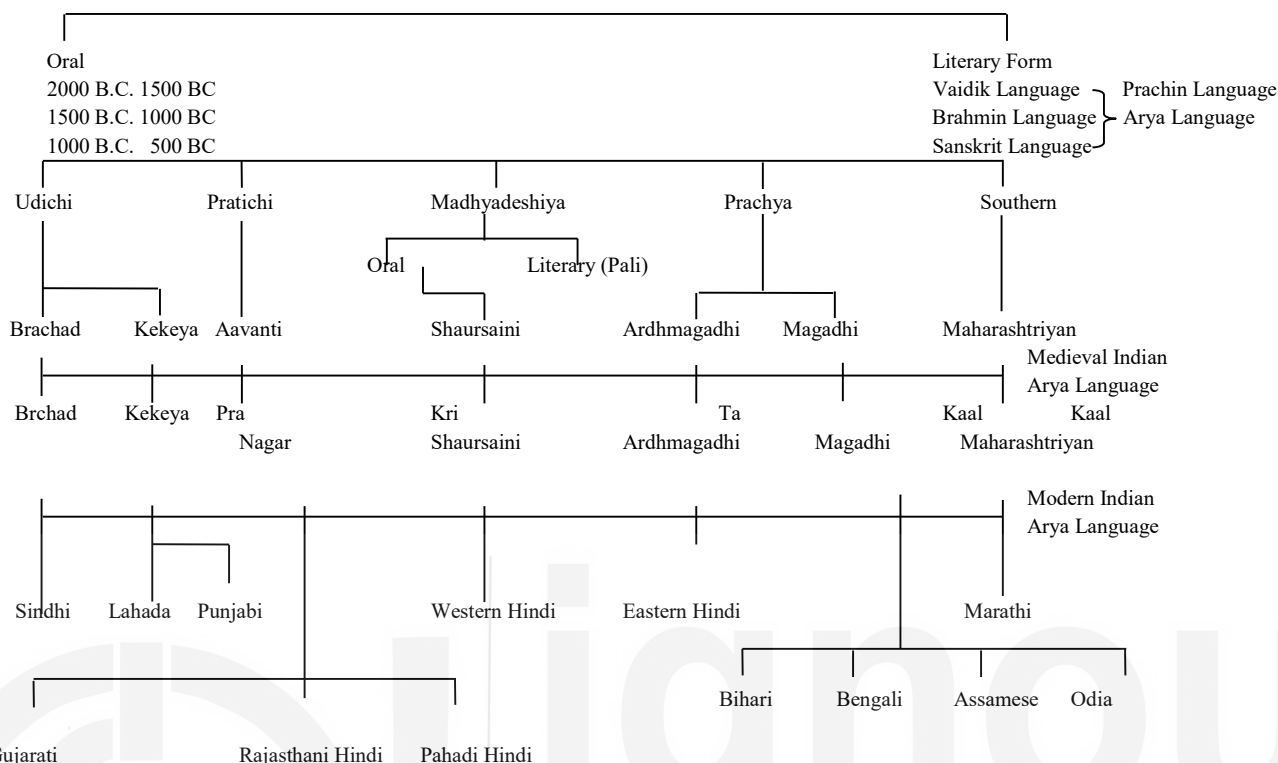
(a) Oral or colloquial form

(b) Literary form.

We still do not know for sure what the oral form of the language was in the ancient Aryan period. We get to see the literary form of that period in the language of Rigveda. There is always a difference between the literary and the oral form, but the literary form is also understood by the public. The life span of each literary language is approximately 500 years. Gradually, the language of the public keeps on developing, while the literary language remains stable and a time comes when the public does not understand the literary language. This is the time when a new literary language is born taking the basis of oral form.

During 500 the language of the people kept drifting away from Sanskrit, of which there is sufficient proof. From this time, medieval Aryabhasha begins. At this time five forms of oral language are found in North India. Most of the literary discussions are found in the oral form of the Middle East. The language in which this discussion takes place is famous in the name 'Pali'. Literary composition is not found in other languages. In its 500 years the language developed further and was named Prakrit. Day-to-day Prakrit emerged from different oral language forms. Udichi (Northern has two forms (1) Brachad and (2) Kaikeya Pratichi (Western) remained only one form Avanti Prakrit. Madhyadeshi also had only one form - Shauraseni Prakrit. Prachya has two forms (1) Ardhamagadhi, (2) Magadhi Prakrit, a form of Dakshini -Maharashtrian Prakrit.

This situation was from 001 BC to 500 AD. After this it developed again in 500 years. The form of Brachad Prakrit evolved and was called Brachad Apabhramsa. Thus from Kekeya Prakrit Kaikeya Apabhramsa, from Avanti to Nagar Apabhramsa, from Shauraseni Prakrit to Shauraseni Apabhramsa, from Ardhamagadhi Prakrit to Ardhamagadhi Apabhramsa, Magadhi from Magadhi and from Maharashtrian Prakrit Maharashtrian Apabhramsa developed. This situation dates back to 1000 AD during the Middle Aryan period. At this time only Apabhramsa language developed and modern Aryan languages developed from it. Among the modern Aryan languages, Sindhi developed from Brachad Apbhramsa, Lahnda and Punjabi from Kaikeya Apbhramsa, Gujarati, Rajasthani and Pahadi from Nagar Apbhramsa, Western Hindi from Shauraseni Apbhramsa, Eastern Hindi from Ardhamagadhi, Bihari, Bangla and Odia from Magadhi Apbhramsa and Marathi from Maharashtrian Apbhramsa. Its dialects have been developed from Nagar, Shauraseni, Ardha Magadhi and Magadhi Apabhramsa. The development of Indo-Aryan languages can be understood from the following diagram. Refer Family Tree 2.



1.4 INDIAN LANGUAGES AND HINDI: INTER-RELATIONS

Due to India's complex linguistic and social structure, scholars have sometimes called India as a socio-linguistic giant and sometimes a 'linguistic area' because of unity existing in the core of linguistic social units.

According to Indian Census, four language families (the languages of the Indian family) are spoken in India. Out of these, 1455 languages are those whose number of speakers is less than ten thousand. It was placed in the Eighth Schedule considering it as the principal language. These languages were - 1. Assamese, 2. Odia, 3. Urdu, 4. Kannada, 5. Kashmiri, 6. Gujarati, 7. Tamil, 8. Telugu, 9. Punjabi, 10. Bengali, 11. Marathi, 12. Malayalam, 13. Sanskrit, 14. Hindi. The number of speakers was 98 percent of the total population of the country. Later on, Sindhi was added in the eighth schedule and later Nepali, Konkani and Manipuri were added. Now in the eighth schedule 18 languages of the country are counted in India. And according to the 1991 census, the number of speakers of these languages is 90.92 percent of the country's population. In conclusion, today the main languages of the country are twenty-two. These languages belong to the three language families of the world, Indo-European families including Assamese, Odia, Urdu, Kashmiri, Gujarati, Punjabi, Bengali, Marathi, Sindhi, Konkani, Nepali, Sanskrit and Hindi, Dravidian families including Tamil, Telugu, Malayalam and Kannada and Manipuri belong to the Sino-Tibetan Chinese family.

Hindi became the language of central importance in these main languages. It got the pride of national language, official language and contact language. The reason for this was that Hindi is the language of the

largest area of the country in terms of area expansion. It is the main language of eleven states of the country - Uttar Pradesh, Madhya Pradesh, Rajasthan, Bihar, Himachal, Haryana, Chhattisgarh, Uttaranchal, Jharkhand, Delhi and Andaman and Nicobar Islands, its speakers are 42.88 percent of the country's population. It is the link language of the country from the point of view of business, mass communication and politics. That is why Hindi is considered to be the representative language among the languages of the country, probably this is the reason that where other Indian languages are known by their respective regional names Punjab - Punjabi language, Marathi - Maharashtra's Bangla - Bengal's Assamese- Assam's, Tamil- Tamil Nadu's where Hindi, Hindi or Hindustani is called the language of Hindustan. Seeing this prevalence of Hindi, Mahatma Gandhi gave the noun of Hindi as the national language and all the leaders of the country who mainly belonged to the Hindi region supported Hindi. Wide publicity was planned and it also became the language of freedom movement of the country.

Most of the languages of India, regardless of which language family they belong to, have amazing similarities. The 18 major languages of India, which are enumerated in the Eighth Schedule of the Indian Constitution, are the languages of three language families, the Indo-European family, the Dravidian family and the Sino-Tibetan language family. The main Indian language of the fourth language family, the Munda family, is Santhali. The study of Indian languages confirms that the Indian languages of the Indo-European family bear more resemblance to the languages of other families in India than to other languages of their own family. For example, Hindi, the modern Indo-Aryan language of the Indo-European family, has many similarities with Tamil, a language of the Dravidian family, while English, German, French, the European languages of the Indo-European family do not have those similarities. Thus Hindi language of different family is closer to Tamil while away from English, German etc. of its own family. Similarly, the similarity of Hindi and Tamil is seen in the languages of the Munda family. This similarity of Indian languages of different language families apart from the languages of their family reinforces the principle of unity in diversity.

If we look at the inter-relationships of Indian languages and Hindi, we see amazing similarities which are not coincidental but a sign of the cultural-social relations of those languages and reinforces the external diversity and internal unity. This similarity in sound, word, sentence and the script is visible at all levels.

There is a similarity in the phonological arrangement of Indian languages irrespective of the language family they belong to. The alphabetic order of the modern Indo-Aryan languages of the Aryan language family and the languages of the Dravidian family is the same. Tamil, the ancient and main language of the Dravidian family, which has a limited number of consonants, the alphabet is also very different from Devanagari, but the order of the consonants in the alphabet are similar to those of Hindi. Except for the last four consonants, which are typical Tamil sounds, all the rest of the consonants have the same alphabet as the Devanagari alphabet.

In the Hindi vowel alphabet, there is a system of order of Hasva and Long. This order system is visible in all Indian languages. In the European languages of the Indo-European family, this sequence is not observed in the vowel alphabet. The sequence of consonants in Indian languages is more scientific even in consonants pronounced from the throat to the lips, the sequence is visible. This scientific consonant is less arranged. Due to this there was very little difference in Indian languages and equality remained.

The pronunciation of 'T' is characteristic of Indian languages. In any other European language of the Indo-European family, the sounds of the T class are not pronounced. The sounds of class T are also present in Indo-European, Dravidian and Munda family languages. Some scholars are of the opinion that the sounds of T class came in Sanskrit and in modern Aryan languages under the influence of Dravidian languages in the languages of India. Aspirated sounds are characteristic of Indian languages. Aspirated consonants in virtually all earlier Indo-European languages was a pronunciation that those languages later abandoned. This pronunciation was preserved in Sanskrit from which it came to other Indian languages. The languages of the Dravidian family adopted this feature from the languages of the Indo-Aryan family. Nasal sound of each group, their pronunciation, their phonology is the specialty of Indian languages whereas European languages have only two Nasal consonants. Five Nasal consonants are found in Indian languages.

At the word level, there is similarity between Hindi and Indian languages. The word group of Hindi is mainly the word group of Indo-Aryan language. Many words in Hindi came from Sanskrit invariable form like fire, moisture, air, water, earth, mother, father, religion, karma, birth, death etc. Some words came to Hindi from Sanskrit through Pali, Prakrit and Apabhramsa and also reached other Indian languages. Those Sanskrit words which brought phonetic changes in Hindi are called Tadbhava words like Agni to Aag, Purva to Purab, Dant to Daant, mayur to more, Chaturdash to Chaudah, Krishna to Kanha, Ratri to Raat. There are calculations whose sources we have no knowledge of today. Such words were originally considered regional words originated in the country (deshaj) or formed.

Since Hindi is the language of a wide area and Hindi speakers are spread almost all over the country, it is natural that Hindi has borrowed words from various Indian languages as well. Somewhere there was a phonetic change in these words and somewhere they were accepted in the same form. Indian languages bordering Hindi are Punjabi, Sindhi, Gujarati, Marathi, Telugu, Oriya, Bangla, Nepali etc. Hindi took novels, fiction from Bangla for literature, Lagu, Chalu, Pragati from Marathi, few words from Gujarati, and Chole Bhature from Punjabi.

One more feature of this vocabulary imported in Indian languages is that Indian languages have formed their form along with Sanskrit words like Tatsam, Tadbhava etc. and words coming from Persian, Arabic etc. We also create words in Indian language by taking foreign suffixes and prefixes. Due to this similar word formation method, the family intelligibility of Indian languages has also increased.

The ever-increasing stock of terminological vocabulary in Indian languages has also kept the Indian languages fresh. All the Indian languages feel the need of technical terminology for the enrichment of the language in the field of science, education, law, administration, management etc. We generally take the support of Sanskrit for the construction of technical vocabulary because Sanskrit has a rich and vast vocabulary, as well as the words created with the help of Sanskrit are many in Indian languages. For example, telephone, Doordarshan, Doorlekh. There is no dearth of technical words taken from English - Academy, Union, Strike, Rocket, Missile, Capsule, Laser, Xray etc. The use of common terminology at the international level and the expansion of the communication range bring the languages closer. With the expansion of the knowledge area of language speakers, the practice of terminological vocabulary increases and this expansion of terminological practice is not only making Indian languages capable, but also playing an important role in bringing them closer. Today, there is a need that an all-India terminological dictionary should be created and its standardization should also be done.

Sentence structure is a complex aspect of language analysis. To illustrate the basic unity or interrelationships of various Indian languages, here are some examples of syntactic similarities of Indian languages. Different languages have different gradations. Both English and Hindi are languages of the Indo-European family. Word order in English language sentence is subject + verb+ object but in Hindi language word order is of subject +object +verb. The hierarchy of English language is also found in German, French, Spanish etc. languages, but the unique thing is that the hierarchy of Hindi is different from the European languages of its family but similar to Indian languages. Hierarchy of Hindi syntax is also found in the languages of the Indian Arya branch, in the Dravidian languages and in the languages of the Munda family.

Along with gradation, another feature related to gradation is the order of words within a phrase. In Hindi we use causative marks after nouns, hence these causative marks are also called prepositions. Preposition means a part to be added after such as in the pocket, to the market, for the home.), Tajindagi (throughout life), Sivaya Aap (except you). Probably the preposition was a common feature of the Aryan languages, but it seems that the preposition developed only in the Indo-Aryan languages. Preposition is used in Indo-Aryan language, Dravidian languages and Indian languages of Tibeto-Burmi group (Manipuri etc.) and in this way all Indian languages can be called preposition languages.

If attention is paid to the arrangement of words in verb composition and if English and Indian languages are compared on this axis of contrast, then we see that in Hindi language the main verb comes first and the auxiliary verb comes later and in English the main verb. The word comes at the end of the verb phrase and the auxiliary verb.

This form of verb formation is common in all Indo-Aryan languages and Dravidian languages. In these, the main verb comes first, the auxiliary verb later, there is uniformity in all Indian languages even from the point of view of the modal verb. 'Do' is a root verb. From this verb, taking tax, giving tax, putting tax, passing tax, keeping tax, etc. are colorful verbs. In the European languages of the Indo-European family, there is no regulation of the dative verb. Sanskrit also does not have modal verbs. This feature is seen only in modern Indian languages. Many scholars are also of the view that modal verbs were the specialty of Dravidian languages and from there the composition of modal verbs that reached other Indian languages is similar in modern Indian languages of both Dravidian and Aryan family. In English, causative verbs are different; they are not formed by changing the form of the basic verb.

Humankind invented script to free oral expression from the limits of country and time. Except Kashmiri, Sindhi and Urdu based on Persian script, Indian languages have no scripts. All these scripts are derived from Brahmi script, so there is similarity of letters in them, the writing trend is also the same. The script of Hindi, Marathi and Sanskrit is Devanagari. The languages have different scripts for Tamil, Telugu, Malayalam, Kannada, Gujarati, Punjabi (Gurmukhi), Bengali (which is used for Assamese and Manipuri). There has been a change in their characters in the process of development, yet a clear similarity is visible.

