

# **LINGUISTIC SURVEY OF NEPAL**

**(LinSuN)**

---

**A Proposal**

**Prepared on behalf of**

**National Planning Commission**

**Government of Nepal**

---

**Drafting Team**

**LinSuN**

**Central Department of Linguistics**

**Tribhuvan University**

**June 27, 2008**

<b>EXECUTIVE SUMMARY (1 PAGER)</b> .....	<b>3</b>
<b>EXECUTIVE SUMMARY (4 PAGER)</b> .....	<b>3</b>
<b>1 BACKGROUND AND RATIONALE</b> .....	<b>8</b>
<b>2 CULTURAL AND LINGUISTIC DIVERSITY IN NEPAL</b> .....	<b>9</b>
2.1 VARIOUS ESTIMATES .....	10
2.2 GENETIC AFFILIATION .....	11
2.3 WRITING SYSTEMS .....	11
2.4 LANGUAGE AND ETHNICITY .....	13
2.5 SECOND LANGUAGES AND LINGUA FRANCAS .....	14
2.6 EXISTING LEGAL PROVISIONS .....	15
<b>3 EARLIER LINGUISTIC SURVEYS</b> .....	<b>17</b>
3.1 HODGSON'S SURVEYS .....	17
3.2 CAREY'S LINGUISTIC SURVEY OF INDIA .....	17
3.3 GRIERSON'S LINGUISTIC SURVEY OF INDIA (1898-1927) .....	17
3.3.1 <i>Grierson's Methodology</i> .....	18
3.4 CONTRIBUTIONS OF SUMMER INSTITUTE OF LINGUISTICS (SIL) .....	20
3.5 POKHAREL AND CHAUDHARI'S SURVEY .....	20
3.6 POKHAREL AND CHAUDHARI'S RESEARCH TOOL .....	20
3.6.1 <i>B K Pokharel's Research Tool</i> .....	20
3.7 BANDHU'S FIELD SURVEY AND RNA'S SURVEY DESIGN .....	20
3.8 RAHUL SANKRITYAYAN'S SURVEY .....	21
3.9 GLOVER'S SURVEY OF GURUNG DIALECTS .....	21
3.10 WINTER'S LINGUISTIC SURVEY OF NEPAL (1981-84) .....	21
3.10.1 <i>Winter's Survey Tool</i> .....	22
3.11 MICHAILOVSKY'S TYPOLOGICAL SURVEY OF NEPALESE LANGUAGES (1988) .....	22
3.12 M POKHAREL'S SURVEY OF NEPALI DIALECTS .....	22
3.13 AUSTIN HALE'S CNAS PROPOSAL FOR SOCIOLINGUISTIC SURVEY .....	22
3.14 RESEARCH AND LANGUAGE DOCUMENTATION PROGRAMS AT CDL (1999-) .....	22
3.15 LINSUN PROJECT .....	23
3.16 TOBA ET AL'S SURVEYS .....	23
3.16.1 <i>Toba et al (2002)'s Survey Tool</i> .....	24
3.17 BIELMEIER'S SURVEY OF TIBETAN DIALECTS .....	24
3.18 MINOR SURVEYS .....	24
<b>4 OBJECTIVES</b> .....	<b>24</b>
<b>5 METHODOLOGY</b> .....	<b>25</b>
5.1 A SOCIOLINGUISTIC DESCRIPTION OF ALL THE LANGUAGES OF NEPAL (OBJECTIVE 1) .....	25
5.1.1 <i>Lexical similarity</i> .....	25
5.1.2 <i>Comprehension of Recorded Texts</i> .....	26
5.1.3 <i>Levels of Competence in the National Language</i> .....	26
5.1.4 <i>Sociolinguistic Questionnaires</i> .....	26
5.2 BASIC LINGUISTIC DESCRIPTION: GRAMMAR SKETCHES (OBJECTIVE 2) .....	27
5.2.1 <i>Linguistic Description: Lexicon</i> .....	29
<b>5.3 DEVELOP AND MAINTAIN A COMPLETE DATABASE OF THE LANGUAGES OF NEPAL (OBJECTIVE 3)</b> .....	<b>29</b>
5.3.1 <i>Database: Written corpora</i> .....	29
5.3.2 <i>Database: Spoken corpora</i> .....	29
5.4 NEPALI SIGN LANGUAGE .....	30
5.5 USE OF MOTHER TONGUES IN EDUCATION AND LITERACY IN NEPAL .....	30
<b>6 MANAGEMENT</b> .....	<b>31</b>
<b>7 WORK PLAN</b> .....	<b>33</b>
7.1 INTRODUCTION .....	33
7.2 TIME FRAME .....	35
7.3 MAN POWER .....	35
7.4 OVERVIEW OF RESEARCH ACTIVITIES .....	36

	2
7.5 PHASE ONE (YEAR 1 AND 2).....	38
7.5.1 Sociolinguistic Component .....	38
7.5.2 Documentation Component .....	38
7.5.3 Literacy Component.....	39
7.6 PHASE 2 (YEARS 3, 4, 5, 6, AND 7).....	41
7.6.1 Sociolinguistics .....	41
7.6.2 Documentation.....	41
7.6.3 Literacy .....	42
7.7 INCOMPLETE DOCUMENTATION .....	44
7.8 ARCHIVING COMPONENT .....	44
7.9 FINAL RECOMMENDATIONS.....	45
7.10 SUMMARY OF WORK PLAN.....	45
<b>8 HUMAN RESOURCES AND TRAINING REQUIREMENTS .....</b>	<b>46</b>
8.1 A PHASE BY PHASE REQUIREMENT .....	46
8.2 SPECIAL TRAINING PROGRAMMES.....	46
8.3 ACADEMIC TRAINING PROGRAMMES .....	47
8.4 ANALYTICAL AND TECHNICAL MANPOWER .....	47
8.5 TRAINING ALLOCATIONS .....	48
<b>9 EXPECTED RESULTS .....</b>	<b>49</b>
<b>10 ETHICAL ISSUES.....</b>	<b>50</b>
10.1 THE FIELDWORK.....	50
10.2 ARCHIVING.....	51
10.3 PUBLICATION .....	51
<b>11 REPORTING REQUIREMENTS .....</b>	<b>51</b>
<b>12 INSTITUTIONAL PROFILE .....</b>	<b>51</b>
<b>13 ADDITIONAL SOURCES OF FUNDING .....</b>	<b>53</b>
<b>14 BUDGET .....</b>	<b>55</b>
<b>REFERENCES.....</b>	<b>63</b>

## EXECUTIVE SUMMARY

(one pager)

### **Vision**

The vision of the Linguistic Survey of Nepal (LinSuN) is to lay a foundation that provides for the linguistic rights of the citizens of Nepal so that all her people, regardless of linguistic background, will be included in the overall fabric of the nation.

### **Rationale**

While the Interim Constitution of Nepal (2007) grants numerous linguistic rights to the citizens of Nepal, there is not sufficient understanding in the diversity of its people and the languages they speak. Even a full identification of the number of languages and dialects has not yet been possible. If efforts in linguistic inclusion will have any success, they must begin first with an understanding of the full extent of the linguistic and ethnic diversity of the country.

### **Objectives**

- Produce a sociolinguistic profile for each of the languages of Nepal that includes a mapping of languages and dialects.
- Produce a basic description of at least ten languages that includes a working orthography, observations on the grammar, a trilingual glossary, and publications of folk tales.
- Develop and maintain a complete database of the languages of Nepal which will include sociolinguistic information, grammatical information, oral and written texts, and vocabulary.
- Develop a description of the use of mother tongues in education (formal and non-formal) as a means to better understand the development needed for mother-tongue curricula in the national educational system.
- High quality written and electronic publications for each of its major research activities.

### **Ownership**

The National Planning Commission, Government of Nepal, has commissioned the Central Department of Linguistics, Tribhuvan University, Kathmandu to conduct this survey.

### **Workforce**

The required work force will be approximately 150 members, some 120 of which will be involved in research tasks. The remaining work force is part of the administrative structure to make the project run.

### **Time Frame and Budget**

The survey will begin in August 2008 and run for a span of seven years.

The overall budget for the seven year span is NRS 44,58,49,200 (forty-four crore, fifty eight lakh, forty nine thousand, two hundred, only). This is approximately equivalent to seven million U.S. Dollars or four million Euros.

## **Executive Summary** **(four pager)**

### **Rationale**

The Interim Constitution of Nepal (2007) grants numerous linguistic rights to the citizens of Nepal, and recognizes the linguistic diversity by recording 92 languages in the Census Report of 2001. There are however varying reports on the number of languages spoken in Nepal, and to clarify the linguistic situation in the country the National Languages Policy Recommendation Commission (1994/2050 V.S.; its English version 2008) had recommended a Linguistic Survey of Nepal to be carried out to determine more precise identification of languages spoken in the country. To date there have been two attempts to survey the languages of Nepal, namely the Grierson-Konow Survey (1898-1927) and the German Research Council Survey (1981-84). These however lack the detailed and comprehensive nature needed to address the current issues facing Nepal. The first Survey did not include all the languages spoken by the indigenous nationalities, and the second remains tentative and inconclusive due to inadequate documentation. There are thus several reasons why a more reliable survey is needed namely, (1) to develop orthographies for unwritten languages, (2) to determine the role of language in primary and adult education, (3) to identify and document minority languages facing extinction, and (4) to implement the socially inclusive provisions made in the Interim Plan, National Planning Commission 2007.

### **Objectives**

The objectives of the Linguistic Survey of Nepal are to :

1. Develop a sociolinguistic profile of all the languages of Nepal.
2. Produce a basic description of at least ten languages (at least one description in each cluster: see section 7 for a clarification of the term cluster) that includes an understanding of the sound system, observations on the grammar, and a trilingual glossary.
3. Develop and maintain a complete database of the languages of Nepal.
4. Develop a description of the use of mother tongues in education (formal and non-formal) as a means to better understand the development needed for mother-tongue curricula in the national educational system.

### **Methodology**

To accomplish the specific goals set out for the LinSuN, the following methods and techniques will be adopted:

- A Sociolinguistic description will involve (1) mapping the geographical location of languages, (2) a study of census reports for a demographic description of speakers of languages, (3) word lists and Recorded Text Tests (RTTS) to determine mutual intelligibility, (4) Sentence Repetition Tests (SRTs) to study multilingualism, (5) questionnaires and observations to record language use and attitudes, and (6) questionnaires to develop language resources for literacy and education.

- Linguistic description of individual languages will involve (1) recording, transcribing, and interlinearizing spoken and written texts, (2) writing grammar sketches based on particular framework and content outline, (3) elicitation of data based on native speakers' intuition of the language about grammaticality and acceptability etc.. Interlinearized Corpora will serve as one of the primary sources for writing grammar sketches.
- Sign language as a non-verbal language will be archived through videography with appropriate methods of linguistic annotations and tagging of Parts of Speech. The users of sign language follow a unique methodology and use a uniform system of signs and gestures to convey the meanings of vocabulary and morphology.
- Compiling a Basic Glossary in the form of a tri-lingual: native language – Nepali – English. In collecting words we intend to use both the corpora and classified lexicon known as Ontology. Each Dictionary will have entries of at least 5000 headwords arranged in a standard format. For Nepali and English-speaking audience Nepali-native and English-native indices are produced, respectively.
- Developing orthographies and publishing indigenous stories as a means of laying a foundation for future work in language development.
- Archiving materials: For archiving purposes, video recordings, digital sound, images, graphic representation and metadata will be used. The use of metadata is very essential for data accessibility. All these types of materials including audio / video recordings will be deposited at the Central Department of Linguistics Library (the host institution of this Project).

### **Management**

There shall be a **Steering Committee** at the National Planning Commission, Government of Nepal. It will have a supervisory and evaluative role and provide general guidance to the Survey. The National Planning Commission member in charge of Social Inclusion shall be the Coordinator of the Steering Committee. In addition, the National Planning Commission shall invite senior linguists and dignitaries from related government agencies to join this Committee.