The similarity at the linguistic level in all the major languages of India explains the interrelationship between Hindi and Indian languages. The reason for this fundamental unity of Indian languages is the cultural unity of the country. The unity of religion, philosophy, literature, art and culture strengthens

linguistic unity. All the languages have their own grammatical features, their own scripts, their own rich literary traditions but they also have mutual similarities which enhance linguistic intelligibility and all are linked to each other so that we can consider India as a vast linguistic region.

More important than proving the interrelationship of Indian languages and Hindi is that how inter-relationship between Indian languages and Hindi can be made more intense and dominant.

In Article 351 of the Indian Constitution, special instructions have been given for the promotion of Hindi. This instruction in the form of a mantra not only indicates the inter-relationship between Indian languages and Hindi, but also explains it. That the development of Hindi and the development of Indian languages are interdependent.

The inter-relationship of Indian languages and Hindi at the literary and linguistic level has been briefly discussed in the unit. It is worth mentioning that in the 22 languages enumerated in the Eighth Schedule of the Constitution, Hindi is also enumerated along with other Indian languages, which has already been mentioned in Article 343 (1) as the official language. Similarly, Sanskrit is also counted, which is not a modern language, nor is it used in oral practice. This is probably because both Hindi and Sanskrit languages are the feeders of all Indian languages and cooperation of both of them is essential for the development of Indian languages.

The direction given in Article 351 of the Constitution of India for the promotion of Hindi indicates that Hindi will be successful and can represent India true and complete only when it becomes the medium of expression of India's social culture and this is possible only when it enriches its vocabulary mainly from Sanskrit and secondarily from other Indian languages by assimilating the used form, style and positions of Indian languages.

1.5 Let Us Sum Up

India is a continent. It is a country of many cultures and a vast linguistic area. All the languages of India, along with their linguistic and literary characteristics, also have their linguistic similarities and their rituals, their vision of life binds them together. The development and expansion of Indian languages lies in the intensity and strength of these interrelationships. In the present unit, while studying about the Indian languages, we learned about the Indo-Aryan language family in detail. Moving ahead in this sequence, in Unit 2, we will discuss in detail about the Dravidian language family.

1.6 Exercise

1. To which language family do most of the languages spoken in India belong? Name that language family and all the languages.
2. What is the difference between oral and literary form of language? How does this affect the development of language?

3. At what levels are there similarities between Hindi and other Indian languages? Explain with example. What is the arrangement in the Indian Constitution to strengthen the interrelationship between Hindi and other Indian languages? Write an article on Indo Aryan language family.

1.7 Suggested Readings

- Ferguson, CA. 1966, National Social Linguistic Profile Formulas, Linguistic Vol. Bright
- Emenyu MB, 1966, India A. Linguistic Area, Language in Society (Hymes).
- Census of India, 1967. New Delhi, Volume 1 Part (01) C (11)
- Census of India, 1991, Series 1-India Part IV V (I.I) (A) Table C, India, States and Union Territories.
- Chaturjya Suniti Kumar, 1965, Indio-Aryan Language and Hindi, Delhi, Rajkamal Prakashan.

UNIT 2 DRAVIDIAN LANGUAGE FAMILY

Unit Structure

- 2.0 Objectives
- 2.1 Introduction
- 2.2 Nomenclature
- 2.3 The Question of the Place of Origin of the Dravidian Linguistic Community
- 2.4 Vocabulary: Historical Evidence
- 2.5 Sanskrit and Dravidian Languages
- 2.6 Language Areas (Area Extension)
- 2.7 Dravidian Languages as a Family
 - 2.71 Linguistic Features
- 2.8 Let Us Sum Up
- 2.9 Exercise
- 2.10 Glossary
- 2.11 Suggested Readings

2.0 Objectives

After reading this unit you will be able to know:

- Origin and nomenclature of the Dravidian language family;
- Native languages of the Dravidian language family as well as other languages;
- Place of Dravidian languages among Indian languages;
- Area of Dravidian languages, and
- Linguistic features of Dravidian languages.

2.1 Introduction

The Dravidian language family is important in India as the second language family after the Indo-European language family. It is the fifth largest language family in the world. Languages belonging to the Dravidian language family are spoken by about 200 million people in different parts of the world, mainly in South Asia. This language family includes about 26 languages, as well as about twenty languages for

which no literature is available. In the present unit, we will discuss the important points related to the Dravidian language family.

2.2 Nomenclature

The term was first used by Robert Caldwell around 1856 as 'Dravidian'. The word is a modified form of Sanskrit 'Dravid' which was traditionally used for the Tamil language and people. Accepting this, Robert Caldwell says that the word 'Dravidian' is made due to the influence of the word *Dravida*. This word was used for Tamil language and people. They acknowledge that there can be differences of opinion about the term and nomenclature. This word was used by Sanskrit philologists to refer to South Indian people and their languages. Caldwell gives the source of this word that the use of this word as a language is found in Kumaril Bhatta's 'Tantravartika'. The use of this word is also found in Manusmriti, Acharya Bharata's Natyashastra and Mahabharata. These include the use of the word 'Dravid' in reference to the people and the use of the word 'Dravidi' in the sense of a small Prakrit language developed from Peshachi.

2.3 The Question of the Place of Origin of the Dravidian Linguistic Community

Historians and linguists have considered the question of the place of origin of the Dravidian linguistic community from different points of view. In this context, Cavalli-Sforza says-

The center of origin of the Dravidian languages may be somewhere in western India. This center can also be in Southern India. This language family is found scattered in northern India and as Brahui in western Pakistan. (Cavalli-Sforza, LL. 2000, Genes, People and Languages)

Cavalli-Sforza also links the Dravidian languages with the Indus Valley Civilization. In fact, there is no archaeological or linguistic evidence of the fact that when the people speaking Dravidian languages came to India. But there is evidence to suggest that they were in India around the 15th century BCE when the Rigvedic Aryans entered northwestern India.

Without any evidence, most scholars accept that about two centuries before the arrival of the Aryans the former Dravidians came to India from the North Western region. Rasmus Rusk first proposed this idea that their relationship can be linked with the primitive tribes of North Asia and Europe.

Dr. Suniti Kumar Chaturjya believes that *the Dravidians were representatives of different branches of the Mediterranean races.* (Indo Aryan Languages and Hindi)

Some scholars also believe that the original Dravidian-speaking people were residents of the West, their original residence is found in some areas of the Eastern Mediterranean Sea, the Lycia region of Asia Minor and the Crete Islands. The ancient name of the Dravidians: *Dramij* or *Dramil*, from which the words *Dramid*, *Dravid*, *Mil* and *Tamil* came out of Arya language. Analysis of Dravidian literature suggests that it dates back to the middle of the first millennium AD. But the period when the Aryans came in contact with the Dravidians of India and outside India has been estimated the middle of the second millennium.

Concluding all this, it can be said that all these hypotheses are based on guesswork. No authentic and effective hypothesis has come to the fore yet. All the speculations that have been made regarding the entry of the East Dravidians into India are based on the fact that the *Brahui* was the result of the first division of East Dravidians and the Indus Valley Civilization are believed to be associated with the

Dravidians. It is also believed that the civilization of Harappa and Mohenjodaro was originally inhabited by the Dravidians. The Brahui caste of Baluchistan was also a Dravidian.

Dr. Suniti Kumar Chaturjya estimates that 'before the advent of the Aryans, it was the Dravidians who built the great civil civilizations of Punjab and Sindh. Overall, this fact can be considered that the Dravidians were the natives of India. At the time when the Aryans entered India, the Dravidians lived in the western and northwestern regions of India. 1500 BC After the Aryans entered India around 1500 BC, they spread to different regions of India.

2.4 Vocabulary: Historical Evidence

Many words of Dravidian origin are found in Rigvedic Sanskrit, the oldest form of Sanskrit, on the basis of which the belief developed that before the entry of the Aryans into India around 1500 BC, the Dravidians had settled in North-Western India. The credit of Retroflex Consonants in Sanskrit in Rigvedic period also goes to the contact of Sanskrit speaking people with Dravid speaking people. According to Russian scholar Nikita Gurov, there are about eighty words of dravid origin in the Rigveda.

Caldwell and many other scholars have discovered many such words in Greek, Latin and Hindu about which they believed that these are basically originated from Dravidian. But the authenticity of such beliefs is doubtful.

2.5 Sanskrit and Dravidian Languages

Many linguists targeted the influence of Dravidian languages on Sanskrit from many points of view like phonym, grammar, words etc. Many scholars accept that the development of 'T' class phonym in Sanskrit was due to Dravidian influence. Compound verbs, prepositions and comparative adjectives are also the contribution of Dravidians. Imitative vocabulary in Sanskrit can also be considered as its contribution.

According to Dr. Hardev Bahri, the scheme of conjunctions, future tense, two words, two genders, use of preposition instead of inflection, additional verbs in passive voice and elements of sentence planning have come from Dravidian language.

2.6 Language Area (Area-Extension)

On the basis of the number of language users, the Dravidian language family is considered to be at the fifth position amongst all the language families of the world. The Dravidian language family is spread over a large area of India in terms of area. Detailed information related to the area expansion of languages that come under the Dravidian language-family is as follows:

Tamil - Sri Lanka, South Africa, Malaysia, Singapore, Mauritius, Fiji, Burma, Tamil Nadu

Malayalam - official language of Kerala

Kannada - official language of Karnataka

Telugu - official language of Andhra Pradesh

Tulu - Southern Kannada district of Karnataka, west coast of Kerala

Kodagu - Kodagu (Coorg) district of Karnataka, bordering Kerala

Irul - Nilgiri Hills

Kurumb – Nilgiri Hills

Tod - western region of Nilgiri Hills

Kot - Nilgiri Shilpi

2.7 Dravidian languages as a Family

Around 1816, Francis Whyte Ellis did research on South Indian languages under the name 'Dissertation on the Telugu Language'. Ellis was an English Civil Service officer. He found that while and class and mass Tamil, grammatical and general Telugu, ancient and modern Carnatic and Kannada, Malayalam and Tulu together form a language family which can rightly be called the dialects of South India.

Earlier in 1786, Sir William Jones gave the concept of language family. About thirty years later, Alice tried to establish the Dravidian languages as a family. In support of his hypothesis, he placed Telugu, Kannada and Tamil under this language-family from the lexical and grammatical point of view. In 1856, Robert Caldwell wrote a work called *Comparative Grammar*, which can be called the first seminal work related to the comparative study of the languages of the Dravidian language family. He accepted twelve languages under the Dravidian language-family - Tamil, Malayalam, Telugu, Kannada, Tulu, Kodagu or Coorg, Tod, Kota, Gondi, Kui, Oraon, Malto.

Caldwell was able to link these languages under the hypothesis of a language family. In the second half of the 19th century, several languages were added to the Dravidian language family - Gondi (Driberg 1849), Kui (Letchmajee 1853), Kolami (Hislop 1866), Kodagu (Cole 1867), Tulu (Brigel 1872), Malto (Droese 1884), Tod (Bernhard Schmidt), Brahui (Leech 1838).

Today about 26 Dravidian languages are known, which are divided into four groups:

South Dravidian - Tamil, Malayalam, Irula, Kurumva, Kodagu, Tod, Kot, Budag, Kannada, Korag, Tulu
South-Central Dravidian - Telugu, Godi, Kond, Kui, Kubi, Pengo, Manda
Central Dravidian - Kolami, Nekri, Neki, Paarji, Ollari, Gadab
North Dravidian - Kuruks, Malto, Brahui

2.7.1 Linguistic Features

- The languages of this family are mainly simple agglutative. Languages have the ability to create the compound. The quality of tone-conformity is present.
- These languages have five short and five long vowels. Only a few languages of the Nilgiris and Kodagu have phonemic centralized vowels, which evolved from a modified form of phoneme initial vowels before double consonants. Circled vowels developed only in Tod and Irul. There is no vowel stress in any Dravidian language. Generally stress is on the first vowel of the word, whether short or long words can start with both vowel and consonant. There are both short and long vowels. Very short 'a' is added after the last consonant. There is the priority of retroflex sounds. It is believed that 't' vowel sounds in Sanskrit were developed from Dravidian languages only.

- The Dravidian languages have a tendency to form structurally compound words, i.e. the languages are structurally compound in their own structure. The words do not have prefixes and suffixes. Grammatical relations in words are established on the basis of suffixing and compounding.
- Dravidian nouns do not have inflectional forms. While using words in a sentence, that is, while forming clauses, such locutions are added to the words which can also be separated. The only difference between singular and plural is that the original form of the spatial word is added in singular, and in the plural indicative form, after adding the plural suffix. All of them, regardless of whether they are of a word or a list, take those spatial forms. Which takes by the singular. No separate word form of these migrants can be considered. They are simply relative nouns and are used as auxiliaries.
- Words are often spoken. There is no law for the use of consonant sounds at the beginning of a word, but it is necessary to pronounce vowel sounds, especially in Tamil.
- The subject is used at the beginning of the sentence and the verb at the end, the adjective comes before the adjective, the adjective goes before the verb, the object of the verb and all its adjectives are used before the verb. Adjective participle is used before a proper noun. The negative of the sentence comes immediately before the participle. The relative factor or word is used before related. In place of the preposition, the postposition is added and then the causative expression is expressed. The ending of the sentence is done by the verb. These are the distinctive features of the Dravidian languages.
- On the basis of these features, Caldwell believes that these Dravidian languages are clearly so different from the Aryan languages that they have to be considered as a completely different family from the Indo-European family, especially from Sanskrit. Although they have been in social, religious contact with Sanskrit speaking languages for ages, there can be no doubt about their different existence.
- There are two words and three genders. Male or female gender words are added for gender. The functions of inflections are taken from suffixes. The forms of adjectives do not work according to the noun, there is an excess of dental forms in verb. There is no passive.
- There are two types of nouns - conscious and unconscious. Conscious means the one who has the power of knowledge-discretion and unconscious means the one who does not have the power of knowledge-discretion.
- In Sanskrit and other Indo-European languages, adjectives also follow the same form as nouns and pronouns. Their Gender, words, factors are all there. The forms of adjectives do not work in dravid languages.
- Adjectives are not noun based that is, the forms of adjectives do not follow the noun. When they are not used as adjectives with substantive but as possessive nouns, they also have the same form as nouns, but when they are used before substantive in adjective form, only the original form before it are kept in.

- There is an excess of participle forms in the verb. Participle forms also have signs of verbs and also of adjectives. In the Indo-European languages, pre-locatives are used to convey the same meaning that the Dravidian languages use pre-locatives. Those on-spaces cannot be considered as a separate word form. These are just simple nouns indicating quality or relationship and are used as meaning-helpers. An adverb is either a noun or a past participle and is used before a verb that they are adverbs.
- There is no use of verb or expression. To express their sense, auxiliary verbs are added which mean to suffer or be forced. There is a greater tendency to use prepositions instead of conjunctive elements.
- There is a tendency of participle forms in place of relative pronouns.

2.8 Let Us Sum Up

In the present unit, while discussing Dravidian languages in detail, we have discussed some important points. We learned that almost all the languages of the southern part of India are included under the Dravidian language family. Apart from the major languages of this family, Tamil, Telugu, Kannada, Malayalam, it also has many important languages which are spoken in South India. Discussing in this sequence, we also learned here that the Dravidian language family is the fifth largest language family in the world, under which about 26 languages are spoken. Apart from this, the origin of Dravidian languages is before Sanskrit. Although scholars do not have any concrete evidence regarding this, but on the basis of estimation, it is said that Dravid had come to India many years before the arrival of Aryans.