For conducting the Survey there shall be a **Survey Team**. It will be assisted by an advisory group of distinguished linguists from Nepal and abroad.

### **Human Resources and Training Requirements**

The success of the LinSuN Survey will depend largely on the well-motivated and fully committed human resources of both the Survey team and the Administrative team. In assessing the manpower requirements for the Survey, provisions have been made for Special Training and Academic Training programmes. At least three different areas of training will be necessary for successful survey work under LinSuN :

- Training and upgradation of skills relevant to the Survey work through regular workshops and task-oriented meetings.
- Development and Specializations in Field Studies in the areas of Sociolinguistic Surveys and language descriptions.
- Orientations to impart skills in data gathering in basic descriptive linguistics, grammar writing, and training in making Areal Maps and Linguistic Atlases.

There will also be several Analytical Teams who will function as specialists of various aspects of language related to different levels of linguistic analyses. These would include the expert services of :

- The language community itself on whose knowledge the whole project depends.
- Anthropologists, Archivists, Phoneticians, Phonologists, Syntacticians, Lexicographers, Sociolinguists etc.
- Key Language Experts of Language Families or particular languages
- Specialists such as Statisticians and Management experts to provide reliable statistics of speakers, language data, and efficient management of the Survey.
- Professionals with technical background, such as Software Developers and experts from the Information Sciences to advise and improve the analytical tools and computational analyses of data, and also to solve certain unforeseen problems that may hamper the Survey work.

### **Expected Outcomes**

The primary outcome of LinSuN is to lay a foundation that provides for the linguistic rights of the citizens of Nepal so that all her people, regardless of linguistic background, will be included in the overall fabric of the nation. By better understanding the linguistic and ethnic diversity of the Nepalese people, it is hoped that no community will be excluded from the activities of the nation because of language, and that each language community will be given a foundation for long term language development.

On the completion of the Survey the following results can be expected:

1. This survey will contribute to a comprehensive database with information about each language of Nepal. Each language will have at least a sociolinguistic profile in the database. Many languages, however, will have a fuller picture with an annotation of previous research, grammatical information, and a glossary size lexicon.
2. Policy makers at the government and political levels, linguists, as well as language communities will have access to this database to help them plan and formulate language policies. It is hoped that this will help interested parties make informed decision about including marginalized languages in nation building efforts.
3. The Survey will foster discussion about what it takes to take a spoken-only language to a standardized written language that can be used in written and other more formalized domains.
4. Through its activities in documentation in a limited set of languages, previously spoken-only languages will have a written form that can be used in limited ways in education, media and local administration. It is believed that extending its use in this way will help language communities keep their languages.
5. The host institution and other academic institutions will benefit from the findings of this study as they incorporate the information from this research into their curricula in teaching linguistics and languages.

6. This survey will help GOs and INGOs who are interested in mother tongue education through the information it gleans about orthographies and materials that have been developed in the languages of Nepal, as well as better understanding what remains to be done.
7. In the initial phase of the Survey, a series of trainings will be conducted to orient the researchers in various methods and techniques (including language technology) required for field work and analysis of data. This will develop capacity and expertise for further linguistic studies in Nepal.
8. A series of high quality publications on the languages of Nepal will be made available to the language communities and other interested parties. This material will be published electronically as well as in book form.

### **Time Frame and Budget**

The proposed budget for this 7-year Survey is NRS 44,58,49,200 (forty-four crore, fifty eight lakh, forty nine thousand, two hundred, only). The Survey will include Sociolinguistic Surveys, Language Documentations, Literacy, and Archiving.

## LINGUISTIC SURVEY OF NEPAL

**1 Background and Rationale**

The greater Himalayan region, which extends for 3,500 km from Afghanistan in the west to Myanmar in the east, sustains over 150 million people and is home to great linguistic diversity and many of Asia's most endangered languages. The great biological diversity of present-day Nepal is matched by its cultural and linguistic diversity. Comprising an area of 147,181 square kilometres with a length of 885 kilometres from east to west and a mean breadth of 193 kilometres from north to south, the topography of Nepal is rich and varied. According to recent census data collected in 2001, Nepal's 92 languages belong to four language families, an impressively large number for a country with a small land mass like Nepal. The Indo-Aryan group of the Indo-European language family forms the largest group in terms of speaker numbers, around 80% (Yadava 2003: 141). The Tibeto-Burman group within the Sino-Tibetan family of languages is represented by 57 languages in Nepal (CBS 2001), the largest number of distinct mother tongues of any linguistic grouping, but with noticeably less speakers than the Indo-Aryan group. Two other language families are also found in Nepal: the Munda branch of the Austro-Asiatic family and the Dravidian family, each represented by a couple of languages along the southern belt of the country. Moreover Kusunda, previously thought to be extinct, is a linguistic isolate spoken in Nepal. As is clear from the facts outlined above, Nepal is not only home to more language families than all of Europe combined, but also has more distinct and individual languages in one country than the whole of the European community. While the Census conducted by the CBS in 2001 has established fairly credible numbers of speakers for each of Nepal's languages, more precise and accurate figures still need to be ascertained through further careful investigation (Yadava and Turin 2007).

It is clear, then, that Nepal is impressively diverse in its linguistic makeup. The diversity is known – known to exist – but the way in which it is diverse, the extent to which it is diverse, is not fully known.

Language relationships in the Himalayan region are exceedingly complex, and there is little agreement about what the higher-order taxonomic relationships really are. Is Newar Bodish or is it Himalayish? Where does Chepang fall? Is Ghale part of the Tamangic complex or not? People have come and gone, kingdoms have risen and fallen, and influences have ebbed and flowed. Tibeto-Burman, Indo-Aryan, and Munda languages (among others) have been spoken within close proximity to one another for many hundreds, if not thousands, of years, and each has exerted lexical and grammatical influence on their neighbors. In some cases, what may have originally been genetically distinct languages have, over time, come to be typologically similar.

Fortunately, genetic relationships at lower levels are much more clear, or at least we have a much greater hope of sorting them out. One of the reasons linguists have been unable to establish higher-order relationships is that there is a general dearth of reliable materials to consult at the lower levels. Still, much progress has been made in recent years, and this is one of the areas where we can make significant contributions. We know now, for example, a good deal about the Kiranti languages and what makes them distinct from, say, the Tamang-Gurung-Thakali languages. We also know what makes the Kham Magar cluster of languages distinct from Standard Magar and why an education program for one would not work for the other. A primary goal of a linguistic survey, then, would be to identify, where possible, some of these basic characteristics — a language's so-called "typological profile".

At one level, a definition of the typological features that bring cohesiveness to a given cluster is at the forefront — i.e. "What makes Kiranti Kiranti, or what makes Tharu Tharu?" At a much finer, more precise level, a definition of the kinds of factors that disrupt

intelligibility between languages of the same cluster (thereby motivating some to treat them as separate languages) is at the forefront. Organized properly, then, a linguistic survey will provide answers to double edged questions like, “In what sense is Language X its own language, and in what sense is it a ‘dialect’ of Language Y?” Linguistic planners need to understand such issues before they attempt to implement policy. A well crafted survey can help provide the answers.

The lack of knowledge and understanding has significant implications for the country of Nepal. One of the implications is that there is inadequate information to address basic issues of language planning – language planning being an important part of a healthy process of inclusion in the nation state. Many within the Nepalese context are calling for the rights of ethnic and linguistic communities, but those in government often find it difficult to act even though the Constitution guarantees the elimination of language based discrimination and the right to mother tongue education for everyone. Policy makers simply do not know how many languages are spoken within Nepal’s boundaries, what languages are spoken, and how the speakers of these languages interrelate with one another and the broader national community, and how distinct the cultures and languages within a broad ethnic group are. In recognition of this uncertainty, the report of National Language Policy Recommendation Commission has made as its first recommendation that a linguistic survey of Nepal be carried out for a more precise identification of languages spoken in Nepal.

In addition to a more precise identification of languages, there are a number of other reasons for which a survey is needed. For example, it is unclear what role languages in Nepal play in communities, especially in primary and adult education, and what factors could contribute to the standardization of the languages of the nation. Because Nepali is spoken so widely within the country, it is particularly important that a clear understanding of its role in these domains be developed. (It should be noted here that the term "Nepali" may lead to confusion, and thus some clarification is in order. The term "Nepali" may be understood to mean "of or about" Nepal (thus, Nepali currency, Nepali culture, etc.). At the same time, "Nepali" (previously known as Ghorkali or Khas) is the name currently used for the dominant language of Nepal. In this document we use the term "Nepali" to refer to this language, but not to indicate that it is "the language of Nepal". It is one of the languages of Nepal.)

It is also important to conduct a complete linguistic survey of Nepal in order to establish orthographies for the recording of the languages of the nation. For viable but unwritten languages this would be an important step forward in their development of literacy and would also contribute towards development of mother-tongue curricula. This information will be valuable for the production of literacy materials in the language communities of Nepal, and in the production of bridge materials and further development of literacy in Nepali.

An accurate and authentic identification of Nepalese languages and their dialects is also needed to implement the socially inclusive provisions made in the Interim Constitution of Nepal (2007) such as equal status to all mother tongues and people's fundamental rights to preserve and promote them through their uses in primary education, media, and local administration. Similar provisions are also made in the 6th strategy for inclusive development (Interim Plan, National Planning Commission, 2007). This is in keeping with the international community’s concerns about the loss of diversity in the world's languages and cultures.

It is with this background in mind, then, that the Linguistic Survey of Nepal (LinSuN) is proposed.

## **2 Cultural and linguistic diversity in Nepal**

Nepal is rich in its cultural and linguistic diversity. The Central Bureau of Statistics (CBS) states that the Nepalese people belong to 103 caste and ethnic groups (CBS 2001) who

are largely Hindus, Buddhists, Kiratis, Animists, and Muslims. As a result, Nepal has evolved as a unique cultural space with numerous religious and philosophical values, customs and practices (Pradhan and Shreshtha 2006).

### 2.1 *Various estimates*

The great cultural, ecological and biological diversity of Nepal is matched by its linguistic diversity. Inhabiting these different places are diverse groups of people who speak a large number of different languages. There have been made some attempts to identify the numerous languages spoken in this country as mother tongues. One of these attempts is the enumeration of languages in the various censuses. Since the 1952/54 census languages have consistently been reported. However, their number shows variation in these censuses. Except 'other', 'unknown' and 'not stated' languages, there were recorded 44 (1952/54), 36 (1961), 17 (1971), 18 (1981), and 31 (1991) languages in the last five censuses (Gurung, 2002:37) but their figures have drastically increased to 92 in 2001. However, there still remain quite a few languages (used by as many as 168,340 strong speakers, i.e. 0.74 per cent of the total population) which are lumped together into an 'unknown' slot in the lack of adequate information. This increase is due to the fact that a large number of languages used as mother tongues have been returned for the first time because of the rising consciousness of several ethnic minorities about their distinct cultural and linguistic identity. Such mother tongues include Baram, Bhujel, Chhantyal, Dura, Ghale, Kaike, Kisan (also called Sadari or Sadani), Kusunda, Munda, Raute, Angika, Yholmo, Khariya, Lhomi, Dungmali, and Sadhani. Nepali Sign Language has been reported for the first time in the census reports. Besides, the earlier censuses recorded all the Rai languages under a single heading, called 'Rai group of languages'. In the 2001 census, however, there have been enumerated 22 separate languages in this group. Similarly, Angika and Bajjika, which were earlier considered the dialects of Maithili and Bhojpuri, respectively, have now been returned as distinct languages. In addition, there have also been reported some foreign languages (mostly spoken in India) for the first time; they are Hariyanwi, Magahi, Kuki, Mizo, Nagamese, Sindhi, Assamese, Oriya, Kurmali, and Koche, apart from English, Dzonkha, and Chinese.

This significant increase in the number of languages spoken in Nepal may be ascribed to a number of reasons. Since the restoration of democracy there has been continual increase in consciousness among linguistic minorities (including indigenous peoples) about their mother tongues. Their ethnic organizations had been playing an active role in creating awareness about preserving and promoting their cultural identity including their languages. Taking cognizance of this reality, the CBS also sought the cooperation and support of these organizations during the enumeration for the 2001 census. Following the enumeration, some linguists were also consulted for the precise identification of Nepal's languages (see Annex E for the enumeration of Nepal's languages in 2001 census).

The uncertainty about the number of Nepal's languages and their reduced enumeration in the last five censuses may be attributed to their lack of required information and awareness about mother tongues and also the "one nation – one language" policy adopted during the Panchayat and earlier regimes.

Setting aside the various censuses, attempts have also been made by some linguists to identify Nepal's languages by linguists. Malla (1989) and Toba (1992) mention 70 while Ethnologue (2005) lists 126 languages, one of which, called Kusunda, has been wrongly claimed to be dead. Quite recently, Noonan (2005:2) has presented an estimate of at least 140 languages spoken in Nepal, 109 of which are Sino-Tibetan. Of them at least 115 are spoken exclusively or primarily in Nepal while quite a few (including Nepali, the existing official language) are cross-border languages spoken in India and elsewhere as well.

Of the more than 92 (as enumerated in Census 2001), 16 languages are spoken by 97 per cent of the total population of Nepal while more than 75 languages are spoken by about 3 per cent, of which half of the languages are spoken by populations under 1000 and another half by less than 10,000 (Noonan 2005).

All these attempts (including censuses) for identification of languages spoken in Nepal as mother tongues are, however, just some sort of approximation and cannot be deemed final.

## 2.2 Genetic affiliation

The languages enumerated in the 2001 census except Kusunda belong to the four language families, viz. Indo-European, Sino-Tibetan, Austro-Asiatic, and Dravidian. Kusunda is a language isolate and its genetic affiliation is yet to be determined.

In Nepalese context, Indo-European family of languages mainly comprise Indo-Aryan group of languages, which forms the largest group of languages in terms of speakers, viz. nearly 80 per cent of the total population of the country. The genetic affiliations of some of the Indo-Aryan languages spoken in Nepal are yet to be identified in the lack of their description. These languages include Tharu, Bote, Majhi, Kumal, Darai, and Danuwar.

Another important group of Nepal's languages is the Tibeto-Burman group of languages. Though it is spoken by relatively lesser number of people than the Indo-European family, it consists of the largest number of languages. Contrary to speakers of Indo-Aryan languages, there has been steady decline in speakers of Sino-Tibetan languages from the 1952/54 to the 1981 censuses, viz. 21.8 per cent (1952/54), 19.26 per cent (1961), 17.16 per cent (1971), and 12.06 per cent (1981). In the last two censuses, they have, however, increased to 16.76 per cent (1991) and nearly 19 per cent (2001). Their decline and increase may also be due to the reason ascribed to those of Indo-European languages.

In addition to these two major language families, Nepal also has small numbers of speakers of two other language families. They are Munda branch of the Austro-Asiatic family and Dravidian family of languages. The Munda languages comprise Santhali of the northern Munda group and Khariya of the southern Munda group. According to the 2001 census, Santhali speakers are 40,193 in number, i.e. 0.18% of Nepal's total population, as compared to 0.20% (1952/54), 0.31% (1961), 0.21% (1971), 0.19% (1981), and 0.18% (1991). Another Austric language of Munda branch is Khariya, which has been introduced in the 2001 census for the first time. This language is spoken by 1575, i.e. 0.01%. All the Munda languages are spoken by groups of tribal peoples from the eastern Terai and make up approximately 0.19 per cent of the total population.

A single language of Dravidian family spoken in Nepal is called Uraon, also called Jhangar in the region east of Kosi river and Dhangar in the region west of Kosi river. It constitutes the northernmost part of Dravidian family of languages. It is said to be a regional variant of Kurux spoken in Jharkhand State of India though it shows divergence in its vocabulary and grammar (Gordon, 1976; Yadava, 2002). According to the 2001 census, it is spoken by 28,615, i.e. 0.13% of the total population of the country. Its speakers have been reported to be 4832 (1952/54), 9140 (1961), and 15175 (1991). But it was not listed in the 1971 and 1981 censuses.

## 2.3 Writing systems

Most of the languages spoken in Nepal are still confined to their oral traditions. Each of them has a rich oral heritage of traditional folk stories and songs handed down from parents to children over a long period of time, e.g. *Salhes* in Maithili and *Mundhun* in Kiranti languages. However, they are disappearing with the growth of literacy and language shift. It is, therefore, time to document them before they are lost to posterity.

Only a few of Nepal's languages have well-developed literate traditions. They include Nepali, Maithili, Tibetan/Sherpa, Newar, Limbu, Bhojpuri, Awadhi, and Lepcha in particular. All of them have long tradition of written literature. These languages have employed various writing systems or scripts. Most of the Indo-Aryan languages such as Nepali, Maithili (originally written in Mithilakshar or Kaithi script), Bhojpuri, Awadhi, Hindi and Rajbanshi are written in the modern Devanagri script with small adjustments and additional letters. The modern Devanagri script has evolved from an ancient writing system, called Brahmi dating back to about 500 BC. though Bengali is written in a modern Bangla script, which developed from the Kutila variety of the northern type of the Brahmi script. The proto-Bangla script degenerated into Maithili Tirhuta script or Mithilakshar and Kaithi with small adjustments and additional characters.

The Sambota script or Tibetan script, emerged from another proto-type of the Brahmi script, is used for Tibetan and Sherpa. This script has also been adopted for Tamang, while Tamang speakers also prefer the use of Devanagri or a simplified version of Sambota, called *Tamhig*. Newar has its own traditional script called Ranjana (also developed from the proto-Bangla script) but it has also adopted the Devanagri script for the sake of convenience in reading and printing. Limbu uses its own Kiranti Sirijanga script. Lepcha is written in Rong script. Another script Perso-Arabic is used in writing Urdu language while Punjabi language is written in Gurumukhi.

Of late some other languages have taken to literate traditions. Initiatives have been taken by various language communities to develop writing systems appropriate to the sound system of their languages and practically acceptable to them. They are Tharu, Tamang, Magar, Gurung, Rajbanshi and Rai group of languages such as Bantawa, Thulung, Chamling, Khaling, Kulung, and so on. Tharu, Tamang and Gurung use the Devanagri script but some of the Gurung speakers advocate for the use of Latin and Khema scripts for this language. Magar has developed its own script, called Akkha. Recently, these languages have started developing some written literature in the form of newspaper, magazine, textbooks for adult literacy and primary education, and folk literature.

As in India, Santhali in Nepal is written in the Santali alphabet, which is also known as Ol Cemet', Ol Ciki or simply Ol, which was created in the 1920s by Pandit Raghunath Murmu as part of his efforts to promote Santali culture. Until the invention of this alphabet, Santali was either not written at all, or was written with the Bangla or Oriya alphabets. Christian missionaries prefer to write Santali with the Latin alphabet. In Nepal some people also write Santhali in Devanagari.

In addition, some of the languages have adopted Devanagri script.

The writing systems used in various languages of Nepal may be schematized as follows:

**Table 1: Scripts**

Script	Language
Devanagari:	Nepali, Maithili, Bhojpuri, Avadhi, Newar, Rajbanshi, Magar, Tamang, Kirat languages
Mithilakshar/Tirhuta:	Maithili
Kaithi:	Maithili
Sambota:	Tibetan, Sherpa
Tamhig:	Tamang
Ranjana:	Newar
Sirijanga:	Limbu
Rong:	Lepcha
Akkha:	Magar
Gurumukhi:	Punjabi

Perso-Arabic:	Urdu
Latin :	Santhali, Gurung
Ol Cemet, Ol Ciki /simply Ol:	Santhali
Khema	Gurung
Bangala:	Bangla

#### 2.4 Language and ethnicity

Nepal is a multiethnic nation. It comprises various ethnic and religious communities. Its ethnic and religious diversity is coupled with its linguistic plurality. They are found to interact with one another, resulting in the threefold ethnic/religious-linguistic structure. First, there are a number of ethno-linguistic communities in Nepal each of which has a common mother tongue. The mother tongues associated with jatis ( also called *janjatis* ‘indigeneous peoples’) include Magar, Tharu, Tamang, Gurung, Limbu, Sherpa, Rajbanshi, Sunuwar, Kumal, Majhi, Danuwar, Chepang, Thami, Thakali, Dhimal, Darai, Lepcha, Byansi, Raute, Raji, Dhangad/Jhangad and so on. Such a situation having one-to-one relation between languages and ethnic groups is “one tribe one language” formula. Besides, Urdu is the mother tongue of a community belonging to a particular religion, viz. Islam. It is, however, to be noted that this formula is not exempt from exceptions. Secondly, there are, on the other hand, ethnic communities each of which speaks several mother tongues. This “one tribe with several languages” in Nepal is the Rai (Kirat) group in the eastern hills and mountain. This single group speaks various Kirati languages such as Bantawa, Chamling, Kulung, Thulung, Sangpang, Khaling, Dumi, Puma, Chhintang, Umbule, Bahing, Nachiring, Koi, Yamphu, Chhiling, Lohorung, Mewahang, Tilung, Jerung, Lingkhim, Sam, etc. This trend is illustrated in the Terai also. In this region, various castes and tribes such as Brahmin, Rajput, Yadav, Teli, Kurmi, Chamar, Khatawe, etc. speak Maithili in the Maithili-speaking area, Bhojpuri in the Bhojpuri-speaking area and Avadhi in the Avadhi-speaking area. Thus, there exists no one-to-one correspondence between ethnic communities and their mother tongues in the Terai. Thirdly, we find the “several tribes with a common language” formula. Different ethnic groups are found to speak a single mother tongue. This formula has been exemplified by Nepali, which is spoken as a mother tongue by various caste groups such as Bahun, Kshetri, Kami, Damai, Thakuri, Sarki, Sanyasi, and so on. Being a *lingua franca*, it has also been adopted as a mother tongue by people from different ethnic groups as well. Hence, it is called an “interethnic language.” Similarly, the Newars consist of various ethnic groups but all these groups speak a single language, viz. Newar or Nepal Bhasa.

In case of one-to-one relation between the tribes and their languages, a comparison of the population of different tribes and their languages can show the extent of language retention by each tribe. According to Gurung (2002: 7-8), there has been considerable increase in the speakers of the languages spoken by various ethnic groups except those of Dhimal, Sherpa and Thakali languages. Rajbanshi and Raji are the two languages whose population exceeds the population of their tribes. Besides, Limbu, Jirel, Thami, and Magar have shown significant increase in their speakers.

The significant increase in ethnic languages can be ascribed mainly to the rise of ethnic consciousness in Nepal which asserts ethnic identity mainly through culture and language. As a result, mother tongue, generally defined as a language learned from parents, has now been understood in Nepalese context as an ancestral language even if someone does not know and speak it. This is evident from the enumeration of the mother tongues in Census 2001.

All these aspects of language-ethnicity relation reflect the dynamism of language shift in Nepalese context. They suggest a common tendency to shift toward regional and ethnic languages. As a result, there has been continuous decline in the numerical strength of Nepali speakers.

### 2.5 *Second languages and lingua francas*

The linguistic diversity existing in the country has given rise to the three situations in the country, viz. monolingualism, bilingualism, and multilingualism. Of these language situations, only monolingualism or the use of mother tongues has been regularly reported since the first modern census in 1952/54. The linguistically homogeneous communities figure 70 per cent of the total population of mother tongue speakers. They are mostly Nepali-speaking people, viz. 46 per cent. Their mother tongue Nepali is also a lingua franca; they hardly need to learn a second language to communicate with a group of people, speaking a different language since they know Nepali. Besides Nepali speaking people, there are people from indigenous groups and also from the Terai who are able to use only their mother tongue. These people are usually old and illiterate and/or hail from remote rural areas with minimum contact with people speaking other languages.