Discussing the place of origin of Dravidian languages, we also studied the views of various linguists. It can be clearly said that the Dravidian language family is an important language family not only of India but of the world. In the next unit, we will study another important language family of India, the Austro-Asiatic language family.

2.9 Exercise

1. The languages of the Dravidian language family are mainly spoken in which part of the world?
2. Give the origin and background of the word 'Dravid'.
3. What has been accepted by the scholars as the original place of the Dravidian linguistic community?
4. Where were the native Dravidian speakers? What is the opinion of scholars and historians?
5. What is the opinion of the historians regarding the connection between the Indus Valley Civilization and the Dravidians?
6. What linguistic features unite Sanskrit and Dravidian languages?
7. Tell about the area expansion of Dravidian language-family.
8. Who was the first scholar to identify the Dravidian languages as a family and with whom did the workers
9. Name the number of languages included in the language family by Robert Caldwell.

10. Which languages joined the Dravidian language family in the second half of the 19th century?
11. Into how many categories can the languages of the family be divided basically? Indicate the names of languages fall in each category.
12. State the main linguistic features of the language.

2.10 Glossary

- **Language family** - Languages that have a common source, geographic background as well as similarity in sound, form, meaning and syntactic structure, form a language family.
- **Prakrit language** - ancient popular language, linguistic community, such a group of people who have the same rules, conduct and behavior regarding the use of language
- **Simple agglutative language** - it means not glued. that is, the language in which the semantic elements and the related elements should be combined in such a way that their form can be clearly identified.
- **Self (short) / long vowel** - sounds take some time to pronounce. In this way, depending on the amount of time (Kaal), there are two types of vowels, long and short. It takes less time to pronounce vowels like a. I. And it takes longer to pronounce long vowels like aa, e, oo
- **Phonemic central vowels** - a group of vowels that are pronounced between the front and the back of the tongue. These vowels also have the ability to change meaning.
- **Circled vowel** - The nature of the vowels is determined on the basis of the position of the circled vowels. The position of eight is circular, they are called Vrittamukhi.
- **Word stress** - We do not give equal emphasis to all the sections that we speak. The vowels and words on which we emphasize while speaking, there the pronunciation becomes intense, which we describe.
- **Vowel** - A vowel is a sound that is pronounced without any obstruction in the mouth.
- **Consonants** - The sounds whose utterance passes through some or the other obstruction in the larynx.
- **Retroflex sounds** - In the pronunciation of these sounds, tongue turns around and touches the cerebrum. This action is also called retroflexion.
- **Aspirated sound** - The airflow may be less or more in the pronunciation of aspirated voiced consonants. When this flow is low, it is called Alp Prana and when it is high, it is called Maha Prana. The passage of the larynx becomes narrow in the pronunciation of low-pitched sounds, due to which it vibrates.
- **Vowel words** - words that end with a vowel.

- Vivrit - Those vowels in whose pronunciation the mouth cavity is very open, that is, there is more and more space between the tongue and the vocal range, like *aa*.
- Participle - Participle means - action + end. Suffixes added at the end of root are called Krita suffixes and thus form participles.
- Suffix - that unit of sound or group of sounds which is used to change the grammatical form or meaning.
- For is added to the root word but cannot be used independently.
- It is readable.
- Active voice - It is that form of action by which the knowledge of the law of supremacy of doer, action or emotion in action.
- Passive Verb - The verb refers to the action of the past tense.
- Expressive - expressive when the emphasis is on emotion.
- Conjunctive elements - The elements that join or combine two words, phrases or phrases are called conjunctive elements in language.
- The words used before the preterite verb are called preterite.
- Combined verbs - Verbs in which there are combined root words.
- Vowel stress - where there is more stress on vowels than on consonants in words.

2.11 Suggested Readings

- Tewari, Bhola Nath, 1973, Hindi Language, Allahabad, Kitab Mahal. Chaturjya, Suniti Kumar, Indian Aryan Language and Hindi.
- Caldwell, Robert, A Comparative Grammar of the Dravidian or South Indian Family of Languages, Delhi, Asian Educational Services
- Steever, Sanford B. (ed.), 1998. The Dravidian Languages, London and Newyork, Routledge
- Ramaiah, L.S., 1994 General and Comparative Dravidian Languages and Linguistics, Vol-1. Madras, T.R. Publications

UNIT 3 AUSTRO-ASIATIC LANGUAGE FAMILY

Unit Structure

3.0 Objectives

3.1 Introduction

3.2 Classification of Austro-Asiatic family

3.3 Munda Languages

3.3.1 Structure of Munda Languages

3.4 Mon-Khmer Languages

3.4.1 Linguistic Features

3.5 Khasi

3.5.1 Linguistic Features

3.6 Other Related Issues

3.7 Let Us Sum Up

3.8 Glossary

3.9 Exercise

3.10 Suggested Readings and References

3.0 OBJECTIVES

In the unit, our objectives will be to discuss:

- Classification of Austro-Asiatic family
- Munda and Mon-Khmer languages
- Languages included in both the groups (Munda and Mon-Khmer)
- Geographical spread and linguistic characteristics like the sound system, lexical and morphological system, and the syntactico-semantic system

- Structural changes in the languages due to contact with other language families

3.1 INTRODUCTION

In this unit we will discuss the third prominent language family of India and Southeast Asia: Austro Asiatic Language family. By now you must have acquainted yourself with two language families, namely, Indo Aryan and Dravidian language families. It is believed that the presence of Austro-Asiatic languages predates Dravidian and Indo Aryan languages. The term Austro-Asiatic indicates the subgroup under Austric Family and it includes Mon-Khmer and Munda and other groups of languages.

This language family has its presence mainly in Southeast Asia. Austro Asiatic family is subdivided into three branches: Mon Khmer, Munda and Nicobarese (Abbi 2001). But there is no uniformity in the classification. Asher (2008) includes Nicobarese under Mon Khmer; he writes “Two of the twelve or so recognized divisions of Mon-Khmer have their home in India, namely Khasi and Nicobarese (2008: 41).” Asher gives two principal groups- Mon-Khmer and Munda and the subdivisions of Mon-Khmer, he writes, are believed to be more complex than the Munda languages. The precise membership of this family is still a matter of debate. The biggest handicap in this field is and always has been the lack of really adequate data (Thomas 1964).

On the whole this family includes 25 languages and 12 million speakers in India (Singh & Manoharan 1993).

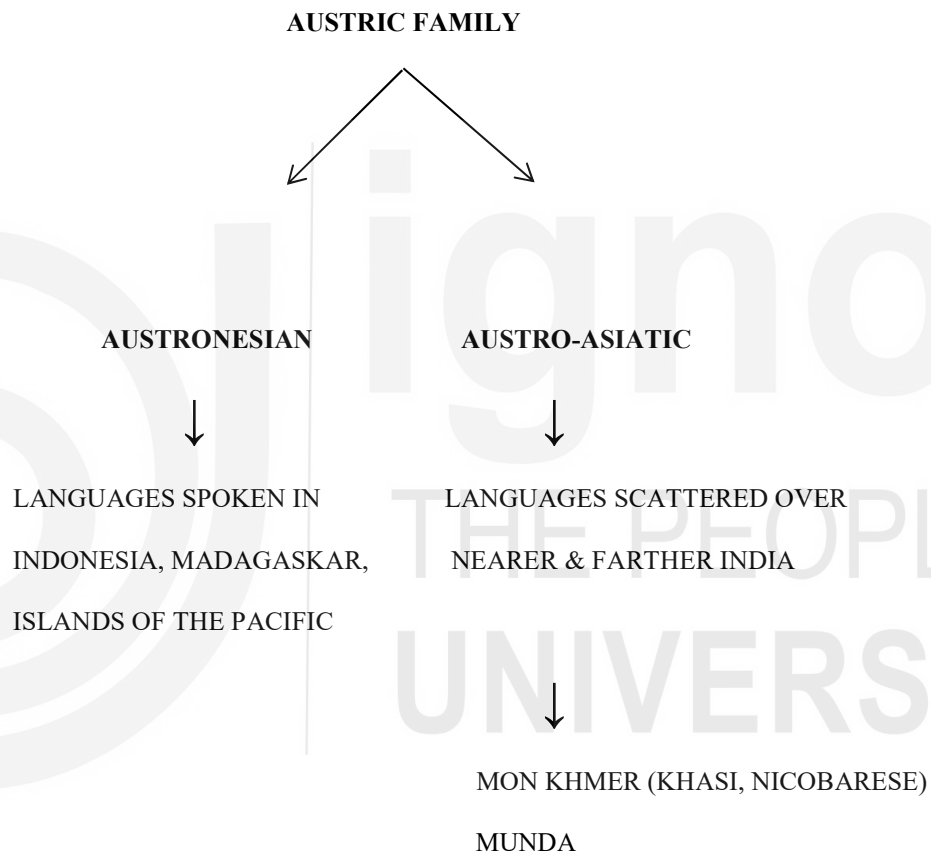
3.2 CLASSIFICATION OF AUSTRO-ASIATIC FAMILY

Beginning with Schmidt in the beginning of 20th century, with the recent classification by Diffloth (2005), the classification of Austro-Asiatic languages is modified again and again; the main drawback has been the lack of adequate data. With more and more literary and linguistic works, we are gaining more and more understanding of the languages of this family. In this section, we will discuss classification from Schmidt to Diffloth.

Wilhem Schmidt, an anthropologist, in a series of articles from 1904-1907 postulated an Austric superfamily composed of Austronesian (Malaypolynesian) and Austroasiatic. The former included

languages of Madagascar, Indonesia, and the islands of the Pacific, while the latter included languages scattered over Nearer and Further India (*LSI*, Vol. 1, p.32). Austro-Asiatic language family is strongly represented in India. Though the first suggestion of this language group was made by Logan (1856); and it was elaborated further by Forbes (1881), Muller (1888) and Kuhn (1889). But Schmidt (1906) *put it on a sound footing and proved the existence of a great family of languages hitherto not recognized* (*LSI*, Vol.1: 32). Schmidt's broad categorization can be presented as:

Fig: 3.1 Austric Language Family



Mon-Khmer, Munda, Khasi and some other languages were grouped under *Austroasiatic languages*; and Austronesic group included languages of Indonesia, Melanesia and Polynesia. Indonesian languages were assumed as a link between Austroasiatic and Austronesic languages. Thus the *Austric family* included these two groups which are further branched in many languages, 'the whole forming one great linguistic family, - called the 'Austric'- extending from the hills of Central India to Easter Island, off the coast of South America, and covering a wider area even than that of the Indo-European tongues' (*LSI*, Vol. I: 15).

Schmidt's sub grouping of Austroasiatic was criticized for its methodology (for initial groupings by Schmidt refer article by David P. Thomas (1964) *A Survey of Austroasiatic and Mon-Khmer Comparative Studies*). In 1926 Schmidt's refined classification was:

- Old Malacca (Semang, Senoi)
- Central (Khasi, Nicobarese, Palaungic)
- Southeast and Northwest (Mon-Khmer, Munda)
- Northeast Mixed (Cham, Sedang etc)

Grierson has discussed Mon-Khmer and Munda as distinct language families. Giving his acceptance *to a common substratum at the bottom of these tongues*, Grierson writes, 'the resemblance of Munda and Mon-Khmer and Nancowry vocabularies is so remarkable that a connection between these languages can not be doubted' (Varma, 1972: 109). He mentioned two important differences in these two sub groups:

- Syllabic Structure: Munda languages contain polysyllables while Mon-Khmer languages are monosyllabic (every word consisting of a single syllable).
- Word Order: Munda- S (subject), O (object), V (verb)

Mon and Khasi- S (subject), V (verb), O (object)

Nicobarese was considered by Grierson to constitute a connecting link between the Munda languages and Mon. The Nicobarese are the only Mon-Khmer speaking community to live isolated from the mainland. Six Nicobarese languages are distinguished: Car, Chowra, Teresa, Nancowry, South Nicobarese, Shampẽ. The most remarkable feature of Nicobarese remarked upon by researchers in the presence of a rich productive morphological system.

H.J.Pinnow (1959) provided a classification of the entire Austro-Asiatic family of languages. Pinnow's achievement is substantial, and marks a major advance in the field. His 1959 classification is as follows:

(A) West-Northwest: Nahali

(B) Nordwest: Munda

(C) Mon-Khmer: a) Mon , b) Kambodja-Sprachen (Khmer, Pear, Stieng and others), c) Chama-Sprachen (Ma, Chrau and others), d) Mnong-Sprachen, e) Bahnar-Sprachen, f) Sedang-Sprachen, g) Brao-

Sprachen, h) Jəru-Sprachen (Bolovent, Nhaheun and others), i) Kuoy-Sprachen, j) Suoy-Sprachen (Kaseng, Alak, Laveh and others)

(D) Nordost: Palaung-Wa-

a) West-Untergruppe (Riang, Palaung, Wa and others)

b) Ost-Untergruppe (Khmu, Lamet and others)

(E) Nord: Khasi

(F) Südwest: Nikobar

(G) Sömang-Pangan (Malaya)

(H) Sakai (Malaya)

(I) Jaku'd (Malaya)

An important feature of this classification is that Pinnow identifies a Mon-Khmer group consisting of 10 languages/groups of languages that are today recognised as Monic, Khmeric, Pearic, Bahnaric and Katuic (all geographically contiguous and under Indic influence in ancient times).

Gérard Diffloth (1974) classified Austro-Asiatic in Munda and Mon-Khmer which further have some subgroups:

- **MUNDA : North Munda:** Korku, Kherwarian;
South Munda: Kharia-Juang, Koraput Munda) and

- **MON-KHMER**

Eastern Mon-Khmer: Khmer (Cambodian), Pearic, Bahnaric, Katuic, Vietic (includes Vietnamese);

Northern Mon-Khmer: Khasi (Meghalaya, India), Palaungic, Khmuic; and

Southern Mon-Khmer: Mon, Aslian (Malaya), Nicobarese (Nicobar Islands)).

But in recent classification Diffloth (2005) has modified his earlier classification with three subgroups:

Table: 3.1 Diffloth's Classification of Austro-Asiatic Language Family 2005
--

Munda

Remo, Savara, Kharian-Juang, Korku, Kharewarian

Khasi- Khmuic (Northern Mon-Khmer)

Khasian: 3 languages of eastern India and Bangladesh Palaungo-Khmuic languages, Khmuic : 13 languages of Laos and Thailand, Palaungo-Pakanic languages, Pakanic or Palyu : 4 or 5 languages of southern China and Vietnam, Palaungic: 21 languages of Burma, southern China, and Thailand

Nuclear Mon-Khmer languages

Khmero-Vietic languages (Eastern Mon-Khmer) Vieto-Katuic languages, Veitic: 10 languages of Vietnam and Laos, including the Vietnamese language, which has the most speakers of any Austro-Asiatic language. These are the only Austro-Asiatic languages to have highly developed tone systems, Katuic: 19 languages of Laos, Vietnam, and Thailand. Khmero-Bahnaric languages, Bahnaric: 40 languages of Vietnam, Laos, and Cambodia, Khmeric languages

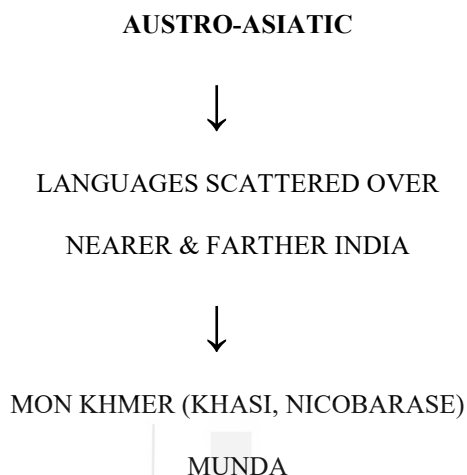
The Khmer dialects of Cambodia, Thailand, and Vietnam, Pearic: 6 languages of Cambodia, Nico-Monic languages (Southern Mon-Khmer), Nicobarese: 6 languages of the Nicobar Island, Asli-Monic languages, Aslian: 19 languages of peninsular Malaysia and Thailand., Monic: 2 languages, the Mon language of Burma and the Nyahkur language of Thailand

Other recent classification is provided by Ilia.Peiros (1998) that is based on similarities in the vocabulary of these languages.

The classification of Austro-Asiatic languages is complicated as the languages are spread over a wide area. The important thing about the classification is not the discrepancy among the views of various scholars but the methodology that they have adopted for the classification. Languages are classified genetically if they are descendants of a common ancestral language. Genetic classification of languages into a language family is based on an abundance of cognates (related words) in the member languages. Using these terms, one may treat the languages of the world according to the following geographic divisions: Europe, South Asia, North Asia, Southwest Asia, East Asia, Southeast Asia, Africa, and the Americas.

Thus following sub grouping is generally followed for Austro-Asiatic language family:

Figure: 3.2 Austro-Asiatic macrofamily



3.3 MUNDA LANGUAGES

‘The Munda languages were first recognized as a separate group distinct from Dravidian in the year 1854 by Professor Max Muller’ (*LSI*, Vol. 1, p.35). Prof Max Muller named this family-Munda. Munda languages constitute a sub group under Austro-Asiatic family. Chota Nagpur plateau (comprising the states of Bihar, Chattisgarh, Jharkhand and Orissa) was recognized as the principal home of the Munda languages. Munda languages are scattered over much of north-eastern, central and eastern India, Nepal and Bangladesh. The languages listed under Munda group (Grierson) are:

- **Kherwari**

Kherwari, i.e. the language of Kherwars, is the cover term used to designate a number of dialects spoken at the northeastern end of the Central Indian plateau as well as adjacent area. Santhali (or Santali), Mundari, Ho, Bhumij and Korwa are some the major dialects. Santhali is the only language from Austro-Asiatic family that is included in the languages under VIII schedule in the Constitution of India (2001 Census). Santhali is credited with the largest number of speakers for any Munda language. The total number of Santhali speakers in India according to 2001 census is 6,469,600. Largest numbers of Santhali speakers in India are in Bihar, Jharkhand and West Bengal with the presence in Assam, Tripura and Orissa also. Two dialects of Santhali are Karmali (or Kalha) and Mahle. Grierson writes, ‘these two dialects do not differ much from the standard form of speech.’ Mundari (2001 census - no. of speakers-1,061,352) is main habitat is Jharkhand and Orissa with

presence of speakers also in Assam, Bihar and West Bengal. Bhumij, Ho, Korwa, Kodaka/ Koraka, Birhor, Turi and Asuri are some of the dialects mentioned in *LSI*.

- **Korku** (574,481)

It is a Western most language of this group i.e., there is no other Munda language spoken to its west in India, mainly spoken by the Korku tribes of Maharashtra and Madhya Pradesh. It had two dialects- Muwasi and Nahali (Nahali is no more listed as a dialect of Korku in CIIL language list). For Nahali, Grierson notes it to be on the verge of extinction. 'It came under Dravidian influence and has become a mixed form of speech, half Munda and half Dravidian. This, in its turn, has fallen under the spell of Aryan tongues, and is now in a fair way to becoming an Aryan language' (*LSI*, Vol. 1, p. 29).

- **Kharia**

According to 2001 census number of Kharia speakers is 239,608. It is spoken chiefly in Ranchi District of Jharkhand. Speakers of this language can also be found in Bihar, Madhya Pradesh, West Bengal, Assam and Orissa; and some in Nepal also.

- **Juang**

Juang is the language of a Munda tribe settled in Dhenkanal and Koenjhar Districts of Orissa. Juang is largely influenced by the neighboring language- Oriya, an Indo Aryan language.

- **Savara (Sora)**

This is one of the language of South Munda group with Kharia, Juang and Gutob. It is mainly spoken in the Ganjam, Koraput, and Phulbani districts of Orissa, and the states bordering Orissa and also Assam. 'Savara has been influenced by Telugu and no more an unmixed form of speech. It is most closely related to Kharia and Juang, but in some characteristics differs from them and agrees with the various dialects of Kherwari. (*LSI: A Summary*, S. Varma: 1972).

- **Gadaba**

It belongs mainly to Koraput District in Orissa and also in some parts of Andhra Pradesh.

For classification of Munda Language learners can read:

- Bhattacharya, S. (1975). *A new classification of Munda*. *Indo-Iranian Journal* 17: 97-101
 - Anderson, Gregory D S (2001). *A New Classification of South Munda: Evidence from Comparative Verb Morphology*. *Indian Linguistics*. 62. Poona: Linguistic Society of India. pp.21-36.
- http://livingtongues.org/docs/Munda_Classification.pdf

3.3.1 Structure of Munda Languages

We will discuss some prominent features of these languages. The languages of this family have undergone immense change due to contact with the languages of other families (mainly Indo Aryan) and assimilated many features of the languages of different genetic stock. We will discuss mainly Munda features in this section.

The Sound System

Munda languages are rich in vowels. All the vowels can be short as well as long. Vowel harmony is found in all the languages of this family. In Kharia *jib* ‘to touch’, *jinib* ‘a touch’, *job* ‘to suck’, *jonob* ‘suck’ after the infix /n+vowel/ is inserted. The vowel of the infix is always the copied vowel of the stem (Abbi 2001: 34).

These languages are also rich in consonants. Grierson has mentioned the occurrence of hard and soft consonants and both the classes can be aspirated. Semi consonants form a characteristic feature of Munda languages. Phonemic occurrence of the glottal stop /ʔ/ is another prominent feature of these languages. (Examples given below are from Abbi 2001: 34)

Santhali: / u:p / - hair

/ ʔu:p / - put

Kharia: /borol-daʔ/ - water

/laʔ-ta/ - feels

Lexical and Morphological System

Comparing Munda languages with Turkish grammar, Grierson writes, ‘suffix is piled on suffix, till we obtain words which, to European eyes, seem monstrous in their length, yet which are complete in themselves’ (*LSI*, Vol I: 36).

Austro Asiatic languages are **highly agglutinating** in nature. In agglutinating languages, ‘there tends to be a more or less one to one matching of morphemes to morph’ (Katamba, 1993: 57). Words are formed from bases or other words by means of reduplication or by adding affixes; insertion of infixes in a root is

very common (Santhali- *dal*-to strike; *da-pa-l* - beating becomes reciprocal. Mundari- *ma-na-ran* – greatness; *maran* - great). See following example from Kharia, nouns are derived from verb bases by infixation:

Table 3.2 DERIVATION OF NOUNS FROM VERBS IN KHARIA

Verb	Gloss	Noun
<i>jib</i>	touch	<i>jinib</i>
<i>job</i>	suck	<i>jo-no-b</i>
<i>juŋ</i>	ask	<i>ju-nu-ŋ</i>
<i>kol</i>	count	<i>ko-no-l</i>
<i>deb</i>	climb	<i>de-ne-b</i>
<i>del</i>	come	<i>de-ne-l</i>
<i>joʔ</i>	sweep	<i>jo-no-ʔ</i>

(From Malhotra 1982 quoted in Abbi 2001: 157)

Various classes of words are not distinguished, ‘any word can be used for any concept’ (Pinnow, 1966: 101). It is the most prominent feature of the Munda languages that they do not distinguish between nouns and verbs as distinct word classes. Information of referent, modification, numeral, quantifiers etc is incorporated in verb. See the example below of a verb complex

Santhali - /εge:r-ked-ijŋ-a /

scold-pst-object marker-[me]-definite

X scolded me

(Abbi: 2001: 34)

Nouns do not differ for **gender**. Distinction of gender is made by adding different words for male and female. Broadly, nouns are classified as animate or inanimate, but it becomes further complex, as the criteria of animate and inanimate depends on specific community. As Abbi has observed (2001: 119)

‘In Munda animate nouns are intellectual beings, animals, super natural powers, and heavenly bodies. The criteria for deciding the gender for each noun vary from community to community, as the decision lies in the cognitive aspect underlying the specific speech community.’

There are **three numbers**: singular, dual, and plural. For example Kharia verb *qoko* – sit, will be marked for number as (From Malhotra 1982):

Table 3.3 NUMBER IN KHARIA VERB *qoko* – sit

	Singular	Dual	Plural
<i>I Person exclusive</i>	qoko-taŋ	qoko-ta-jar	qoko-ta-le
<i>I Person inclusive</i>		qoko-ta-naŋ	qoko-ta-niŋ
<i>II Person</i>	qoko-ta-m	qoko-ta-bar	qoko-ta-pe
<i>III Person</i>	qoko-ta-φ	qoko-ta-kiyar	qoko-ta-ki/moy

(in Abbi 2001: 35)

These languages are also **rich in personal pronouns**. These pronouns are distinguished for number, distance between the addresser and the referent and the person-addresser component.

Table 3.4 PERSONAL PRONOUNS IN SANTHALI

	<i>I Person inclusive</i>	<i>exclusive</i>	<i>II Person</i>	<i>III Person</i>
<i>Singular</i>	<i>i: ɲteɲ</i>		<i>a: mteɪ</i>	<i>ai?tte</i>
<i>Dual</i>	<i>ala: ɲt:ɲ</i>	<i>ali: ɲteɪ:ɲ</i>	<i>abenteɪben</i>	<i>akintekin</i>
<i>Plural</i>	<i>abutebu</i>	<i>aletele</i>	<i>apetepe</i>	<i>akuteku</i>

(ibid: 125)

Other features discussed (Abbi 2001) are:

- Lack the dichotomy between morphology and syntax
- Verb Complex incorporates all the information about the subject, indirect and direct object. Within a verb complex order of the morphemic elements is: Subject Marker + Verb + Tense + Object marker + Definitiser. (Verb complex follows SVO word order).
- But NP + VP structure has developed due to contact with languages of other families.

Syntactic System

The concept of verb complex in these languages makes sentence an altogether different entity from what we generally know a sentence is. A clear distinction between morphology and syntax is not possible for these languages. 'The logical form of Munda sentence is altogether different from that of Aryan languages, and hence it is impossible to divide it into the parts of speech with which we are familiar' (LSI, Vol. I: 37). One of the characteristic features of Munda grammar is the use of particle 'categorical *a*.' For example, 'John came' is a representation of some meaning that is mentally represented, but a Santhali speaker would add the categorical *a*, to say that it is a fact or in other words, the reality of the event is denoted by suffixing categorical *a*.

- The features of Munda languages discussed above are some major features, for detailed study of features, learners can consult literature on these languages:

Zide, Norman 1966. *Studies in comparative Austro Asiatic Linguistics*. Mouton, The Hague

Linguistics Information Services, India website www.lisindia.net

We will discuss about change in Munda languages due to contact later in the unit. Languages as we know have never been static entities. Languages are entities that have vibrant life. They grow, develop, and are influenced by various factors. The fact that the languages that we have discussed have become marginalized, due to policies, due to technology and lack of development (there may be various factors involved for the lack of development) and even if development takes place, their languages might suffer more, as these are not the languages used in schools, work places and literatures. How can these

languages get the benefit from translation, if good texts are made available in their languages? Think and discuss.

3.4 MON-KHMER LANGUAGES

“The MK languages form the oldest and most diverse language family of Mainland Southeast Asia. MK speakers colonized all available ecological riches the region has to offer, from mountain slopes to tropical islands. Along the way they diverged widely; some barely subsisted in near-stone-age conditions, while others built great civilizations still recognized as hallmark achievements of mankind. Enriched by their complex history of contact with other Asian societies, the MK languages present a deeply evocative record of the region that can be revealed by comparative-historical analysis.” (Sidwell 2007: 1)

Languages of Mon-Khmer subfamily are spoken on Indonesia Peninsula, in Southern China and Northeastern India and on the Nicobar Islands. The name is derived from two ancient literary languages, Khmer and Mon. ‘The German comparativist and ethnologist Peter Wilhelm Schmidt (1868–1954) effectively established the field of comparative MK studies with a series of four major publications at the beginning of the twentieth century (1901, 1904, 1905, 1906). Taken as a whole these four works form a more or less coherent account of the MK family as it was known at the time.’ (Sidwell 2007: 2) He showed that Mon-Khmer languages form a link between the Munda languages of India proper and the languages of Indonesia. “Today it exists as a patchwork of more than a hundred languages spread across an area that ranges from eastern India to Vietnam, and from Yunnan to Malaysia and the Nicobar Islands of the Andaman Sea. In only two countries are Mon-Khmer languages the official tongues: Cambodia (Cambodian) and Vietnam (Vietnamese). In all other cases they are minority languages, ranging from communities with only hundreds of speakers (e.g. such as some Aslian languages) to communities of more than a million (e.g. Khasi, Katuic etc.).”

(Sidwell <http://people.anu.edu.au/~u9907217/languages/languages.html>)

As classification is already discussed, we are giving here only the names of the languages:

(1) **Southeastern** (in Cambodia, Vietnam, Thailand, Laos, and Southern Burma): Khmer (the official language of Burma), Mon, Poar, Chong, Stieng, Mnong, Sre, Ma, Chrau, Bahnar, Brao, Chru, Kui, Katu, Souei, Bru and Pacoh.

(2) **Northeastern** (in Laos, Thailand, Burma, and Southern China): Palaung, Wa, Kawa, Lawa, Rein, Benlun, Danaw, Khmu, and Lamet

(3) **Northwestern** (Northeastern states in India): Khasi

(4) **Southwestern** (Nicobar Islands): Nicobarese and its numerous dialects

Pointing out the similarities and differences of Mon-Khmer and Munda languages, Grierson (LSI Vol. 2, p.1) mentions the vocabulary of these languages (together with Nancowry of Nicobar Island). He observed that the similarities between vocabularies 'are so remarkable and of such frequent occurrence, that a connexion between all these tongues cannot be doubted', but these two groups differ in two important aspects (From LSI Vol. 2, p.1-2):

- **Word Structure:** The Mon-Khmer languages are monosyllabic. Every word consists of a single syllable. When, in Khasi for instance, we meet an apparent disyllable we find on examination that it is really a compound word. On the other hand, the Munda, Nancowry, ... languages contain many undoubted polysyllables.
- **Word Order:** Munda languages follow SOV word order, while in Khasi and Mon it is SVO.

3.4.1 Linguistic Features of Mon-Khmer Languages

Study of this language group has come a far way since the foundational work of Wilhelm Schmidt-developments in mid 20th century (1960s and 1970s), from pre-computer era to the present time when the work has comparatively become easier (it is nonetheless simple but IT has surely made it less time consuming and less complicated), and with the publication of Shorto's *A Mon-Khmer Comparative Dictionary* (2006), 'a wealth of extensive descriptive and lexical data is now available'. (Sidwell 2007)

Under MK group we will discuss Khasi in detail, but briefly overview the features of MK languages on the whole:

- The set of phonemes in MK languages can be quite large and have some unusual types of articulation in comparison to other languages of the world, including imploded and/or glottalised consonants, voiceless nasal and lateral sounds, and there may be tones or so-called 'registers', distinguishing breathy and/or creaky vowels. Among these tongues we find the largest inventories of vowels of any language in the world, with some having dozens of monophthongs and diphthongs.