Bilingualism or the use of mother tongues and their speakers' second languages (i.e. the languages most commonly used by them) has, however, been reported since the 1991 census. The 2001 census has reported only 12 second languages used by 92 mother tongue speakers. The total percentage of bilingual speakers is reported to be 30 per cent.

Nepali as a second language has been mostly returned by speakers of Tibeto-Burman languages in 2001 census. In such language communities it serves as a lingua franca. There are also quite a few Indo-Aryan languages whose speakers have returned Nepali as their second language in a large number. Such speakers also live in close proximity with Nepali speakers. They are by and large bilingual except for a few isolated communities. Apart from contact with Nepali-speaking community, non-Nepali speakers also use Nepali as medium of education, media and administration.

The choice of a speaker's second language is not constrained by the condition that it must belong to the language family to which his mother tongue belongs. For example, a Tibeto-Burman (e.g. Newar) speaker may use an Indo-Aryan language (e.g. Nepali or Maithili) as his second language. Instead, the use of a second language is conditioned by contact and pressure for inter-community communication.

It is thus found that many people make vigorous use of second languages while conversing with speakers of other mother tongues. However, this fact does not suggest any details about their level of proficiency in the second language they use. There is a need to evaluate their level of bilingual proficiency.

Multilingualism, i.e. the use of more than two languages, is not as widespread as bilingualism in Nepal. However, it does exist in the country to a large extent. A number of Nepalis are found proficient not only in their mother tongues but also in their *lingua francas* such as Hindi, Tibetan and English and/or neighbouring indigeneous languages. But Census data is not available about third languages. It does not mean that they are equally proficient in all these languages. Such a multilingual situation can be attributed to electronic media, employment and education in India and abroad.

Most speakers of Nepal's languages have been found to be in close contact. As a result, these languages tend to converge through mutual borrowing and influences and gradually share a number of lexical and grammatical features. For example, we find Tibeto-Burman languages "Indo-Aryanized" and Indo-Aryan languages "Tibeto-Burmanized". There is a need for an intensive study of convergence as well as divergence to ascertain the precise nature of Nepal as a linguistic area.

The Indo-Aryan languages spoken in the Terai (e.g. Rajbanshi, Maithili, Bhojpuri, Avadhi, and Tharu) constitute a curious case, in that they are linked uninterruptedly from the viewpoint of intelligibility. That is to say, it is rather difficult to say where one language ends and another begins. However, the degree of intelligibility goes on decreasing with the

increase in distance between the adjacent languages. In such a case, people often adopt Hindi as a link language. Thus, these languages form a continuum (NLPRC 1994).

In the hills and mountains as well, Nepali language forms an uninterrupted cline of its various regional varieties. However, speakers of linguistically distinct Tibeto-Burman languages generally fail to communicate through their mother tongues. Instead, they adopt Nepali as a link language for their interethnic communication. Especially in the mountains bordering with Tibet (China), speakers of various languages are found to use Tibetan as their lingua franca.

A comparison of the various census reports (1952/4 –2001) indicates two conflicting tendencies in language use, viz. language shift and language maintenance. Language shift is a change from the use of one language to the use of another language. This tendency is reflected by decreases in languages and increase in Nepali, the language of the nation, in the censuses ranging from 1952/4 to 1981.

This language shift may be attributed to “one language – one nation” government policy to prescribe a single language for use in administration, education, and media. It may also be due to the pace of migration from one speech community to another (esp. a multilingual community often in urban areas) with the increasing facilities of transport and communication. Besides, it may also happen because Nepali, a single language, is required for employment, opportunities and wider communication as a lingua franca.

There has, however, been perceived a different tendency since the 1991 census. This tendency, called language maintenance, occurs when a group is loyal to its native language and tries to promote, preserve and use it even in a bilingual or multilingual situation. As a result of this tendency, there has been decrease in Nepali and increase in minority languages by 1.7 per cent. It is evident from the comparison of population by mother tongues in the last three censuses (1981-2001). The language maintenance may be due to a number of socio-cultural factors such as restoration of democracy, constitutional provisions, linguistic (including ethnic) awareness and so on.

## 2.6 Existing legal provisions

Despite being multicultural and multilingual, Nepal enshrined 'ethnic', instead of 'civic', nationalism in its task of nation-building (Oakes 2001). It has been reflected in various regimes in the country. Following the Gorkha conquest, the language of ruling elites, also spoken by hill people as lingua franca or mother tongue, assumed the new nomenclature 'Nepali' with an intent to transform it into the national and official language. Rana regime further perpetuated this 'one nation-one language' policy.

With the restoration of democracy there has been growing awareness among non-Nepali speaking people about their culture and languages since they consider them as the symbols through which they strive to assert their identity and recognition. The constitution of Nepal (1991) framed after the restoration of democracy recognized languages other than Nepali and made the following provisions about the non-Nepali languages:

(1) The Nepali language in the Devanagari script is the language of the nation of Nepal. The Nepali language shall be the official language. (*Constitution of Nepal, Part 1, Article 6.1*)

(2) All the languages spoken as the mother tongue in the various parts of Nepal are the national languages of Nepal. (*Constitution of Nepal, Part 1, Article 6.2*)

In addition, the constitution also made a provision for the use of mother tongues in primary education (*Constitution of Nepal, Part 1, Article 18.2*). It also guaranteed Nepalese as a fundamental right to preserve their culture, scripts and their languages (*Constitution of Nepal, Part 1, Article 26.2*).

The greatest weakness of these provisions was the lack of any explicit plan and policy to implement them.

To tease out these constitutional provisions about language, a recommendation commission for formulating policy for national languages was formed by the government in 1993. The main objectives of the commission were twofold: promotion of national languages and their use in local administration, primary education and media. There have been some advances in the use of minority languages in primary education and media. The CDC has so far developed textbooks in 16 national languages to be taught as subject. Another implementation of the commission report has been the introduction of 18 different languages in Radio Nepal though they have been made available very limited time for broadcasting.

The recommendation for the use of minority languages in local administration was later enacted into a law, the *Local Self-Governance Act* of 1999 which deputed to local bodies the right to preserve and promote local languages. Nevertheless, on June 1, 1999, the Supreme Court announced its final verdict and issued a *certiorari* declaring that the decisions of these local bodies to use regional languages were unconstitutional and illegal.

Government of Nepal (GON), which participated in different international fora like Jomtin Convention and Dakar Forum, is committed to making quality primary education accessible to all children including children from indigenous and minority language groups. In pursuance of this commitment, one more goal for Nepal has been included besides other six universal goals of EFA for *ensuring the right of indigenous people and linguistic minorities to basic and primary education through mother tongue*.

To achieve the EFA/Nepal (2004-9) goals a policy of transitional multilingual education policy has been proposed. According to this policy, a child will acquire basic educational skills through the medium of his/her mother tongue and gradually switch to a lingua franca/an official language so that s/he can “feel at home in the language in which the affairs of government are carried on” and finally learn a foreign language (e.g. English) for broader communications and access to science and technology.

Quite recently, *The Interim Constitution of Nepal* (2007), an outcome of the *Andolan II*, has enshrined the following provisions for languages:

(1) All the languages spoken as the mother tongue in Nepal are the national languages of Nepal.

(2) The Nepali Language in Devanagiri script shall be the official language.

(3) Notwithstanding anything contained in clause (2), it shall not be deemed to have hindered to use the mother language in local bodies and offices. State shall translate the languages so used to an official language.

(The Interim Constitution of Nepal (2007), Part 1, Article 5)

Regarding the education and cultural right, the constitution enshrines the following provisions:

(1) Each community shall have the right to get basic education in their mother tongue as provided for in the law.

(2) Each community residing in Nepal shall have the right to preserve and promote its language, script, culture, cultural civility and heritage.

(The Interim Constitution of Nepal (2007), Part 3, Article 17)

In pursuance of the spirit of these constitutional provisions National Curriculum Framework for School Education in Nepal (2007:34) ensures the use of mother tongues in elementary education (Grades 1-3).

To sum up, Nepal is a mosaic of linguistic diversity. However, a single language Nepali was entrusted with all power and prestige to forge assimilation among diverse cultural and linguistic groups while minority languages were looked down as inferior and suppressed.

With the increasing consciousness of individual rights there has been focus on accommodation of minority cultures and languages.

### 3 Earlier linguistic surveys

#### 3.1 *Hodgson's Surveys*

The first attempt of the linguistic survey of Nepal was made by Hodgson (1828-88). During the 19th century B H Hodgson published several papers on the Nepalese languages in the issues of the Journal of the Asiatic Society of Bengal. He published comparative wordlists of Newar (1834), Koch, Bodo (Meche), Dhimal (1847), Gurung, Kiranti, Lepcha, Lhoke, Limbu, Magar, Murmi [Tamang], Newari, Rong [Lepcha], Sherpa, Sunuwar (1847-48), Bhotia, Lepcha, Lhopa, Newari, Murmi [Tamang], Gurung, Kiranti (1853), Bhramu [Baram?], Chepang, Pahari, Thaksya [Thakali], Vayu/Hayu (1857a), Vayu [Hayu] (1857b), Rodong [Chamling], Rungchhenbung, Chingtangya, Nachhereng, Waling, Yakkha, Chourasya [Umbule], Kulungya, Bahingya, Lohorong (1857c), Bahing and Balali (1857-58). In the comparative wordlist of Kiranti (1880) he has also included Thulungya, Lambichhong, Sangpang, Dumi, Khaling and Dungmali. In his 1847 paper, comparative vocabularies of Dravidian and Munda languages also appear. Some of the Munda and Dravidian languages are also spoken in Nepal. In the wordlists he has also given short introductions to the speakers of those languages, which are useful for the sociolinguistic information of the languages of Nepal. In the 1880 wordlist there are about 130 words in the comparison.

Hodgson's methodology is not clear. He seems to have collected data by appointing other people, because there is no consistency in the errors of the same words published at different dates (Toba, personal communication).

#### 3.2 *Carey's Linguistic Survey of India*

According to G A Grierson (1927:11-12), Carey was the first man to undertake Linguistic Survey of India in 1816. Out of 33 languages in Carey's survey, three languages which are also spoken in Nepal are Nepal[i], Mythilee [Maithili] and North Koshala [Awadhi?].

Nepali and some of the Indo-Aryan languages spoken in Nepal are included in the typological study of Indo-Aryan languages by Hoernle (1880).

#### 3.3 *Grierson's Linguistic Survey of India (1898-1927)*

Grierson's Linguistic Survey of India (LSI) is the first comprehensive linguistic survey of the South Asian region.