- Structurally Mon-Khmer languages tend to have 'sesquisyllabic' word structure. Words may begin with clusters of three or even four consonants, but the set of possible word finals is always much more restricted. The range of word shapes in a typical Mon-Khmer language is illustrated with the following examples from Jru' (West Bahnaric):

CV *kaa* 'fish'

CVC *daak* 'water'

CCVC *plaj* 'fruit'

CCVC *pteh* 'earth'

CCCVC *kbreh* 'blink'

- The languages are analytical in their grammar, with only a modest morphological systems, utilising prefixation and infixation for mostly derivational purposes. Historically there is an evident tendency towards morphological simplification, an extreme example being Vietnamese—its lexicon is basically monosyllabic and it relies upon word order and compounding to express grammatical functions.
- Syntactically MK languages have characteristics typical of the SEAsian linguistic area, such as:
 - Subject-Verb-Object word order
 - Noun Classifiers
 - Serial Verb Constructions
- Generally MK languages show a strong tendency to assimilate syntactically to their areal neighbors. The more distant (geographically and genetically) Munda languages are markedly different in many respects.

3.5 KHASI

Khasi is the official language of Meghalaya and is spoken mainly in northeastern states. Khasian (or Khasic) is the umbrella term that is used to refer to the languages spoken in Meghalaya and neighbouring states and in Bangladesh which includes dialectal varieties of Khasi. *Standard Khasi* (or *Cherrapunji Khasi*) is most documented and relied for linguistic studies. Cherrapunji (or Sohra to the Khasis) is the home of the standard Khasi as it was the place where London Missionary Society and Welsh Calvinistic Methodist Mission (or Welsh Presbyterian Mission) established a number of schools in mid eighteenth century (in 1840s).

The number of speakers of Khasi language in India (2001 census) is 1,128,575. The following Table 3.5 shows the groups under Khasi language:

Table: 3.5 Various Mothertongue groups under Khasi

KHASI	SPEAKERS
	1,128,575
Bhoi Khasi	14,882
Khasi	828,545
Pnar / Synteng	243,441
War	25,886
Others	15,821

(From: www.lisindia.net/Khasi/Khasi)

Khasi is differentiated from the surrounding non-Indo-Aryan languages of India in its complete system of gender, expression of grammatical relations by prefixes (suffixes in Dravidian, Munda and Tibeto-Burman), possessor being placed after the thing possessed, and the presence of relative pronoun.

3.5.1 Linguistic Features

Sound System

The Standard Khasi vowel inventory is typically Mon-Khmer, with a full set of short and long vowels. 'The sounds in Khasi are relatively few, and a 23 letter alphabet was, and is considered to be adequate for the spelling of words in common use in the language. These letters are: *a b k d e g n g h i j l m n ñ o p r s t u w y*. There is no velar voiced stop / *g* / in Khasi. It was introduced in Khasi by the inclusion of foreign words spelt with this phoneme e.g. *gadha* 'donkey' in the earlier translations of the New Testament.it is interesting that while a 'foreign' phone could be used in the beginning of the word, the medial consonant *dh* was replaced by *dd*. It is true that segment *dh* is very rare in Khasi words, yet it does exist, although in borrowed words, initially. But the initial phoneme / *g* / the word *gadha* has long passed into disuse and replaced by more natural Khasi phoneme / *k* / - (*kada*).' (Simon 2008: 63)

Certain sounds do not normally occur at the end of words viz. /*l*, *s*/, fricatives, aspirates, conjunct consonants and voiced stops. The frequency of occurrence of palatal and velar nasals is striking feature. Khasi also offers various kinds of initial consonants cluster paradigms with the initial velar [ŋ]. Like other Austro-Asiatic languages glottal stop is phonemic /*ka-waʔ*/ - 'river'. [f] is not present in the language. But it is commonly used in free variation with /*pʰ*/, for example word *football* is pronounced as *phutbɔ:l*.

In Khasi, consonant clusters occur only in the initial position whereas in the final position no clusters occur at all, for example, *br* - *brɛt* - 'throw', *tw* - *twa* - 'collapse' etc.

The Khasi language has a special feature called *onset simplification*. Onset simplification is the process of reduction of initial consonant(s) of a syllable in a word. Onset simplification in Khasi takes place in noun roots which occur in a compound. The condition for onset simplification is that the root should be monosyllabic and have consonant clusters in the initial position. This is the minimum requirement for onset simplification. (www.lisindia.net)

For example:

Underlying form		Surface form	Gloss
b re:u + sta:d	→	re:usta:d	"wise person"
bre:u + spa:ʔ	→	re:uspa:ʔ	"rich person"

Lexical and Morphological System

Khasi would be considered as an Analytic language, having the following structural criteria:

1. All the words (Nouns, Verbs, Adjectives, etc) are invariable.
2. There is no Inflectional Morphology, i. e, Nouns, Verbs, Adjectives etc are not inflected.
3. Words do not reflect grammatical relationships morphologically.
4. Syntactic and grammatical relationships are primarily shown by word order.
5. Words are mostly monomorphemic, a property of most isolating languages.
6. There are very few class maintaining and class changing Derivational prefixes.
7. Because of the absence of Inflectional Morphology in Khasi, particles, especially proclitics, have important grammatical functions. Particles, defined as invariable terms with grammatical functions, form closed sets of linguistic elements different in structure and functions. They resemble words, but cannot stand on their own, being structurally dependent on a neighboring word, hence they are clitics.

The following are some examples of particles in Khasi:-

Negative particles 'ym/-m' (not), 'Em' (no), 'Khlem' (not, past).

Modal Verb particles 'dei' (must), 'lah' (can/may), 'hap' (must).

Past participles 'lah' (lah wan: came)

8. **Number:** In Khasi cannot be defined as word classes displaying morphological contrasts such as boy (singular), boys (Plural), since 'Number' in Khasi is indicated by the participles, where 'ngi' indicate more than two.

	Singular	Plural
First Person:	'nga' (I)	'ngi' (we)
Second Person:	'pha' (you, F, intimate/non-formal 'me' (you, M, intimate/non-formal	'phi baroh' (you all) }
	'phi' (you, respect/formal)	
Third Person:	'i' (he/she, respect, diminutive) 'u' (he, M)	'ki' (they) }
	'ka' (she, F)	

Non-countable or Mass Nouns are treated as 'singular' e.g. u khaw (rice), ka ja (cooked rice) ka um (water), u shyiap (sand) ('u', 'ka' are singular and 'ki'-plural (they)).

9. **Gender:** Gender is indicated in singular by the 'article' (The pronoun of the third person is commonly described as article. Its forms are, singular, mas. *u*, fem. *ka*, diminutive or familiar, *i* ; plural (common gender) *ki*.); and in plural by adding words denoting gender. The great majority of inanimate and abstract nouns are feminine.
10. **Adjectives** are formed by prefixing *ba* to the root. Thus, *bha*-goodness; *ba-bha* –good. The adjectives always follow the nouns.

Verbal root may be simple or compound. Tense and mood are indicated by prefixes. When the subject is a noun, the pronoun is inserted before the verb.

u-mohan *u-rukhey*
msg clt-Mohan msg clt-laughs-msg

Mohan laughs

ka-bharti

ka-ruai

fsg clt-Bharti

fsg clt-sing-fsg

Bharti sings

(Abbi: 2001: 141)

11. **Order of Words:** The usual order of words is SVO, but sometime (for emphasis) verb is put before the subject.

3.6 Other Related Issues

Knowing about languages also involve knowing about the phenomenon of contact and convergence. Languages are dynamic and living entities. Structural properties of languages do change due to contact with other languages. The processes of change and assimilation of features varies for different languages. But all the aspects of grammar are susceptible to change, some relatively easy to change (lexicon) and other relatively take much longer to change. Whether changes are in positive direction or negative, it may be a topic of debate among scholars, but the fact is that languages do change. The processes of contact and convergence are more important for the languages which are minority languages (in Indian reference not scheduled in Indian Constitution) and for the languages which do not have wider acceptance as medium of instruction or medium of trade outside a limited area. Most of the Munda and Mon-Khmer languages constitute minority or tribal languages. The daunting task for scholars is to keep these languages alive and enrich them. In this context, Khasi written literature is an example which has grown in great measure in spite of it being born only a little over a century and a half ago. It is due to the writers who started translating texts of other language into Khasi. 'The year 1896 marked an important milestone in the realm of development of Khasi literature. It was on this particular year that translation works were carried out by a prominent Khasi writer U Jeebon Roy Mairom, known to Khasi literary world as the father of modern Khasi literature. He translated some classical Indian literature into Khasi' (Marak & Shangliang 2008: xii).

Khasi is compared to an island representing MK family surrounded by the sea of Indo-Aryan and Tibeto-Burman language. This sea is coming into the island in the form of waves of words and breezes of sounds and all is getting assimilated into the veins of Khasi language. It is interesting to note that a bulk of Khasi vocabulary has come from borrowing. Khasi vocabulary has enriched itself by borrowing vocabulary

items, coinage of new words and through the productivity of derivational morphemes. See some examples: (War, 2008: 20-21)

(a) Borrowing

Trade and Commerce – Pisa, tyngka, hazar, pawa, mon, ser, etc. (Indo-Aryans); Kilo, ton, bang (bank) (English).

Administration – dorbor, myntri, jylla, muluk, riti, muhor, dakhhol, ilaka etc. (Indo-Aryans); Ophis, elecshon, komiti, bud (vote) etc.

Philosophy and religion – dujok, niam nusip, hok, pap, sot, duwai, duk, nurok ijat etc. (Indo-Aryan); phadar, krismas, haleluia etc (English).

Borrowed words (lexical items) are adapted to the phonological system of the language- *bek* (bag), *ben* (bench) *ketli* (kettle), *pulit* (police) etc. Some words are translated into the target language, as

Korshujain (Sewing Machine), Narsdeh (Frying Pen), Jaisiang thiah (Bed sheet), Jainsiang meij (Table cloth)

(b) Coinage of New Words

Shongknor (Chairman), Kongsan (Chief Guest), Sengbhalang (Social welfare organization), Jiamphang (Iambic metre) Namding (Missile) Dheng Kali (Traffic Jam), Lad Pathai Khubor (Media) etc.

Munda languages also show change, borrowing and assimilation at various levels. 'In the case of Munda and Dravidian languages of Jharkhand, one can safely say that 50% of the basic words are replaced by Indo-Aryan words' (Abbi 2001: 46). Presence of aspirates and retroflex sounds in Munda languages shows influence of Indo Aryan languages. IA infinitive suffix *na* in Munda languages (Kharia and Santali) shows the impact on language morphology. Munda languages have also deviated from earlier lack of distinction between morphology and syntax to analytic type (NP+VP constructions).

The processes of contact, convergence and retention in a multilingual society like India are important issues to be explored further. What are the changes that have been incorporated and what is the direction of change, from which language other languages are taking elements. All these issues involve institutional, social and political mechanisms. These issues are dealt in with the branch of linguistics-

sociolinguistics. We can end our discussion with the lines from Sidwell for Mon-Khmer language but it can be generalized on other languages too:

'The conclusion to be drawn from reviewing the past century of work on the MK languages is surely that no single methodology or resource (and certainly no individual researcher!) is sufficient to reconstruct the history of this extraordinarily ancient and diverse language family. But by the same token, no method, resource, or researcher should ever be seen as being marginal enough to be safely ignored. Again and again it is the combination of method, resources, and individual drive that ultimately leads to lasting results' (Sidwell 2007).

3.7 Let Us Sum Up

In this unit we discussed Austro-Asiatic language family which forms a subgroup in Austric language family. Two main groups of Austro-Asiatic are Mon-Khmer and Munda languages. Mon-Khmer languages are mainly spoken in Southeast Asia and Khasi is the only Mon-Khmer language represented in India. Munda languages are mainly spoken in eastern and Northeastern states of India. Recent classifications, publication of dictionaries, growing number of literary texts have added and will also add more to the linguistic studies of the languages of this family.

3.8 Glossary

AUSTRIC FAMILY- superfamily composed of Austronesian (Malaypolynesian) and Austroasiatic. The former included languages of Madagascar, Indonesia, and the islands of the Pacific, while the latter included languages scattered over nearer and further India.

AUSTRO-ASIATIC- the sub group under Austric Family and it includes Mon-Khmer and Munda and other groups of languages. This language family has its presence mainly in Southeast Asia.

NICOBARESE- the only Mon-Khmer speaking community to live isolated from the mainland-spoken in Nicobar Island. Six Nicobarese languages are distinguished: Car, Chowra, Teressa, Nancowry, South Nicobarese, Shampẽ.

VOWEL HARMONY- found in all the languages of Austro-asiatic family. Infix /n+vowel/ is inserted to form noun from verbs. The vowel of the infix is always the copied vowel of the stem, e.g. Kharia *jib* 'to touch', *jibib* 'a touch', *job* 'to suck', *jonob* 'suck'

AGGLUTINATING LANGUAGES- one to one matching of morphemes to morph is found.

CATEGORICAL ‘a’ -One of the characteristic features of Munda grammar is the use of particle ‘categorical a’, a speaker would add suffix *a* to a sentence (utterance) to denote that the event is fact.

MON-KHMER- MK Languages are spoken on Indonesia Peninsula, in Southern China and Northeastern India and on the Nicobar Islands. The name is derived from two ancient literary languages, Khmer and Mon.

SESQUISYLLABLIC- in sesquisyllabic word structure words may begin with clusters of three or even four consonants, but the set of possible word finals is always much more restricted.

KHASI - the official language of Meghalaya and is spoken mainly in northeastern states. Khasian (or Khasic) is the umbrella term that is used to refer to the languages spoken in Meghalaya and neighbouring states and in Bangladesh which includes dialectal varieties of Khasi.

STANDARD KHASI (or **CHERRAPUNJI KHASI**) is most documented and relied for linguistic studies.

ONSET SIMPLIFICATION- the process of reduction of initial consonant(s) of a syllable in a word. Onset simplification in Khasi takes place in noun roots which occur in a compound. The condition for onset simplification is that the root should be monosyllabic and have consonant clusters in the initial position.

3.9 Exercise

1. Who identified the Austric family as a distinct linguistic family?
2. Write the names of macro-families under Austric Family.
3. How many major groups are there in Austro-Asiatic languages?
4. What are the two differences between Munda and Mon-Khmer languages discussed by Grierson?
5. There are 22 languages listed under VIII Schedule, Article 343-351 of the Constitution of India. Find out how many languages from Austro-Asiatic language family are listed in our constitution and write the name of the language/s?
6. Which Munda language is spoken in the states of Maharashtra and Madhya Pradesh?

7. Under Kherwari languages how many languages are listed in Linguistic Survey of India (Grierson). Write their names.
8. Who was the first to recognize Munda languages as a separate group?
9. What is vowel harmony? Give an example.
10. What do understand by agglutinating language? Why Munda languages are agglutinating?
11. Write the personal pronouns used in Santhali. Q.4 Explain the statement: Munda languages lack the dichotomy between morphology and syntax
12. What do you understand by 'sesquisyllabic' word structure'?
13. What do you understand by the term 'monosyllabic'?
14. How does Munda and Mon-Khmer differ in word order?
15. In which parts of the world are Mon-Khmer languages spoken? Write two languages from Mon-Khmer languages that are represented in India.
16. Give classification of Austro-Asiatic language family.
17. Name the languages included in Munda languages. Also discuss structural features of Munda languages.
18. In what ways Mon-Khmer languages differ from Munda and other non Indo Aryan languages?
19. What are the languages (language groups) included in Mon-Khmer languages? Why is the classification of Mon-Khmer languages more complex than Munda languages?
20. Write main features of Khasi language.
21. It is observed that the languages of Indo Aryan family have affected the language of other families more than their being affected by other language families. Give two reasons for unidirectional flow of change, if it exists.

3.8 Suggested Readings and References

- Abbi, Anvita (2001) *A Manual of Linguistic Field Work and Structures of Indian Languages*. Lincom Europa
- Asher, R. E. (2008) *Language in Historical Context*. In Kachru, B. B., Kachru, Y. & Sridhar, S. N. (2008). *Language in South Asia*, pp. 31-48. Cambridge: CUP
- Diffloth, Gérard. 2005. *The contribution of linguistic palaeontology to the homeland of Austroasiatic*. In: Sagart, Laurent, Roger Blench and Alicia Sanchez-Mazas (eds.). *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. Routledge/Curzon. Pp. 79-82.
- Grierson, G. A. (1927) *Linguistic Survey of India*. Vol.I. (Part-I)- Introductory . Reprint-1990: Delhi
- Grierson, G. A. (1928) *Linguistic Survey of India*. Vol.2. Mon-Khmer and Siamese-Chinese Families (including Khasi and Tai). Reprint-1990: Delhi
- Kachru, B. B., Kachru, Y. & Sridhar, S. N. (2008). *Language in South Asia*. Cambridge: CUP
- Malhotra, Veena (1982) *The Structure of Kharia: A Study of Linguistic Typology and Language Change*. Unpublished Ph.D Thesis. Jawaharlal Nehru University, New Delhi
- Marak, C. R. & Shangliang, J. S. (ed.) (2008). *Growth and Development of Khasi and Garo Languages*. New Delhi: Sahitya Academi
- Simon, I. M. (2008). *Loan Words in Khasi: Orthography and Pronunciation*. In Marak, C. R. & Shangliang, J. S. (ed.) (2008). *Growth and Development of Khasi and Garo Languages*, pp. 61-77. New Delhi: Sahitya Academi
- Thomas, David P. (1964) *A Survey of Austroasiatic and Mon-Khmer Comparative Studies*. In *Mon-Khmer Studies 1*: 149-163 (<http://sealang.net/sala/archives/pdf4/thomas>)
- Varma, S. (1972). *G. A. Grierson's Linguistic Survey of India: A Summary*. Hoshiyarpur (Punjab): Vishveshvaranand Institute
- War, B. (2008). *The Growth and Development of Khasi Language: A Linguistic Perspective*. In Marak, C. R. & Shangliang, J. S. (ed.) (2008). *Growth and Development of Khasi and Garo Languages*, pp. 18-22. New Delhi: Sahitya Academi

Web Resources

- Paul Sidwell (2006) <http://people.anu.edu.au/~u9907217/languages/languages.html>
- Paul Sidwell (2007) *Comparative Mon-Khmer Linguistics in the 20th Century: Where From, Where To?* <http://sealang.net/monkhmer/sidwell2007comparative.pdf>
- Linguistic Information Services, India (CIIL) www.lisindia.net



UNIT 4: CHINESE –TIBETAN AND OTHER LANGUAGE FAMILIES

Unit Structure

4.0 Objectives

4.1 Introduction

4.2 Chinese –Tibetan Languages

4.2.1 Chinese

4.2.2 Tibeto–Burman

4.2.2.1 Number of Languages and Language Names

4.2.2.2 Classification

4.2.2.3 Languages of Northeast India

4.3 Structural Features

4.3.1 Structural Features of Tibeto-Burman Languages

4.4 Other languages Families of World

4.5 Let Us Sum Up

4.6 Exercise

4.7 Glossary

4.8 Suggested Readings and References

4.0 Objectives

In this unit our objective will be to discuss:

- About the two main groups of Chinese-Tibetan languages: Sinitic and Tibetan, about their geographical spread, classification of both the groups and the structural properties of these

languages

- About other language families of the world.

4.1 Introduction

In this unit we will discuss about the fourth language family represented in India, that is Chinese-Tibetan language family. Together with this, we will also briefly outline other language families in the world. Chinese-Tibetan language family (also called Sino-Tibetan or Tibetan-Chinese) is one of the largest language families in the world with the largest number of first language speakers. Today this group exists in two sub-groups: Sinitic (or Chinese) and Tibeto-Burman. Sinitic refers to the various Chinese languages and Tibetan languages are spoken mainly in Tibet, Myanmar and North East India. It is the Tibetan part of these languages that is represented in India (North-East India). There are no records of the original (or parent) Sino-Tibetan language but through the similarities of languages, linguists have reconstructed the Proto-Sino Tibetan from which the later forms (the forms we know today) have been assumed to be derived. Languages as you know are dynamic entities; they borrow, give and change with various influences and with the flow of time what we see must have taken years to come into that shape. The same is true for this language family, ‘What was originally a single language has developed into vast, diverse family of language under the pressure of natural change, intermingled with frequent and often intimate contact with speakers of other languages’ (Thurgood 2003: 3).

4.2 Chinese-Tibetan Languages

The languages of this family are spoken in Northeast India, Myanmar, Bangladesh and northern Thailand, Tibet, China, Korea and Taiwan. Sino Tibetan (ST) speakers are associated with the ‘Neolithic Yang-Shao culture which originated in Yellow river valley in the central plains of China’ (Thurgood 2003: 3). Later this family split into Sinitic (Chinese) and Tibetan. The family has come to be the way it is because of multiple migrations, often into areas where other languages were spoken (LaPolla, 2001). Proto Sino Tibetan is assumed to be spoken in Yellow river valley around 6000 years ago.

Sino-Tibetan has the second largest number of speakers (after Indo-European) due largely to the over one billion Sinitic speakers; except for Burmese, most Tibeto-Burman languages have relatively few speakers (LaPolla 2006: 393).

Commenting on the vastness of ST languages, Grierson writes:

‘Excepting the Austric, no great family of speeches is spoken over so wide an extent of the Eastern Hemisphere from Central Asia to Southern Burma, and from Baltistan to Peking – as that formless, ever

moving, ant-horde of dialects, the Tibetan-Chinese. The number of its speakers far exceeds those of the Austric and even those of the Indo-European family. So vast is the area covered by it, and so apparently infinite is the number of its members, that no single scholar can hope to master it in its entirety. A few of them, such as Tibetan, Burmese, Siamese, or Chinese have been more or less thoroughly investigated by specialists; of others we have only a few words, single bricks, each of which we have to take as specimens of an entire house; while of others, we know only names, or not even that' (1927: 40).

Matisoff (2003) has grouped the languages of this family into the 'Sinosphere' and 'Indosphere' due to the linguistic and political influence of China and India respectively on the groups of these languages. Subgrouping and classification of Sino-Tibetan is highly controversial, mainly due to (LaPolla 2006):

- i. the differences in criteria for subgrouping
- ii. the paucity of reliable data (particularly on morphosyntactic patterns)
- iii. the fact that the development and distribution of these languages has been greatly influenced by migration and language contact

Apart from Sinitic and Tibetan groups, Tai-Kadai (Zhuang-Dong) and Hmong-Mein (Miao-Yao) languages were also used to be included in ST languages but the similarities among these groups are now believed to be a result of contact influence.

Sino-Tibetan is commonly thought to consist of two primary branches: Sinitic and Tibeto-Burman. Both the group are further classified to include many subgroups. At first we will discuss the classification of Chinese group of languages and then move to Tibeto-Burman classification.

4.2.1 Chinese

Chinese family of languages is a group of related dialects. The dialect groups include many dialects which are sometimes non-mutually intelligible. In the words of Norman (2003: 72), "Chinese is a vast dialectal complex containing hundreds of mutually unintelligible local varieties, each of which can be viewed as a distinct object for comparison." Norman comments that the term dialect is used in different context for Chinese languages from that of the European context. In Chinese context local forms of speech are referred to as dialects of a particular place. One reason for using the term dialects for Chinese languages lies in their use of common script and sharing a common literary and cultural history (Crystal,

1996). Generally agreed number of dialect groups in Chinese is six (seven and eight dialect groups are also mentioned). These groups of dialects are initially distinguished on the basis of the reflexes of the historically voiced initial consonants (Li, 1936-1937). These groups are (LaPolla, 2006) :

- i. Mandarin (northern and southwestern China)
- ii. Wu (Jiangsu and Zhejiang)
- iii. Xiang (Hunan)
- iv. Gan (Jiangxi)
- v. Yue (Guandong and Guangxi)
- vi. Min (Guangdong, Fujian, Hainan Island, and Taiwan)

Some scholars classify the Hakka group (Guangdong, Fujian, Jiangxi, Sichuan and Taiwan) as a part of Gan group and others identify it as a separate group.

Norman (1988, 2003) has further grouped these dialect groups as

- i. Northern group (Mandarin)
- ii. Central group (some Xiang dialects, Wu and Gan), and
- iii. Southern group (Yue, Hakka and some Xiang dialects)

Min group was not included in the above classification by Norman as in his view the Min dialects lay 'outside the mainstream of Chinese linguistic development'. Mandarin has the largest geographic spread and population while Min does not have a large spread and population. More commonly the word 'Chinese' is used for the national standard language: *Putonghua* (*P<<ut – onghu \ a*) which literally means common language. It is based on the phonology of the dialect of Beijing and the lexicon and the grammar are based on northern vernacular. The movement for developing a common language was started in the early 20th century, the reason for the movement were

- i. to simplify the characters of classical written Chinese
- ii. to provide a single means of communication
- iii. to introduce a phonetic alphabet

In 1956, Putonghua became the medium of instruction in all schools. It is now most widely used form of spoken Chinese and is the normal written medium for almost all kinds of publications (Crystal, 1996: 315).

For more on Chinese language learners can refer :

Norman, J. (1988) *Chinese*. Cambridge: CUP

Norman, J. (2003) *The Chinese Dialects: Phonology*, p. 72-83 In Thurgood and LaPolla (eds.). (2003). *The Sino-Tibetan Languages*. London & New York: Routledge

4.2.2 Tibeto-Burman

The other group of Sino-Tibetan language family, the Tibeto-Burman family, is the largest of any languages family in South Asia in terms of number of languages. These languages are spread over a vast geographical area, ranging from Jammu and Kashmir in the west to Northeastern India, Nepal, Myanmar, Tibet, China, Thailand, Vietnam and Laos.

In India, densest concentration of Tibeto-Burman is found in the Northeast. “By comparison with most Indian states, the seven states of the northeast are small in area and low in population, but even by Indian standards they are very high in linguistic and ethnic heterogeneity.The great majority of the indigenous people of Arunachal on the north and of Nagaland, Manipur and Mizoram in the east speak Tibeto-Burman languages” (Burling, 2003: 169). Kokborok was once the dominant language of Tripura. Meghalaya is the only northeastern state where Khasi language of Mon-Khmer family (Austro-Asiatic language family) is spoken. Garo is the Tibeto – Burman language spoken in western Meghalaya except Khasi. In Assam, where dominant language is Indo-Aryan Assamese, Tibeto-Burman speakers can be found in minorities.

4.2.2.1 Number of Languages and Language Names

The number of languages in the NE states of India, Burling writes, cannot be counted because some regions are not known at all and even in better known areas it is difficult to distinguish between language and dialect. The answer for how many languages are there varies from author to author. In Burling’s words, “Estimates of the number of languages spoken in Northeastern India run as high as two hundred, but I suspect that this number can be reached only by counting a good many mutually intelligible dialects as separate languages. Nevertheless, even the most modest estimate would have to recognize at least fifty or sixty languages.” (2003: 171).

Names of languages also add to difficulty in knowing about languages. There are multiple names for the same language and dialect: name by which speakers of the language know their language (autonym), and what others call them (exonyms), there are old names (paleonyms) and new names (neonyms), and languages are also known by name of the village or nearby river where the language is spoken (loconyms). Matisoff writes, 'the 40,000 speakers of Lotha Naga are called Chizima, Choimi and Miklai by the neighbouring Angami, Sema and Assamese respectively. Conversely, quite different languages are known by the same or very similar names: Nung is both a central Tai language and a Tibeto Burman language of Nungish group; Mru is a TB language of the Kuki-chin group, but Maru, also TB, belongs to the Burmish group; Kham is both a dialect of Tibetan and a separate language of central Nepal.' Sometimes the names are in both a narrower and a broader sense, i.e. it may refer to one specific language or to the whole group of linguistically and culturally related languages. 'Terminology is made even more difficult by confusion between ethnic and linguistic affiliation' (Burling, 2003: 171). Burling discusses some names as misleading as 'Kamrupan' (Matisoff 1991) and 'Naga'. In Burling's views, the former groups' languages as if there is some genetic relatedness between the languages when there is no evidence of such grouping have been found. Naga is a term which denotes ethnic identity and the languages spoken by Nagas fall into at least two or possibly several completely distinct branches of TB.

4.2.2.2 Classification

The existence of Tibeto–Burman family was put forward in 1850s when many words in Written Tibetan were found to be cognate to the forms in Written Burmese. The first attempts to study these languages were by British administrators in India and Burma. Robert Shafer's (1955) classification of Sino–Tibetan gave systematic account of the languages for the first time. He assembled all the material then available on TB languages and presented a division of the family into subgroups. Rather than recognizing an autonomous group of Tibeto–Burman, Shafer presented six primary divisions of the Sino–Tibetan family. These divisions are:

- i. The Karenic (Central and South Burma)
- ii. The Baric (Assam)
- iii. The Burmic (Indo – Burman Frontier, Burma, Indo China, Tibet and South West China)
- iv. The Bodic (stretching from western Himalayas through Nepal and into Assam, Tibet and Western China)
- v. The Daic (West China, Tonkin, Laos, Thailand, parts of Burma)
- vi. The Sinitic (China)

In this classification ‘Tibetan’ languages are included in the Bodic division. The Baric, the Bodic and the Burmic divisions represents the languages spoken in South Asia. Under each division are sections and branches. Here we do not go into the details of the classification given by Shafer; we will discuss recently offered classification of these languages.

Other well-known classification is given by Matisoff’s (1978). Matisoff’s subgroups of TB languages are: Kamarupan, Qiangic, Himalayish, Kachinic, Lolo-Burmese and Karenic. Following subgroups of TB are now generally accepted (LaPolla 2006: 394):

- Qiangic: Qiang, Pumi, Muya, Namuyi, Shixing
- Lolo-Burmese:
 - Burmish languages – Burmese, Lawngwaw [Maru], Ngo Chang [Achang], Zaiwa, Lachik [Lashi]
 - Loloish Languages – Northern : Nosu [Yi, Yunnan or Sichuan], Nasu, Nisu;
 - Central: Lahu, Lisu, Nusu, Jinuo
 - Southern: Hani, Bisu, Phunoi, Mpi
- Bodish: Tibetan, Dzongkha, Tamang, Tashangla, Takpa
- Kuki-Chin: Lushai, Asho Chin, Tddim [Chin, Tidim], Anal, Hamar
- Bodo-Koch: Bodo, Garo, Dimasa, Kachari, Koch, Rabha
- Konyak: Tangsa [Naga, Tangsa], Chang, [Naga, Chang], Konyak [Naga, Konyak] Nocte [Naga, Nocte], Wancho [Naga, Wancho]
- Tani: [Apatani, Mising [Miri], Adi
- Karenic: Pwo [Karen, Pwo], Karenni, Sgaw [Karen, S’gaw]

Learners can refer the works of the writers given here for classifications of Tibeto-Burman languages:

Grierson 1909, Shafer 1955, Benedict 1972, DeLancey 1987, Sun 1988, Dai, Liu & Fu 1989, Bradley 1997, Matisoff 1996, 2003, and Thurgood 2003; Hale 1982 for comparison of the older proposals.