Following Nepalese languages have been surveyed by LSI (p 389):

1. Austric Family
  - 1.1 Munda Branch
    - 1.1.1 Austroasiatic Subfamily: Santali and Kharia
  - 2 Tibeto-Chinese Family
    - 2.1 Tibeto-Burman Subfamily
      - 2.1.1 Tibetan Group: Bhotia and Kagate
      - 2.1.2 Pronominalized Himalayan Group
        - 2.1.2.1 Western Subgroup: Byangsi, Chaudangsi

- 2.1.2.2 Eastern Subgroup: Dhimal, Thami, Limbu, Yakkha, Vayu, Chepang, Kusunda, Bhramu, Pahari, Rong or Lepcha, Toto and Khambu dialects (Bahing, Balali, Sangpang, Lohorong, Lambichhong, Waling, Chhintang, Rungchenbung, Dungmali, Rodong or Chamling, Nachhering, Kulung, Thulung, Chaurasya, Khaling and Dumi)
- 2.1.3 Nonpronominalized Himalayan Group: Gurung, Murmi, Thaksya Sunwar, Magari, Newari,
- 2.1.4 Assam-Burmese Branch: Bodo or Meche
3. Dravidian Family: Kurukh or Oraon
4. Indo-European Family
- 4.1 Indo-Aryan Branch
- 4.1.1 Sanskrit
- 4.1.1.1 Outer Sub-branch
- 4.1.1.1.1 Eastern Group: Maithili, Tharu Maithili, Magahi, Bhojpuri, Tharu Bhojpuri, Rajbangsi,
- 4.1.1.2 Mediate Sub-branch, Mediate Group: Awadhi
- 4.1.1.3 Inner Sub-branch, Pahari Group, Eastern Pahari: Khaskura or Naipali (standard & Palpa)

Details of Grierson's description of a language (like say Limbu, pp 283-304) cover the following:

- a short introduction to the people
- geographical location
- endonym-exonym(s)
- number of speakers
- census figures
- writing system(s)
- authorities
- bibliography of major publications
- phonemic inventory
- notes on pronunciation
- grammatical notes
- verb paradigms
- language or dialect
- glossed text of each dialect without free translation
- information about literature
- classified list of all the languages
- comparative vocabulary of 168 selected words and
- historical classification into language families

### 3.3.1 Grierson's Methodology

Grierson's (1927:18-19) work sequences were as follows:

1. Preliminary lists of languages
2. Compilation of the lists
3. Local language nomenclature
4. Collection of specimen

## 5. Editing

Grierson's 'basis of the survey' (Grierson 1927:17) was the collection of three specimens:

- The first specimen was a 'standard passage' from the Bible ('The Prodigal Son' story) which was to be translated into each of the languages or dialects or sub-dialects under survey
- The second specimen was to be 'a piece of folklore or some other passage in narrative prose or verse selected on the spot and taken down from the mouth of the speaker'
- The third specimen was a standard list of words and test sentences originally drawn for the Asiatic Society of Bengal by George Campbell.

Grierson collected data with the help of district officers and not through language experts:

- Forms were sent to district officers and political agents to be filled out the names of every language and the number of speakers of each language or dialect.
- The data thus collected were tallied against census reports for the local language nomenclature.
- Foreign languages were separated from native languages.

Following is the Table of Contents of Grierson's product:

Vol I	Part I. Introduction
	Part II Comparative Vocabulary of Indian languages
Vol II	Mon-Khmer and Tai Families
Vol III	Part I Himalayan Dialects, North Assam Group
	Part II. Bodo-Naga and Kochin Groups of TB languages
	Part III Kuki-Chin and Burma Groups of the TB languages
Vol IV	Munda and Dravidian languages
Vol V	Indo-Aryan languages (Eastern Group)
	Part I Bengali and Assamese
	Part II Bihari and Oriya
Vol VI	Indo-Aryan languages, Mediate Group (Eastern Group)
Vol VII	IA lgs, Southern Group (Marathi)
Vol VIII	IA lgs, Northwestern Group
	Part I Sindhi and Lahnda
	Part II Dardic or Pisacha languages (Including Kashmiri)
Vol IX	IA lgs, Central Group
	Part I Western Hindi and Panjabi
	Part II Rajasthani and Gujarati
	Part III Bhil lgs including Khandeshi, Banjari or Labani, Bahrupia, etc.
	Part IV Pahari lgs and Gujuri
Vol X	Eranian languages
Vol XI	Gipsy languages
Appendix I:	Classified list of languages
Appendix I A:	Details of languages and dialects
Appendix I B:	Summary of general table
Appendix II:	List of grapheme records
Appendix III:	Index of language names

### 3.4 *Contributions of Summer Institute of Linguistics (SIL)*

During 1965-75 in Nepal SIL appointed many linguists to describe twenty one different languages. SIL published wordlists, phonemic summaries, texts, sketch grammars and sociolinguistic information. For this SIL made many short term field surveys as well.

### 3.5 *Pokharel and Chaudhari's Survey*

Bal Krishna Pokharel brought an idea of linguistic survey of Nepal during 1965-1968 while he was teaching at Nepali Department in Tribhuvan University, Kirtipur. A Bengali linguist named Shambhu Chandra Chaudhari was teaching in Tribhuvan University those days. He was also involved in the project. Dhruva Chandra Gautam and Madhav Bhattarai were to assist them. According to Gautam (personal communication) they surveyed language varieties spoken in Mahakali and Seti zones of the Far Western Development Region. They surveyed Rana Tharu and Dangaura Tharu apart from the Nepalese dialects spoken in the region.

### 3.6 *Pokharel and Chaudhari's Research Tool*

According to Bandhu and Gautam their mode of data collection was interviews. They had a set of questionnaire consisting of wordlists, grammatical information, paradigms, basic sentences and sociolinguistic details.

The final report of the survey is yet to come out. A mimeograph ["nepaali bhaasaa ko naapi"] by Bal Krishna Pokharel is found in Nepal Academy. Bal Krishna Pokharel is pioneer in the study of dialects of Nepali (Pokharel 2022 VS). This linguistic survey may have been foundational to the study of dialects of Nepali by him. According to him there are five major dialects of Nepali.

#### **3.6.1 B K Pokharel's Research Tool**

Bal Krishna Pokharel's research tool was an extract from Guru Prasad Mainali's story 'Naaso'. He got the questionnaire translated by the informants and analyzed his findings.

### 3.7 *Bandhu's Field Survey and RNA's Survey Design*

In 1968 Chura Mani Bandhu undertook a field survey of Bagmati and Janakpur zones. He was assisted by Daivagyan Raj Neupane and Ram Prasad Bhattarai in 1968-69. According to Bandhu (personal communication) the team surveyed the languages of Bagmati and Janakpur zones. In that expedition the team surveyed eleven different languages. Among the languages surveyed were Thami, Jirel, Majhi, Danuwar, Pahari, Dolakha Newar and Tamang.

Royal Nepal Academy also designed a Linguistic and Cultural survey of Nepalese languages under the leadership of Bandhu, but that plan could not come out due to present political disturbance of the country.

### 3.8 *Rahul Sankrityayan's Survey*

According to Janak Lal Sharma (personal communication) Rahul Sankrityayan also tried to survey some of the languages of Nepal while he was preparing to write a book on Nepal. His methodology was translating a paragraph into the Nepalese languages. Janak Lal Sharma got the translations in Kathmandu. He found rare informants in the Supreme Court, Police Headquarter and Army Headquarters.

### 3.9 *Glover's Survey of Gurung Dialects*

Warren Glover and John K Landon (1980) conducted a survey of Gurung dialects, in appreciation of which Toba (1998:28) has following remarks:

This is the first-full blown survey report I've found on Nepali languages. Contents: local opinion, word lists, isoglosses, sound shifts, intelligibility testing (Casad method modified), conclusions. Charts are included for the distribution of the languages, etc.

Glover (1971) has also done a Swadesh list calculation on thirty Tibeto-Burman languages. He has published a comparative Swadesh 100 wordlists of Tamangic languages and has designed a comparative reconstruction of the languages in a tree model.

### 3.10 *Winter's Linguistic Survey of Nepal (1981-84)*

Werner Winter came to survey Nepalese languages from 1981 to 1984. The title of the project was 'Linguistic Survey of Nepal'. The survey was financed by German Research Council. Twenty two university graduates were given some basic orientations and trainings and sent for collecting data (Winter 1984, 1986). Field surveys were started in Eastern Development region and gradually moved from East to West. It covered the languages of both Terai and highland. Following languages of eastern Nepal were surveyed by the project:

#### Languages in Winter's survey

- |                               |  |
|-------------------------------|--|
| 1. Athpare and Belahare       | 2. Bahing (Rumdali)                          |
| 3. Bantawa                    | 4. Chamling                                  |
| 5. Chhintange                 | 6. Chhilung                                  |
| 7. Dungmali (Bhojpur)         | 8. Dimmali/Dummali                           |
| 9. Dumi                       | 10. Hayu                                     |
| 11. Jerong                    | 12. Khesang (Bhojpur)                        |
| 13. Khaling                   | 14. Koyu                                     |
| 15. Limbu                     | 16. Kulung (Sankhuwa Sabha)                  |
| 17. Nachhiring, Para, Khotang | 18. Kulung (Sotang) in Chheskam, Sagarmatha  |
| 19. Sangpang, Khotang         | 20. Lohorung, Mane Bhanjyang, Sankhuwa Sabha |
| 21. Nechali                   | 22. Mewahang, Sankhuwa Sabha, Tamku          |
| 23. Puma                      | 24. Sunuwar                                  |
| 25. Thulung                   | 26. Umbule/Bonu                              |
| 27. Waling, Bhojpur           | 28. Yakkha                                   |
| 29. Thami                     | 30. Yamphu                                   |
| 31. Dhimal                    | 32. Meche                                    |
| 33. Danuwar                   | 34. Newar                                    |
| 35. Newari                    | 36. Uraon                                    |

- |                     |                     |                    |            |
|---------------------|---------------------|--------------------|------------|
| 37. Satar           | 38. Sotang, Bhojpur |                    |            |
| 39. Other languages |                     |                    |            |
| a. Amdo             | b. Gurung           | c. Nyorung Tibetan | d. Parali  |
| e. Tamang           | f. Jirel            | g. Telugu          | h. Madhesi |
| i. Burmese          | j. Manipuri         |                    |            |

The project did not cover all the languages and all the geographical and development regions of Nepal. It was limited only to the eastern Nepal. At CNAS in Tribhuvan University the whole data of Winter's survey are alphabetically ordered in different folders and one of the folders is captioned 'Madhesi'. The survey was not extended for the languages spoken to the west of the Kathmandu valley. Only one survey report of Rai Kiranti languages (Hansson 1991) has come out of this project.

### 3.10.1 Winter's Survey Tool

Winter designed a questionnaire which covers primary wordlist (250 words) and syntax and morphology in addition to some sociolinguistic information.

### 3.11 Michailovsky's Typological Survey of Nepalese Languages (1988)

Boyd Michailovsky (1988) has published a survey of phonological typology of Nepalese languages.

### 3.12 M Pokharel's Survey of Nepali Dialects

During the years 1990-95 Madhav Pokharel supervised dissertations of five Nepali MA students (Timilsina 1993, Niraula 1993, Subedi 1994, Acharya 1997 and Dhungana 1997) for the dialect surveys of Nepali. The students were given different sets of questionnaire to survey the dialects. Yamanath Timilsina based his study on the variation of pronominals and numerals, Yagyeshwar Niraula conducted his survey on the variation of verb forms, Sakhi Sharan Subedi surveyed variations in the syntactic structures, Bhagavat Acharya limited his survey to the variation in pronunciation and Lavanya Prasad Dhungana checked the variations in the 100 word list of Morris Swadesh.

M Pokharel's students collected data in Kathmandu by consulting students at hostels and apartments in Kirtipur, on Tahachal campus and parliamentarians at Sinha Darbar.

### 3.13 Austin Hale's CNAS Proposal for Sociolinguistic Survey

Austin Hale (1993) has designed a sociolinguistic survey proposal with a set of questionnaire. It was proposal for a joint project to be conducted 'under the auspices of the Centre for Nepal and Asian Studies with the sponsorship of the Summer Institute of Linguistics'.

### 3.14 Research and Language Documentation Programs at CDL (1999-)

Linguistic Survey of Nepal has been on top priority since the inception of Central Department of Linguistics (CDL) at Tribhuvan University in 1996. Thesis writing has been

compulsory at MA level and students are encouraged to work on undocumented languages. By now students have written their MA and PhD level theses on different topics of the following languages: Bahing, Bhojpuri, Bhujel, Bote, Boto Boli or Khanchi, Byangsi, Chamling Rai, Chepang, Chhatthare Limbu, Chhintang, Churaute, Danuwar, Darai, Dhimal, Dura, Hayu, Jhangad, Kaike, Kumal, Limbu, Magar, Manipuri, Meche, Nepali, Newari, Sanskrit, Sunuwar, Tamang, and Thakali.

Besides, students and teachers of CDL have produced dictionaries and descriptive sketch grammars of different indigenous languages with support from NFDIN.

### 3.15 *Linsun Project*

Around 2002 the Department designed a proposal of Linguistic Survey of Nepal (LinSuN Project). The idea was primarily initiated by Professor Michael Noonan of University of Wisconsin. The idea of Linsun was shared by native and foreign linguists at University of Berne, Switzerland. David Watters came with a idea of Encyclopedia of Nepalese Languages to be a part of LiSuN. Chhintang-Puma Documentation Project and Baram Documentation Project were also included in the LinSuN, but it could not emerge as a full-fledged Linguistic Survey of Nepal.

In addition to these attempts research works of many native and foreign linguists included and yet to be included in the bibliographical works like Toba (1991, 1998), have published many works which prove to be of significant contributions to the proposed Linguistic Survey of the country.

### 3.16 *Toba et al's Surveys*

Toba *et al* (2002) have produced a sociolinguistic profile of the following 59 Nepalese languages:

#### a. Tibeto-Burman languages:

- |             |             |              |                |              |
|-------------|-------------|--------------|----------------|--------------|
| 1. Athpare  | 2. Bahing   | 3. Bantawa   | 4. Baram/Bramu | 5. Belahare  |
| 6. Byasi    | 7. Chamling | 8. Chepang   | 9. Chhintang   | 10. Dhimal   |
| 11. Dolpali | 12. Dumi    | 13. Dura     | 14. Ghale      | 15. Gurung   |
| 16. Hayu    | 17. Jirel   | 18. Khaling  | 19. Kham       | 20. Koyu     |
| 21. Kulung  | 22. Lepcha  | 23. Limbu    | 24. Lhomi      | 25. Lohorung |
| 26. Magar   | 27. Meche   | 28. Mewahang | 29. Mugali     | 30. Newar    |
| 31. Pahari  | 32. Puma    | 33. Raji     | 34. Raute      | 35. Sangpang |
| 36. Sherpa  | 37. Sunuwar | 38. Tamang   | 39. Thakali    | 40. Thami    |
| 41. Thulung | 42. Wambule | 43. Yakkha   | 44. Yamphu     | 45. Yohlmo   |

#### b. Indo-Aryan languages:

- |           |             |          |           |               |
|-----------|-------------|----------|-----------|---------------|
| 1. Awadhi | 2. Bhojpuri | 3. Bote  | 4. Darai  | 5. Danuwar    |
| 6. Kumal  | 7. Maithili | 8. Majhi | 9. Nepali | 10. Rajbamshi |
| 11. Tharu |             |          |           |               |

#### c. Austroasiatic language: Santal

#### d. Dravidian language: Jhangad

#### e. Language isolate: Kusunda

### 3.16.1 Toba et al (2002)'s Survey Tool

According to Novel Kishore Rai (personal communication) Toba et al (2002) used the questionnaire of the *UNESCO World Languages Report (China)*.

Sueyoshi Toba (1973) had also collected a comparative vocabulary of about fifteen Kiranti languages.

### 3.17 Bielmeier's Survey of Tibetan Dialects

Professor Roland Bielmeier (To Appear) of University of Berne, Switzerland conducted a survey of Tibetan dialects and compiled a Dictionary of Tibetan Dialects. In his survey several Nepalese Bodish languages of the Tibeto-Burman family are also studied.

### 3.18 Minor Surveys

Sociolinguistic surveys of some of the Nepalese languages have been undertaken by some SIL linguists, like Warren Glover (1980), Jeff Webster (1992, Ghale and Bodish dialects of North Gorkha), Bryan Varencamp (1996, Eastern Tamang), Nirmal Tuladhar (1999, Jirel), Epele et al (2001, Rajbamshi), Dhimal et al (2001, Dhimal), Kwang-ju Cho et al (2003, Bantawa Rai), Glover et al (2003, Gurung), Maureen Lee (2005, Bayung), Jeff Webster (Limbu, 2001) and Steve Watters (1988, Kham and 1998, Mugu).

Two running projects namely, Chhintang-Puma Documentation Project supported by Volkswagen Foundation and Baram Documentation Project funded by SOAS at Central Department of Linguistics are also parts of the Linguistic Survey of Nepal project.

KR Khambu (?) has published a comparative Swadesh 100 word list of 26

Kiranti languages. Kirat Rai Bhasha Tatha Sahitya Parishad Jhapa (2062 VS) has also published another comparative Swadesh wordlist of 28 Kiranti languages.

## 4 Objectives

In spite of the surveys that have already been conducted in Nepal, a clear understanding of the diversity of language communities still persists. Given the lack of understanding, the primary goal of this Survey is to give clarity to the sociolinguistic situation of the language communities of Nepal. In support of this goal, we seek to develop an archive which houses a database of information about the languages of Nepal. We will also seek to empower language communities to use their language in written domains. More specifically, LinSuN aims to:

1. Develop a sociolinguistic profile of all the languages of Nepal. Such a description will include:
  - a. information relating to the geographical location of languages (including maps),
  - b. a demographic description of speakers of the languages,
  - c. lexical variation
  - d. mutual intelligibility of dialects,
  - e. an increased understanding of what it would take to standardize the language,
  - f. the extant of multilingualism
  - g. attitudes of mother tongue speakers to their own language and others commonly used,
  - h. language use and vitality,
  - i. and an increased understand of a language's readiness for use in education and governance.

2. Produce a basic description of at least ten languages (at least one description in each cluster: see section 7 for a clarification of the term cluster) that includes an understanding of the sound system, observations on the grammar, and a trilingual glossary.
3. Develop and maintain a complete database of the languages of Nepal which will include:
  - a. the sociolinguistic data mentioned above
  - b. recordings of oral texts
  - c. transcriptions of these texts, and if such exists texts which are already written
  - d. videos of sign languages
  - e. basic grammatical data
  - f. glossaries
4. Develop a description of the use of mother tongues in education (formal and non-formal) as a means to better understand the development needed for mother-tongue curricula in the national educational system which will include:
  - a. a literacy in use survey to understand what orthographies and materials have been developed in the languages of Nepal
  - b. the development of orthographies and the publication of folk tales in at least ten languages.

## 5 Methodology

While the survey will employ methodologies for studying the sociolinguistic profile of a language community, one foundational aspect we would like to highlight is the importance of community involvement. The success of this endeavor will depend on the ability of the survey to interact with the language community and include it in the decision making of language planning. As such, an important initial aspect of the survey will be to involve community leaders in planning certain aspects of the survey.

To accomplish the specific goals set out for the LinSuN, the following methods and techniques will be adopted:

### 5.1 *A sociolinguistic description of all the languages of Nepal (Objective 1)*

Census reports from the government of Nepal and previous surveys and linguistic publications will be used as a starting point to determine the geographic location of the languages spoken in Nepal. To provide an accurate sociolinguistic profile of the languages of Nepal, the following methods will be employed:

#### 5.1.1 Lexical similarity

On the basis of standard word lists elicited from mother tongue speakers, we will determine thresholds of lexical similarity uniting groups of languages and dialects at various percentage levels. The results can be presented in a table which illustrates the relative linguistic distances among various speech communities, and lexical differences can be compared in an exhaustive matrix of pairs. The technology for evaluating, comparing, and interpreting word lists is relatively mature (Blair 1990:21-32; Simons 1984). The use of word lists in providing a picture of language area intelligibility is well illustrated in O'Leary 1992. (See Annex B).

### **5.1.2 Comprehension of Recorded Texts**

Recorded Text Testing (RTT) is an established technique for determining the levels of intelligibility that obtain among language groups. The technology for this is also quite mature (Casad 1974; Blair 1990:33-34). This approach to intelligibility involves a far greater investment of time and energy than the collection of word lists and sample paradigms and utterances since it involves the construction of valid tests and the administration of these tests to a representative sampling of speakers in each language community under study. Word lists and other collected samples, however, give only a rough indication as to how well speakers of one community comprehend the speech of another community.

The comparison of true cognates between dialects can be an even less reliable indicator of intelligibility than comparison by way of lexical similarity. True cognate pairings in the mid ninety percent range have correlated with low intelligibility within the Kham-Magar language area.

In any event, if one is serious about obtaining reliable, quantified measures of mutual intelligibility, there is at present no viable substitute for testing comprehension directly by means of recorded speech.

### **5.1.3 Levels of Competence in the Nepali Language**

Sentence Repetition Testing (SRT) developed recently by survey investigators working in South Asia operates on the assumption that one cannot correctly repeat long complicated utterances which one does not understand (See Annex D). The technology, described in Radloff 1992 and in Blair 1990: 93- 96 is relatively new, but is sufficiently well tested to be incorporated as part of a strategy for intelligibility survey. The heavy investment in this technique is in the construction and calibration of the initial test. Once a valid test has been constructed, it is possible to administer the test in a very short time to large enough numbers of people to obtain statistically significant results reflecting a cross section of the population in a given speech community. If such tests were devised to measure levels of competence in Nepali possessed by a cross section of mother tongue speakers of other languages throughout Nepal it would be a significant contribution to understanding the effectiveness of communication by the media in Nepal.

We propose that such a test be devised to measure the more formal register of Nepali used on Radio Nepal news broadcasts and that another test be devised to test the register used in the less formal language of the Radio Nepal agriculture programs. One could propose other levels and registers to study as well.

A questionnaire has been developed and will be piloted for developing standardized forms, understanding languages functioning as link languages, investigating attitudes of mother tongue speakers to their own language and others commonly used, discovering orate/literate development within languages, and evaluating their vitality and viability, and their readiness for use in education and governance (See Annex C).

### **5.1.4 Sociolinguistic Questionnaires**

Language vitality has to do with the prediction based on current language attitudes and language use patterns that a language will continue to be spoken by mother-tongue speakers in the foreseeable future. This is an important area to study when any decisions that have to do with language are being made.

In this survey, we will seek to understand language vitality through the use of questionnaires and observation. Language use refers here to what languages people speak in what domains of life--the choices people make about which language to use in different situations. Language attitudes refers here to the attitudes people have towards their own

language and other languages around them. The questionnaire will elicit three types of information:

1. Biographical information
2. Questions seeking to elicit factual information about language use
3. Questions eliciting opinions about the language

It has been noted that questions seeking objective information are often answered subjectively where it is common for a subject to answer in the way he/she thinks that the questioner wants it answered. This is particularly true for questions of type 2. However, if there is uniformity of answers this still reflects attitudes to the language, even if it diverges from actual behavior.

While in the field, surveyors will also observe patterns of actual use, and make notes of those observations. It will be instructive to compare these notes with the language use reported in questionnaires.

## 5.2 *Basic Linguistic Description (Objective 2)*

### 5.2.1 **Grammar Sketches**

In spite of its interest, the Survey will need to limit the number of grammatical sketches it produces. The Survey will produce a basic description of at least ten languages (at least one description in each cluster: see section 7 for a clarification of the term cluster).

The Survey will use the framework that was developed by Watters and Yadava (2003) for the Encyclopedia of Nepalese Languages for writing grammar sketches for this survey. The framework proposed for the Encyclopedia is Basic Linguistic Theory, which they describe as “a natural framework for preparing the grammatical sketches of the various languages. It is a “cumulative” framework in the sense that it employs insights from a broad array of theoretical approaches, from generative grammar to functional-typological grammar — the kinds of insights that have proven, over time, to be the most useful and the most widely accepted in describing individual languages. The framework is also “user friendly” in that it is not theory bound or theory laden, but employs natural language prose in its descriptions” (2003: 2).

Watters and Yadava (2003) proposed the following outline as a something to work off while writing a grammar sketch. But they warn against following it as an iron clad template without exceptions but to use it as a suggestive of the kinds of things a full grammatical sketch might include.