4.2.2.3 Languages of North East India

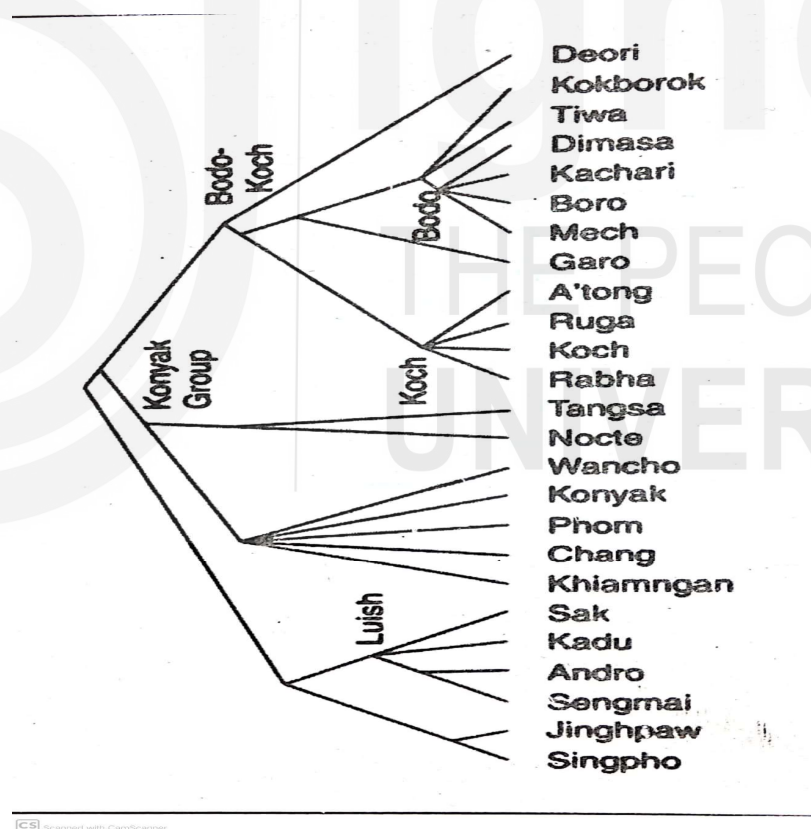
Affiliation of many of the languages of NE India is controversial. Burling (2003) has discussed fourteen

subgroups of TB languages spoken in NE India. But he also cautions not to assume these subgroups as co-ordinate branches of the family. He adds, 'it is surely unlikely that the North Eastern India harbours fourteen totally separate branches of Tibeto-Burman, and as we learn more about these languages, we should be able to merge some groups' (p.174). The groups are discussed under the broad geographic area representing the languages:

1. THE CENTRAL AREA

- i. **Bodo-Konyak-Jinghpaw:** most recognizable group based primarily in NE India. Given below (Figure 4.1) is the probable relationship of the languages of the group through a stambaum (family tree):

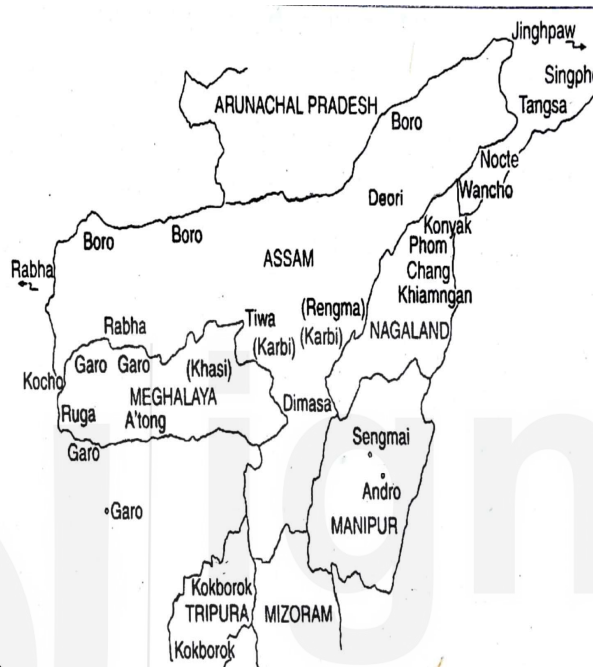
FIG: 4.1 RELATIONSHIP AMONG THE BODO-KONYAK-JINGHPAW LANGUAGES (SOURCE-BURLING: 2003: 175)



Bodo-Konyak-Jinghpaw has three main branches: Bodo-Koch [Bodo, Bodo-Garo, Barish]. It further has four parts: Bodo, Garo, Koch, Deori. Bodo includes several languages or dialects: Boro (or

Bodo), Mech, Dimasa, Kachari and Hill Kachari. Bodo-Konyak-Jinghpaw are spoken in (Map 4.1) Meghalaya, Nagaland, Meghalaya, Tripura and Assam.

Map 4.1 Bod-Konyak-Jinghpaw languages (Source-Burling 2003: 176)



It is assumed that the Bodo-Koch languages were spoken in much wider and continuous geographic area earlier but the gradual spread of the Indic languages has affected their presence in the area and as a result we find these languages separated in their representative areas. These languages are also assumed to be the dominant languages of the Assam and some parts of Northern Bengal in earlier periods.

Six languages have generally been recognized as belonging to the Konyak group: Tansa, Nocte [Namsangia], Wanchao [Banpara] (in southern most Arunachal Pradesh); Konyak [Tableng], Phom [Chingmengnu, Tamlu], and Chang in Northeastern Nagaland. Burling suggests Khamngan as the seventh member of this group. Jinghpaw has often been placed with Lolo-Burmese as it shares vocabulary with northern Burmish and with Burmish. It is used as a lingua franca in Kachin state of northern Myanmar.

2. THE NORTHERN AREA

Languages of Arunachal and Assam are mainly included in this group. (See Map 4.2)

MAP 4.2 LANGUAGES OF ARUNACHAL AND NEIGHBORING ASSAM, BURLING 2003: 179



ii. **Tshangla-Takpa**

Tshangla-Western Arunachal, near the border of Bhutan and Sharchopkha-Similar language spoken in southeastern Bhutan

Takpa [Dwags] - western tip of Arunachal, referred there as Northern Monpa, in Tibet-known as Cuona Menba, resembles Bumtang languages of Central Bhutan

iii. **Sherdukpen, Bungun/khoa, Sulung, Lishpa** – Arunachal Pradesh to the east of Tshangla

iv. **Hrusish** – Hruso [Aka] and Dhammai/Miji north of Bungun and Sherdukpen.

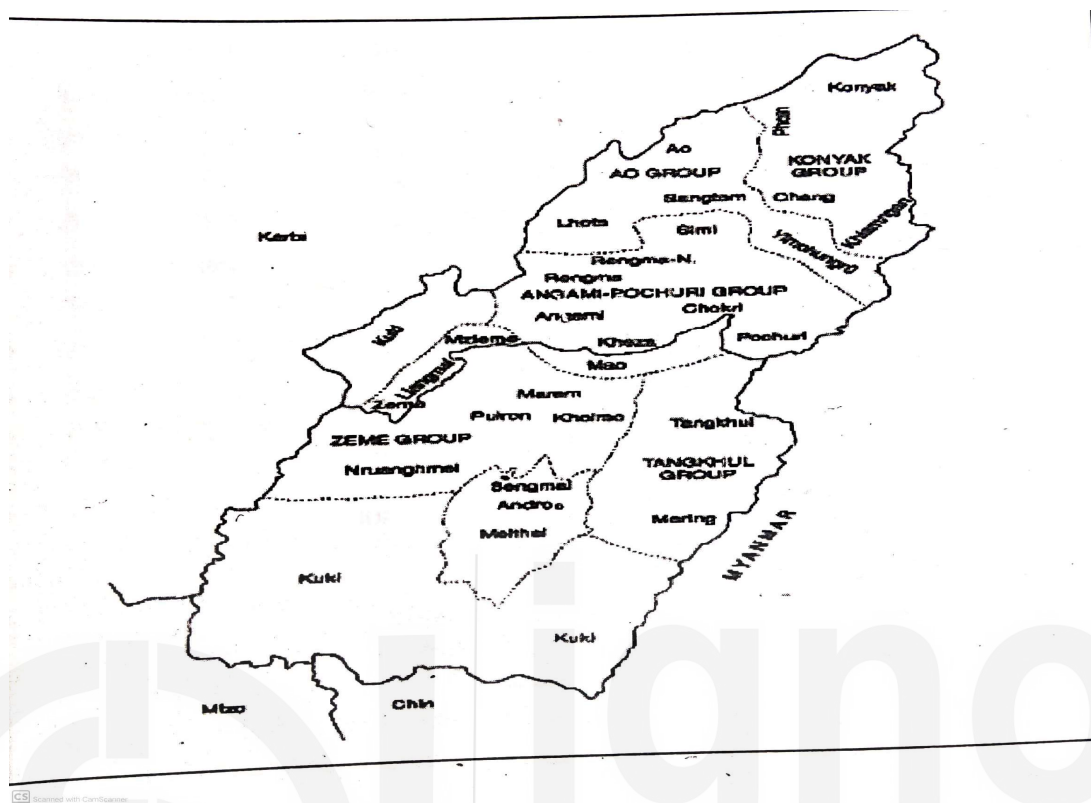
v. **Tani** [Mirish, Misingish, Abor-Miri-Dafila]- the term Tani (Sun 1993) is given to the languages of Central Arunachal Pradesh. Mising [Plains Miri]- in plains of upper Assam.

vi. **Idu-Digar**- Northeastern Arunachal

vii. **Miju**- Southeastern Arunachal Pradesh

3. **THE EASTERN BORDER AREA:** Linguistically most heterogeneous region of NE India. Location of this group (as shown in the Map 4.3) is ‘relatively accurate for Nagaland; Manipur has required much more guess work’ (Burling 184)

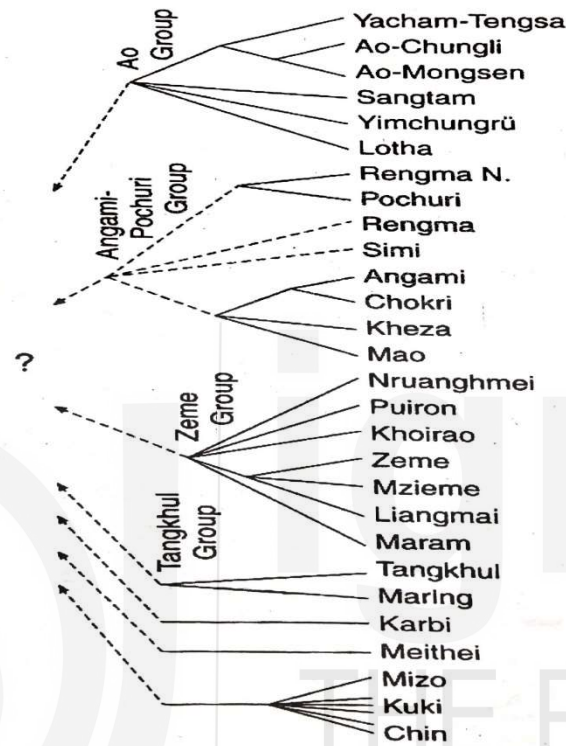
MAP 4.3 LANGUAGES OF THE EASTERN BORDER, BURLING, 2003: 185



- viii. **Ao Group**- Chungli, Mongsen, Yacham, Tengsa, Lotha, Sangtam [Tukumi], and Yimchungrü [Yachumi]
- ix. **Angami-Pochuri group**: Angami, Chokri, Kheza, Mao, Simi, Rengma, Rngma N., Pochuri
- x. **Zeme group**: Zeme [Empeo, Kachcha], Mzieme, Liangmai [Kwoireng], Nruanghmei [Rongmei, Kabui], Puiron, Khoirao, Maram
- xi. **Tangkhul group**: Tangkhul, Maring
- xii. **Karbi [Mikir]**
- xiii. **Meitei**- Meitei or Meiteilon (or Manipuri)
- xiv. **Mizo-Kuki-Chin**- Mizo [Lushai]- Majority language of Mizoram, Chin – Chin state of Myanmar, Kuki- Mizoram, Manipur, also in surrounding states

For the relationship of the languages of this group refer the Figure (4.2) below.

FIG 4.2 RELATIONSHIP AMONG THE LANGUAGES OF THE EASTERN BORDER, BURLING 2003: 184



Solid lines in the above figure show most reliable relationships, broken lines represent probable relationship and the dotted lines are speculative. The relative height of the branching nodes represents degree of differences between languages (Burling). In terms of internal diversity the groups can be arranged like Angami-Pochuri group > Ao > Mizo-Kuki-Chin.

We can end our discussion on the classification of the languages of the North-East India with the words of Burling (2003: 189):

“Much more work is needed before we will have a definitive classification of the languages of the eastern border. The similarities among these languages are tantalizing but difficult to pin down, and until more solid work is done, it seems best to withhold judgement. The individual groups seem sound but how they relate to each other and whether they fall into one or several branches of Tibeto-Burman awaits further data and further analysis.”

4.3 STURCTURAL FEATURES

Sino-Tibetan languages are monosyllabic, isolating and tonal languages. In isolating languages words do not change form with inflectional morphemes. Tibeto-Burman languages are less monosyllabic and isolating than the other Sino-Tibetan languages. Word order is important in expressing grammatical relationship in ST languages. The number of tones differs in each language: three tones in Burmese, four in Mandarin Chinese and nine in Cantonese Chinese. (The details given here are from LaPolla, 2006).

Writing about the structure of these languages, LaPolla (2006) writes, at the earliest reconstructable stage of the development (possibly eight thousand years ago), “the Proto-Sino-Tibetan was monosyllabic, but with much more complicated syllable structure than most of the modern languages. Matisoff (1991) reconstruct the syllable canon as *(PREF) (PREF) C_i (G) V (:) (C_f) (s). (C_i = initial consonant, G=glide, : = vowel length, C_f = final consonant, s = suffixal *-s); parentheses mark items that do not appear in the syllables). There was no relational morphology, but there was derivational morphology in the form of prefixes, suffixes, and voicing alterations of the initial consonants. Modern languages has moved much more towards bisyllabic or polysyllabic words, although they are often reduced again to sesquisyllables (syllable and a half) or monosyllabic forms, and tone systems have developed in Sinitic and many of TB languages (either through contact, through independent innovation or both). For example, in Sinitic the tones developed out of consonant suffixes and loss of initial voicing.” Chinese became more isolating around roughly 3500 years.

The clause in Proto Sino-Tibetan was verb focused, i.e, verb was the key element and noun phrases were optional. In Sinitic the clause is largely verb medial, as the verb has come to function as the divider between topical (preverbal) and non topical (postverbal) elements (there has clearly been a progressive change away from verb-final order over time).

We find diversity in ST languages: in phonology as well as in morphology. The Qiang language having 36 initial consonants, a complex system of consonant clusters in initial and final positions and no tones, while Lahu has only 24 consonant initials, a simple (C)V syllable structure, and phonemic inventories. Qiang is agglutinative whereas Lahu is isolating. Qiang has comlex affixal system of direction marking, and evidential marking on the verb and definite marking in the noun phrases, whereas Lahu has none of these features.

In Indospheric languages (TB languages of India, Nepal) we often find development of relative pronouns and correlative structures and also of retroflexive initial consonants. In the Sinosphere the development of

tone systems and more analytical structure can be observed. In Chinese contact influence can be observed from the Altaic languages in the north and Austro-asiatic, Tai-Kadai, Hmong-Mein languages in the South.

4.3.1 Structural Features of the Tibeto-Burman Languages

Given here are some structural features of TB languages discussed in Abbi (2001: 39-42):

- The use of tones at phonemic levels is found in all the languages of this family. For example, words from Tangkhul Naga - kəhún (rising tone) = red, kəhuŋ = rotten, kəhùŋ (falling tone) = lime, are differentiated only on the tone contour.
- These languages provide vowel clusters. Meithei- 'məun' means skin, 'ain'- rules, laws and 'məruoyobə' - important. Chelliah (1997) mentions 36 vowel sequences in Meithei.
- Most of the languages of this family are isolating in nature.
- Tense distinction is maintained between future and non future which is symbolized by an adverb of time. Verb indicates only aspect.
- Do not distinguish a separate word class for adjectives.
- Languages of this family SOV in nature. Notions of subject and predicate are alien to Sino-Tibetan. ST languages are topic prominent languages.
- All TB languages are OV, except for Bai and Karen languages, which are VO (more specifically SVO) and share other characteristics of word order of OV languages such as they employ postpositions rather than prepositions, genitive modifiers are placed before the possessed noun. (Dryer 2003: 43).
- Verbs have intricate system of evidential particles that characterize the nature of speaker's information. The pragmatic features such as hearsay, preconceived knowledge, first hand knowledge, visual vs. auditory information and other emotional attitudes are encoded in the verbal morphology.