1. Introduction
  - a. Language name & alternate names
  - b. Notes on subgrouping
    - i. Dialects
    - ii. Geographical Locations
    - iii. Map of language speaking area (with village names)
  - c. Language use and maintenance
  - d. Notes on language and language research history
2. Phonology
  - a. Consonants
  - b. Vowels
  - c. Syllable structure
  - d. Tone, register, accent

- e. Phonological alternations
- f. Basic orthography
- 3. Inflectional and derivational morphology
  - a. Nouns and pronouns
    - i. Noun classifiers
    - ii. Alignment of cases
      - Ergative / Split Ergative
      - Anti-ergative
      - Accusative
      - Primary and secondary objects
    - iii. Local case markers
    - iv. Compounding
  - b. Adjectives
    - i. Derived adjectivals
      - Verbal sources
      - Nominal sources
  - c. Verbs
    - i. Transitive/intransitive/copulas
    - ii. Verbal compounds
    - iii. Complex predicates
    - iv. Tense-aspect-modality
    - v. Agreement
      - Hierarchical agreement
      - Subject–Object agreement
      - Conjunct/Disjunct
      - Inverse marking
    - vi. Evidentials
    - vii. Mirative
    - viii. Voice systems
      - Reflexive, reciprocal
      - Passive, medio-passive/middle voice, anti-passive, etc.
      - Deponents
    - ix. Causatives, applicatives, benefactives
- 4. Syntax
  - a. Structure of the NP
    - i. Relative clauses
  - b. Structure of the clause
    - i. Word order
    - ii. Serial verbs
    - iii. Clause chaining
      - Switch reference
      - Sequential vs. simultaneous
    - iv. Major sentence types
      - Direct affirmative declarative
      - Negation
        - Scope
    - v. Questions
      - Content questions
      - Yes–no questions
      - Tag questions

- c. Imperatives
  - d. Subordinate clauses
  - e. Sentential complements
5. Basic vocabulary
  6. Bibliography

### 5.2.2 Lexicon

In the same set of languages in which the survey conducts grammatical research, we will also develop a glossary. Toolbox or FLEX will be the software tools of choice. One source of the lexicon will be the vocabulary that comes from interlinearizing native texts. Another source will be through a method referred to here as ontology. This is a method that has been used with considerable success in African languages, and has been used with two language communities in Nepal. This is an approach that uses a group of people from the language community in a workshop setting to generate large numbers of words based on a semantic domain based questionnaire. In Nepal, this approach has been successful in generating ten thousand words in a matter of two weeks.

We will seek to have at least 5000 headword entries, each of which will contain fields such as lexeme, pronunciation (using IPA symbols), pronunciation (in Devanagari), part of speech, definitions (in Nepali and English), example sentences (selected from the corpus, context,) etc..

For Nepali-speaking audience, we will produce a Nepali-native language index. This index will help not only non-Nepali speakers but also native speakers of Nepali to learn and understand this language.

For international communication, we will also produce an English- native language index.

## 5.3 *Develop and maintain a complete database of the languages of Nepal (Objective 3)*

### 5.3.1 Database: Written corpora

In spite of an interest in developing a comprehensive written corpora of the languages of Nepal, our efforts in this area will draw upon existing corpora where these can be identified, like the Nepali National Corpus, developed by the Bhashasanchar Project and other digitized materials.

This collection will be digitized or converted (if already digitized) into Unicode.

For the purpose of computer processing, these texts will then be manually formatted using XML tagging in the body, paragraph, sentences and foreign words appearing therein.

Each text will be provided with the metadata or bibliographical details such as book/article/ issue title, author, publisher, publication date, publication place, name of the typist, etc. in XML header. This annotation will be done for the purpose of archiving and to be used as resources by various language communities and researchers.

### 5.3.2 Database: Spoken corpora

The spoken corpora will be developed from the texts used in the sociolinguistic survey (RTTs), as well as from a limited set of domains and genres. These corpora will be presented with their audio-video recordings.

These recordings will be annotated and phonologically transcribed for use in analyzing their linguistic and extralinguistic features.

They will be transcribed using ELAN software and interlinearized using Toolbox or FLEX. This will be helpful for compiling dictionaries and grammars.

To specify metadata relevant to a text, an international standard will be used. This standard, which contains information such as genre, author, year of publication, publisher, etc., is found to be useful for digitally archived language documentation

To manage and integrate the corpus data, each of its text will be annotated using Toolbox or FLEX, which is of tremendous help for field linguists working with lexical, grammatical and cultural analysis of texts. In annotating the corpora, transcription will be done using IPA fonts and where possible Devanagari Unicode.

#### 5.4 *Nepali sign language*

As sign languages will be archived through videography, a sign language corpora will be created and appropriate methods of linguistic annotation and tagging developed. The study of sign languages will call for a unique methodology as the signers of these languages are not only predominantly monolingual but also typically become users of these languages only when they come into contact with other deaf people.

Because the Nepali sign language is different than the spoken and written languages, different sets of survey tools will be prepared for the sociolinguistic survey of the Nepali sign language.

The survey questionnaires will be prepared for the deaf, their family members, the interpreters and experts.

Necessary visual signs for the development of sign will also be collected to make Nepali sign language more effective for communicative and educational purposes.

Necessary information about the bilingual and multilingual deaf people will also be collected.

A critical review of the development of Nepali Sign Language and the works done in the past for its development will be assessed.

As a high level of proficiency in the Nepali Sign language and good background in linguistics as well as Sign Language Linguistics will be essential for the linguistic description of the Nepali Sign Language, Nepali students will be encouraged to study and analyze the structures of the Nepali Sign Language.

#### 5.5 *Use of mother tongues in education and literacy in Nepal*

As in sociolinguistic survey, questionnaires will be used to conduct a literacy in use survey.

In addition to a literacy in use survey, the Survey research team will work with the community to develop an 'ad-hoc' (working) orthography. This orthography will be based on the phonological analysis developed as part of documentation efforts.

However, it is important for the survey to help the community understand some of the broader issues of orthography development. In the multilingual setting, a common script helps develop national integration and facilitates inter-group understanding. Common diacritics for the modification of the Devanagari script should also be understood. The survey can also help the community understand the value of using the Devanagari script to help it aid in transitional material to Nepali. Though the speakers of the language are final decision makers about the script of their language, interests of the larger community should be taken into consideration while selecting a script for a particular language.

It will not be possible to do much in the way of literacy in this survey. In order to introduce the idea in communities where there are no written texts, this survey will put to writing folk tales and other native stories, and publish these as a model of literature in the mother-tongue.

## 6 Management

There shall be a **Steering Committee** at the National Planning Commission, Government of Nepal. It will have a supervisory and evaluative role and provide general guidance to the Survey. The National Planning Commission member in charge of Social Inclusion shall be the Coordinator of the Steering Committee. In addition, the National Planning Commission shall invite senior linguists and dignitaries from related government agencies to join this committee.

For conducting the Survey there shall be a **Survey Team** which will consist of the following linguists:

**Table 2: Survey Team members and their research interests and functions**

Name	Position	Affiliation	Research interests	Potential functions
Prof. Dr. Yogendra P. Yadava	Team Leader	Central Department of Linguistics, Tribhuvan University	Research on Maithili syntax. Publications on Syntactic Theories & Descriptive Grammars of Indo-Aryan languages including Maithili and indigenous languages	Sketch grammars of Indic languages as a key language expert. Overall coordination as the administrative Head of the Survey; Survey Analysis & write-ups; Training & Editing.
Prof. Dr. Chura Mani Bandhu	Member	Central Department of Linguistics, TU	Research on syntax in Nepali. Conducted field work on languages and folklore.	Literacy, Sign language, Mother-tongue Education; Coordination of Survey of Nepali dialects and other areas; Training of RA's & Editing
Prof. Dr. Tej R. Kansakar	Member	Central Department of Linguistics, TU	Research on Phonology of Kathmandu Newari; Tibeto-Burman Linguistics, lexicography & translation studies	Survey of Newar dialects, Phonetics & Phonology of minority languages, lexicon development; Typological studies; Training & Editing
Prof. Dr. Madhav P. Pokharel	Member	Central Department of Linguistics, TU	Research on the Phonetics & Phonology of Nepali language; Indo-Aryan & Tibeto-Burman languages of Nepal.	Survey of Nepali dialects, Sociolinguistic surveys, Dialectology & Typological studies, Training & Editing Development of orthography and training mother tongue speakers how to produce grammar, dictionary and writing system.
Prof. Dr. Novel K. Rai	Member	Central Department of Linguistics / Centre for Nepal & Asian Studies, TU	Research on syntax of the Bantawa language; Chief Coordinator of Puma-Chintang Project	Descriptive & comparative studies of the Kirat languages; Training & Editing
Prof. Nirmal	Member	Central	Sociolinguistic	Sociolinguistic

M. Tuladhar		Department of Linguistics / Centre for Nepal & Asian Studies, TU	studies; Comparative lexicon of Newar language; Sociolinguistic Profile of Jirel language.	surveys; Literacy & Mother-tongue education; Training & Editing.
Dr. Carl Grove	Member	Central Department of Linguistics / SIL Intl.	Research on Language Planning; Software programming, Management coordination	Sociolinguistic survey methods; Data processing; Training of Field workers; Editing & preparations for publications
Mr. Steve Watters	Member	Central Department of Linguistics / SIL Intl.	Research on Dzongkha language of Bhutan & Tibeto-Burman languages of N-W Nepal; Sociolinguistic profile of Mugu language, Research on Bodic languages	Acoustic phonetics, Phonemic analyses of undocumented languages; Sociolinguistic survey. Descriptive grammars; Training of Field workers; editing & preparations for publications.

The Survey Team introduced above will function as a policy-making and monitoring body under the administrative auspices of the Central Department of Linguistics, Tribhuvan University. This Survey Team will also hold regular meetings and ensure the availability of required personnel to be recruited, trained and assigned specific duties. As the LinSuN Survey progresses and we gain more knowledge and experience with the language situations in particular areas, the Team will have a guiding function in the step-by-step development of the Survey.

The Survey Team has also proposed a National and International Advisory Group. The role of this Group will be consultative and will consist of the following scholars and specialists :

- Dr. Ramawatar Yadav, Vice-Chancellor, Purbanchal University, Biratnagar
- Dr. Subhadra Subba, Senior Linguist / CNAS, Tribhuvan University
- Prof. Dr. Kamal P. Malla, Senior Consultant, Tribhuvan University
- Dr. Austin Hale, SIL International, USA
- Prof. Dr. George van Driem, Leiden University, The Netherlands
- Prof. Boyd Michailovsky, CNRS, Paris, France
- Prof. Dr. Balthasar Bickel, University of Leipzig, Germany
- Prof. Dr. Carol E. Genetti, University of California, Santa Barbara, USA
- Dr. David Watters, SIL Int'l and Research Centre for Linguistic Typology, USA
- Prof. Michael Noonan, University of Wisconsin, Milwaukee, USA
- Prof. Dr. Scott DeLancey, University of Oregon, Eugene, USA
- Prof. Dr. Tony McEnery, Lancaster University, UK
- Dr. Andrew Hardie, Lancaster University, UK
- Prof. Dr. David Bradley, La Trobe University, Australia
- Prof. Sun Hongkai, Chinese Academy of Social Sciences, Beijing, China
- Prof. Yoshio Nishi, Kagoshima University, Japan
- Mr. Sueyoshi Toba, SIL Int'l, Japan
- Dr. Mark Turin, Digital Himalaya Project, University of Cambridge, England

Prof. Dr. Udaya Narayan Singh, Director, Central Institute of Indian Languages, Mysore, India.

Prof. Dr. K.V. Subbarao, University of Delhi, India

Prof. Peri Bhaskararao, Research Institute for Languages & Cultures of Asia & Africa, Tokyo University of Foreign Studies, Japan

Dr. Chris Wilde, SIL Intl, Finland

## 7 Work Plan

### 7.1 Introduction

While the instruction of the National Planning Commission (NPC) is to complete a survey within seven years, by all expert opinion, this is rather ambitious. The original work plan of the survey, in fact, included a number of research and work tasks which went beyond a simple profile kind of approach to survey. But in the current proposal many activities have been scaled back, and a more modest plan is proposed.