4.4 Other Language Families of The World

So far, you have read about four language families which are represented in different parts of India: Indo Aryan, Austro-Asiatic, Dravidian and Tibeto-Burman.

Languages of one more family are spoken in India, namely Andamanese Languages. Andamanese

languages are spoken in the Andaman Island in the Bay of Bengal. Two distinct groups – the Little Andamanese (LA) and Great Andamanese (GA) speak different languages. LA group speak Onge, Jarawa, and Sentinelese; GA group speak ten different languages which are spoken by ten different tribes. We know very little about the languages of this family, notes Abbi (2001: 42), ‘Very little is known about the languages of this family. Not more than a few hundred speakers remain today. The Andamanese language family is considered oldest of the Indian language families and has been spoken by the Negritos in the area for last 3000 years.’

Indo Aryan language family is one of the largest group of Indo-European language family. Other families of languages related to Indo-European are: Albanic, Anatolian, Armenian, Balto-Slavic, Germanic, Greek, Italic and Celtic.

There are other language families represented in various parts of world other than Indo European. Some non-Indo European languages are shown in the table 4.1 below:

Table: 4.1 Some of the non-Indo-European languages of the World

Family	Language	Principal area where spoken	No. of speakers in million
Afro-Asiatic	Hausa	West Africa	23
	Amharic	East Africa	10
	Arabic	North Africa	155
	Hebrew	Israel	3
Altaic	Mongolian (Khalkha)	Mongolia	2
	Turkish	Turkey	45
Austro Asiatic	Vietnamese	Vietnam	45
Austronesian	Indonesian-Malay	Indonesia, Malaysia	115
Caucasian Dravidian	Georgian	Caucasus	3
	Kannada	India	32
	Malayalam	India	31
	Tamil	India, Sri Lanka	59
	Telugu	India	60
Finno-Ugric	Finnish	Finland	5
	Hungarian	Hungary	13

Japanese	Japanese	Japan	119
Korean	Korean	Korea	60
Niger-Congo	Swahili	East Africa	32
	Igbo	West Africa	12
	Yoruba	West Africa	14
Sino-Tibetan	Cantonese	Southern China	55
	Mandarin	Northern China	726
	Burmese	Myanmar (Burma)	26
	Tibetan	Tibet	6

Source: Akmajian, Demers, Farmer and Harnish (2006: 334)

We will briefly outline some language families of the world.

URALIC

The 30 languages of this family are spoken in the region of north Ural Mountains in Russia over 7000 years ago (Crystal 1996). Two main branches of the family are:

- i. **Finno-Ugric:** Central Europe and northern territories where Europe and Asia meet. Finnic Branch: Finnish (Finland, Sweden, and Russia; and through emigration in USA); Estonian

Ugric Branch: Hungarian (or Magyar) – in Hungary and other parts of the world through emigration; Khanty (or Ostyak) and Mansi (or Vogul) - east of Urals around the River Ob. Ingrian, Livonian and Votic- spoken within Russia (and have very few speakers) Karelin.
- ii **Samoyedic:** Siberia and Arctic Russia. Widely spoken language is Nenets (or Yurak). Selkup (or Ostyak Samoyed), Nganasan (Tavgi or Aram) Enets (only 100 speakers in 1989 and seems to have recently died out, Crystal, 1996).

CAUCASIAN

A small area between the Black Sea and the Caspian Sea surrounding the Caucasus Mountains ‘contains one of the highest concentrations of languages in the world’ (Crystal 1996: 307). 40 languages belong to Caucasian family of languages. These languages are classified into four regionally based types:

- i) North West Caucasian (or Abkhazo-Adyghian group)
- ii) North Central Caucasian group

- iii) North East Caucasian group
- iv) Southern Caucasian group (or Kartvelian)

Ubykh – one of language of this family was reported to have only one speaker in 1980s. The language has 80 consonants and it gives this language family a special place in phonological studies. Languages of this family are – Kartvelian (one of the dialects of Georgian), Zan, Svan, (in the south), Kabardian, Adyghe and Adkhaz, Ubykh (in the north – west); and in the north – east two groups of languages are spoken: Dagestanian group and Nakh group. The main languages of Dagestanian group are Avar, Lezghian (Kuri), Dargwa (Khjurkili), Lak and Tabassaran. The main languages of Nakh group: Chechen, Ingush and Bats.

PALAEOSIBERIAN

The groups of languages spoken by inhabitants of north eastern Siberia have been studied as Paleosiberian. ‘The groups are not genetically related to each other and do not constitute a family in the linguistic sense. Nor are the links with other families clear’ (Crystal 1996: 308). In the north – east Luoraweltan group consists of Chuckchi, Koryak (or Nymylan) Kamchadal (or Itelmen), Aliutor and Kerek. To the west, Yukaghir group with a single language – Yukaghir or Odul. Yeniseian group along the Yenisey River in the further west people speak Ket (or Yenisey-Ostyak). In the south Gilyak (or Nivkhi) people speak which has no other relatives.

ALTAIC

Altaic family covers a vast area- from the Balkan Peninsula to the north-east of Asia. The family has got its name from Altai Mountain. This family consists of three groups: Turkic, Mongolian and Tungusic. Among 60 languages of the family, over half belong to Turkic group. Turkic group includes Turkish (spoken in Turkey and surrounding territories), Azerbaijani and Turkmen in south west (Russia, Iran and Afghanistan), in the south east Uzbek (Uzbekistan), Uighur (China and nearby Russia), in the north- west Tatar, Kazakh, Kirghiz and Bashkir (Kazakhstan, Kirghizstan, and Russia), and in the north-east Yakut, Tuva and others with very less number of speakers.

Mongolian languages include Mongol (in Mongolian republic of China), Buryat, Dongzhang (or Santa), Dagur and Tu (Monguor) and in the west Oirat, Kalmyk and Mongol.

Tungusic – Evenki (formerly Tungus), Lamut (or Even), Nanai and Manchu

TAI

Tai languages are spoken in South East Asia, mainly in Thailand with other countries Laos, North Vietnam, China and Myanmar. Tai languages are divided into three groups: South Western, Central and northern. Main languages of this family are: Thai (official language of Thailand), Lao or Laotian (Thailand and Laos, official language of Laos), Isan (Thailand), Zhuang (China) Lanna or Yuan (Thailand), Shan, Tho, Buyi, Dong and Nung.

Languages of Africa

‘Africa contains more languages than any other continent – some 2,000 spoken by over 400 million people. The language total is uncertain, because many areas are inaccessible and many dialect groups have not been well investigated.’ (Crystal 1996: 316). Very few of these languages are spoken by large numbers. Africa is the continent of Lingua francas (Crystal): Arabic, Swahili, English, and French and some Pidgins and Creoles have been used to foster communication between various communities. African languages are classified in four groups: Niger-Congo, Nilo-Saharan, Khoisan and Afro-Asiatic.

- i) **Niger – Congo:** Largest African family with around 1,350 languages and several thousand dialects. This group is further divided into several sub-groups. Benue–Congo is the largest group (central and southern Africa), Kwa (southern part of West Africa), Adamawa-Ubangi group (remote, northern part of central Africa), Gur group, Atlantic, Mande and Kordofanian are some subgroups and each sub group contains various languages.
- ii) **Nilo-Saharan:** Also called Chari-Nile as the languages of this group are spoken in the upper parts of the Chari and Nile rivers. Some 180 languages are included in this group.
- iii) **Khoisan:** Smallest language family in Africa with less than 40 languages. The languages are well known for the use of click consonants.
- iv) **Afro-Asiatic:** Languages of this family, formerly known as Hamito-semitic, are spoken in North Africa and south-west Asia. This group is further divided in five (or six) sub groups:
Semitic languages – spoken in South west Asia and North Africa. Arabic is the major language of this group. Tigrinya, Amharic and Gurage are other well known languages of this group. Egyptian (now extinct), Berber, Cushitic, Chadic and Omotic branches are less wide spread.

Some Other Language Families and Language Isolates

Apart from the language families discussed, there are other language families: Austronesian, Indo-Pacific,

American-Indian languages and Australian Aboriginal languages. Some languages for which no relationship can be established with any other language are known as language isolates. A language can be classified as an isolate when no resemblance with other languages can be determined. But sometimes a particular language can be classified as isolate when little is known about the language both historically and linguistically and it may be that further research indicates some relationship to other languages. Iberian (southern and eastern Spain), Basque (south west Europe) and Burushaski (north west Kashmir) are some of the language isolates.

4.5 Let Us Sum Up

We have discussed Sino-Tibetan language family in this unit. Classification of ST language family is not agreed upon by scholars. There are different groups proposed by various linguists. But two generally accepted groups of languages are: Chinese and Tibeto-Burman. Tibeto-Burman languages are mainly represented in Northeastern India and adjacent areas. It is difficult for linguists to tell exactly how many languages of Tibeto-Burman group are spoken in Northeastern India. As some of the areas are inaccessible and same languages are known by different names by the speakers and the outsiders. Sometimes it is difficult to differentiate languages and dialects both in Chinese languages and in Tibeto-Burman languages. Chinese languages are called dialect groups and the connotation of the term dialect is not the same as in the European context. Most of the languages of ST family are monosyllabic, tonal and isolating. The languages are also classified as Indosphere and Sinosphere as the two groups show contact influence of the surrounding areas.

Other language family present in India is Andamanese. Very little is known about the family at present. There are many languages spoken all over the world. All the languages have been classified into diverse groups but there are certain languages for which no linguistic or historical relation has been established. These languages are called language isolates.

4.6 Exercise

1. Write the geographical spread of Chinese–Tibetan language family.
2. Why the exact classification of Chinese–Tibetan languages is difficult?

3. Name the two groups widely accepted under Chinese–Tibetan language family.
4. Why the term dialect group is used for Chinese languages instead of languages?
5. Name major dialect groups of Chinese language family.
6. Why it is difficult to count the languages of NE India?
7. What do you mean by the term language isolate?
8. Where are the languages of Andamanese language family spoken?
9. Write geographic spread of Tibeto-Burman language family.
10. Which group of languages is known for click consonants?
11. Write an essay on Sino-Tibetan language family.
12. Give a classification of Sinitic (Chinese) languages.
13. Give a classification of Tibeto-Burman languages.
14. Write an essay on the language families represented in India.
15. Write an essay on the languages of Northeast India.

4.7 Glossary

- **Chinese-Tibetan language family:** (also called Sino-Tibetan or Tibetan Chinese) is one of the largest language families in the world and consists of two sub-groups: Sinitic (or Chinese) and Tibeto-Burman.
- **Sinitic:** refers to the various Chinese languages. Chinese family of languages is a group of related dialects as these languages share a common script and a common literary and cultural history. Generally agreed number of dialect groups in Chinese is six.
- **Putonghua** (*P^ut – onghu \ a*) : literally means common language, the National standard language of China. It is based on the phonology of the dialect of Beijing and the lexicon and the grammar are based on northern vernacular.
- **Tibetan languages:** The languages of this family are spoken in Northeast India, Myanmar, Bangladesh and northern Thailand, Tibet, China, Korea and Taiwan. Tibetan group of languages of ST family is mainly represented in India (North-East India).
- **Proto Sino Tibetan:** linguists have reconstructed the Proto-Sino Tibetan from which Chinese and Tibeto- Burman (the forms we know today) have been assumed to be derived. Proto Sino Tibetan is

assumed to be spoken in Yellow river valley around 6000 years ago.

- **Andamanese:** Andamanese languages are spoken in the Andaman Island in the Bay of Bengal. Two distinct groups – the Little Andamanese (LA) and Great Andamanese (GA) speak different languages
- **Language Isolate:** The languages for which no resemblance with other languages can be determined are classified as isolate.

4.8 Suggested Readings and References

- Abbi, Anvita. (2001) *A Manual of Linguistic Field Work and Structures of Indian Languages*. Lincom Europa.
- Akmajian, A, Demers, RA, Farmer AK, & Harnish RM (2006). *An Introduction to Language and Communication*. New Delhi: Prentice Hall of India.
- Burling, R (2003). *The Tibeto-Burman Languages of Northeastern India*. In Thurgood & LaPolla (eds.).
- Chelliah, Shobhana, L. (1997). *A Grammar of Meithei*. Berlin, New York: Mouton de Gruyter.
- Crystal, David (1996). *The Cambridge Encyclopedia of Language*. Second Edition. Cambridge: CUP.
- Dryer, Matthew, S. (2003). Word Order in Sino-Tibetan Languages from a Typological and Geographical Perspective. In Thurgood & LaPolla (eds.)
- Grierson, G. A. (1927). *Linguistic Survey of India*. Vol.I. (Part-I)- Introductory . Reprint-1990: Delhi. DK Publishers & Distributors.
- LaPolla, R J (2001). The Role of Migration and Language contact in the development of the Sino-Tibetan Language Family. In Dixon R M & Aikhenvald A Y (eds.) *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*. Oxford: OUP. 225-254.
- LaPolla, R J (2003). *Overview of Sino-Tibetan Morphosyntax*. In Thurgood, G. & LaPolla, R. J. (Eds.). 2003. *The Sino-Tibetan Languages*, pp. 22-42. London and New York: Routledge. Paperback edition 2007.
- LaPolla, R J (2006) *Sino-Tibetan Languages*. v2.linguistlist.org/~lapolla/rjlapolla/papers/Sino-Tibetan_Languages.pdf.
- Li F –K (1936-1937). *Languages, Dialects*. The Chinese Year Book, 121-128. Reprinted 1973 in *Journal of Chinese Linguistics* 1 (1).

- Matisoff JA (2003). *In Defence of Kamrupan*. Linguistics of the Tibeto-Burman Area 22 (2).
- Matisoff JA (1996). *STEDT Monograph Series 2: Languages and Dialects of Tibeto-Burman*. Berkeley: Centre for South and Southeast Asian Languages.
- Matisoff JA (1991). Sino-Tibetan Linguistics: Present State and Future Prospects. Annual Review of Anthropology, 20.
- Matisoff, JA (1978). *Variational Semantics of Tibeto-Burman. The Organic Approach to Linguistics Comparison..* Philadelphia: Institute for the Study of Human Issues Publication.
- Norman, J. (2003). *The Chinese Dialects: Phonology*. In Thurgood & LaPolla (eds.).
- Shafer R. (1955). *Classification of the Sino-Tibetan languages*. Word 11 (1).
- Thurgood, G. (2003). *A subgrouping of the Sino-Tibetan languages: the interaction between language contact, change and inheritance*, in Thurgood, G. & LaPolla, R. J. (Eds.). 2003. *The Sino-Tibetan Languages*, pp. 1-21. London and New York: Routledge. Paperback edition, 2007.
- Thurgood, G. & LaPolla, R. J. (Eds.) (2003). *The Sino-Tibetan Languages*. London and New York: Routledge. Paperback edition 2007.
- Verma, S. (1972). *G. A. Grierson's Linguistic Survey of India: A Summary*. Hoshiyarpur (Punjab): Vishveshvaranand Institute.

Web Resources

11. Description of the Sino-Tibetan Language Family. <http://stedt.berkeley.edu/html/STfamily.html>
12. LaPolla, R J (2006) *Sino-Tibetan Languages*. v2.linguistlist.org/~lapolla/rjlapolla/papers/Sino-Tibetan_Languages.pdf.