The survey has four main research activities (hereafter referred to as components). These are sociolinguistics, documentation, literacy<sup>1</sup>, and archiving. Each of these components are in and of themselves significant research activities, but this survey will be done in a modest scale. It is believed that the integration of sociolinguistics, documentation, literacy, and archiving will make a more compelling contribution to a language community than would otherwise be possible with one activity in isolation. It is hoped that this approach will provide a foundation for future efforts of the Government of Nepal in reaching its aim of granting linguistic rights to the citizens of Nepal.

Given the complexity of the linguistic situation in Nepal, and the logistical difficulties in achieving a thorough study of each linguistic community, it won't be easy to achieve even a scaled back set of goal's for this Survey. However, the basic idea that we would like to follow is something we're calling the 'cluster approach.' The idea of a cluster approach is simply that we treat groups of closely related languages as a single unit (albeit an often loosely affiliated unit), and that several research teams develop a specialization for that group of related languages, and that these research teams carry out the work in that cluster for the entirety of the survey. This is an attempt to recognize that communities and the languages they speak are not isolated from one another and that to understand their situation well, one has to understand how language communities interact with the communities around them. This approach is also an attempt at a good management practice of developing local expertise

---

<sup>1</sup> In this document, we use the term literacy as a catch-all term for mother-tongue development in education. This education may be part of the formal system for children, as in the Multi-Lingual Education movement, or it may part of informal adult education, often referred to as Non-Formal Education.

and specialization in the members of the survey team and in the members of the language communities.

The cluster approach is in contrast to an approach where languages are studied based on some other criteria, like starting with languages with the smallest number of speakers and moving to the largest, or having research teams with one narrow specialization that look at only that one specialization for a whole set of languages. This is an attempt at a more holistic approach to research.

As noted elsewhere in the proposal, there are four language families spoken in Nepal.<sup>2</sup> Some of these families have large numbers of languages spoken in them and some have only a small number of languages. In terms of number of speakers, the Indo-Aryan language family is the largest, but in terms of the number of distinct languages, the Tibeto-Burman language family is the largest. The Tibeto-Burman language family has at least two major branches – the Himalayish branch (the pronominalizing branch) and the Bodish branch (the branches closely related to Tibetan), and then within these broad branches the languages can be divided by finer criteria. We will refer to these finer divisions as clusters.

The clusters that are proposed for this survey are (in no particular order):

- Chepang-Bhujel
- Tamangic
- Kham Magar
- Standard Magar
- Newar
- Nepali and its dialects
- Tibetan
- Kirat (four sub-clusters)
- Terai (four sub-clusters)<sup>3</sup>

This is to say that research teams will focus on these groups of languages, and that unless their expertise is needed somewhere else in the survey, they will devote their time to that group of languages.

There are four main research components in each cluster. These are sociolinguistics, documentation, literacy<sup>4</sup>, and archiving. The focus of this survey is sociolinguistic in nature –

---

<sup>2</sup> In addition to these four language families, Nepal is also home to a few remaining speakers of Kusunda. Their language is regarded as a language isolate.

<sup>3</sup> The designation of ‘Terai’ is geographical in nature. Most of the languages in the Terai belong to the Indo-Aryan family, but there are a few languages that belong to the Austro-Asiatic and Dravidian families. As these are few in number, they are lumped together with the other I-A languages for the purposes of a work plan.

most teams will be involved in this endeavor. But the survey will also involve at least one research team in each cluster. It will be important that each research team works in a collegial and collaborative spirit with the other teams of the Survey and the language communities themselves to facilitate the overall goals for each community.

The sequence of research activities for a component in a given cluster will need to be timed. They cannot be entirely sequential in nature, and will need to be partially integrated. That is, the survey will not do sociolinguistics, then documentation, then literacy, and finally archiving in chronological order with no overlap between them. Rather the activities of a given component will feed into and benefit from the activities of other components.

The focus of the initial research in a cluster will be sociolinguistic in nature, as the first goal of the survey is identification and classification. But soon after the initial assessment efforts, documentation can begin, as well as literacy. While these activities are underway, it will be important to enter and maintain all the information in easily accessible data bases (i.e. archiving). So, at any given time, there may be several research teams working simultaneously on a given cluster. When the survey is fully underway, there will be several research teams working in as many as nine major clusters (and several geographical regions within a cluster).

In summary, then, the work plan is to carry out largely concurrent activities in the four main components of the survey in each of the linguistic regions of Nepal. Each region will have several research teams working in parallel, and these teams will work together in collaboration to fill out the overall linguistic profile for a linguistic community and begin to provide a foundation for future work.

## 7.2 *Time Frame*

The time frame of the survey will span seven years. This is divided into two phases: a two year phase followed by a five year phase. The motivation for these time frames is to fit into the National Planning cycle of the Government of Nepal. The current National Plan will end in two years, and hence the first phase of LinSuN will be two years. The next National Plan will be five years in duration, and LinSun will fit its goals into this.

## 7.3 *Man Power*

The man power needed for this survey is not inconsequential. As noted above, there will be several research teams for a single cluster at any given time. If one multiplies this by the number of clusters, the survey will need at least 119 researchers during the height of its

---

<sup>4</sup> Before the survey begins it isn't clear which languages are regarded as standard and which are not. In some clusters, there is diglossia, such as in the Tibetan cluster. It may be that the Nepali cluster should be viewed in this way as well. Whatever the case may be, appropriate literacy strategies will need to be employed for that particular cluster.

production. The survey plans to draw on man power from both students of Tribhuvan University and members of the language communities themselves.

A Sociolinguistic Team will consist of a team leader and two research assistants. The other kind of research team will conduct a combination of documentation and literacy activities (see goals of documentation and literacy to understand the limited nature of these activities). This latter team will consist of four members: a team leader and three research assistants. The survey, then, will conduct research with each cluster with the following man power (119 staff in 32 teams):

- 2 Chepang-Bhujel Research Teams
  - 1 SL, 1 Doc + Lit
  - 7 total members
- 2 Kham Magar Research Teams
  - 1 SL, 1 Doc + Lit
  - 7 total members
- 2 Standard Magar Research Teams
  - 1 SL, 1 Doc + Lit
  - 7 total members
- 2 Newar Research Teams
  - 1 SL, 1 Doc + Lit
  - 7 total members
- 2 Nepali and its dialects Research Teams
  - 1 SL, 1 Doc + Lit
  - 7 total members
- 3 Tamangic Research Teams
  - 2 SL, 1 Doc + Lit
  - 10 total members
- 3 Tibetan Research Teams
  - 2 SL, 1 Doc + Lit
  - 10 total members
- 8 Kirat (four sub-clusters) Research Teams
  - 1 SL, 1 Doc + Lit for each sub-cluster
  - 32 total members
- 8 Terai (four sub-clusters) Research Teams
  - 1 SL, 1 Doc + Lit for each sub-cluster
  - 32 total members

Given that the research activity generated by this scale of man power will be significant, the survey will also require a solid administrative staff. The administrative structure of the project is summarized in the Table 3 (p. 49) and on page 56.

#### *7.4 Overview of Research Activities*

In Phase 1 (year 1 – 2), the survey will lay a solid foundation for the coming years through training, research in the work that has already been done, and building an archiving system that will serve the project throughout its entirety. However, we also seek to have done enough research to make initial language and dialect determinations. In addition, we will publish a number of volumes: a volume with the word lists collected up to that point in the survey, an annotated bibliography of research done on the languages of Nepal, the results of a Literacy in Use Survey, a bibliography of mother tongue materials available in Nepalese languages, and of particular note a volume on the major typological features of the languages of Nepal.

In Phase 2 (year 3 – 7), the project will be in full production mode. Sociolinguistic work will be concurrent with basic documentation and literacy. At the end of this phase, the survey will produce sociolinguistic profile of all the languages of Nepal. This will include maps. In addition to publishing the results of the sociolinguistic survey, this project will complete a basic description in at least ten languages. In the literacy component, orthographies will be developed in at least ten languages, and a compendium of published folk tales made available to the speakers of those languages.

In addition to producing the final volumes of the sociolinguistic and grammar research, the survey shall make a comprehensive set of recommendations to the GoN for language planning and inclusion of linguistic communities in the national fabric of Nepal.

As part of the training component, three candidates will be selected for further education – two for an MA and one for a PhD – and these will be given scholarships for their academic study. In addition, funds will be made available to members of the project to attend two language related conferences each year throughout the Survey – one South Asia based conference and one internationally based conference.

Finally, an outside evaluator shall be invited to assess the success of the survey. In particular, it shall be evaluated for achieving its objectives. In addition, the evaluator will look at how well those objectives have had an effect on the outcomes of the Survey.

## OVERVIEW OF SURVEY PLAN

(Underlined items are items that will be published)

- TRAINING
  - international level training in best practices of Field Linguistics
  - training in sociolinguistic survey techniques
  - professional development
    - professional conferences, and MA and PhD programs
- SOCIOLINGUISTICS
  - sociolinguistic survey of each of the major genetic groupings of languages
    - lexical similarity, dialect intelligibility, bilingualism, and language use and language attitude questionnaires
  - Publications
    - Publish word list and lexical similarity analysis (phase 1)
    - Sociolinguistic profiles of each language (phase 2)
- DOCUMENTATION
  - compilation of previous research
  - basic documentation in at least one language in each major genetic grouping of languages
    - phonological analysis, small text corpora, glossary, basic grammatical observations
  - Publications
    - Annotated bibliography of previous linguistic research (phase 1)
    - Typological overview of the Tibeto-Burman languages of Nepal (phase 1)
    - Basic documentation published in at least ten languages (phase 2)
- LITERACY
  - Literacy in Use Survey
  - ‘Ad-hoc’ (working) orthography in at least ten languages
  - Publications
    - Annotated bibliography of mother-tongue materials (phase 1)

- Summary of LiU Survey (phase 1)
- Folk tales in at least ten languages (phase 2)
- OVERALL RECOMMENDATIONS
  - a comprehensive set of recommendations to the GoN for language planning and inclusion of linguistic communities in the national fabric of Nepal.

### 7.5 PHASE ONE (Year 1 and 2)

The activities of each research component for phase one are summarized at the end of section 7.5

#### 7.5.1 Sociolinguistic Component

The primary purpose of the sociolinguistic component is to identify the number of languages, to determine the number of dialects spoken within each language, to determine domains of language use, and to determine language use and vitality. These inter-related areas will be investigated using word lists, recorded text tests (RTTs), sentence repetition tests, and interviews with questionnaires.

In phase 1, before beginning a full-fledged sociolinguistic survey, the survey will conduct a pilot survey as a means of testing and practicing its own methodologies. This survey will serve as a training ground for the whole survey, as well as for staff that will be employed for the sociolinguistic component.

After the initial pilot survey, and an assessment and modification of the practices employed during that survey, research teams will fan out into the different geographical regions to begin taking word lists. They will begin by meeting with community leaders, and developing community profiles through interviews with the leaders. After initial planning time with community leaders, the research teams will collect word lists and more full community profiles. The focus of word list collection will be to determine the extent of lexical variation throughout the linguistic communities of Nepal. The goal is to get a word list from every 'distinct' linguistic community. Some languages may have little variation over a limited geographical area and word list collection will go quickly. Other languages may have immense variation over an extended area, and word list collection may take considerable time.

We anticipate word list collection being completed in the first phase for some of the smaller clusters like Chepang-Bhujeli, Kham Magar, Standard Magar, Newar, Tamangic, Austro-Asiatic/Dravidian, and Nepali and its dialects. For the other clusters (Tibetan, Kirat, and Indo-Aryan), this work may not be complete by the end of phase 1.

The word lists will then be analyzed for lexical similarity, and this determination will serve as a basis for where to begin recorded text tests during the second phase of the survey.

The man power needed for this component for this phase will be 17 teams (of three members each).

#### 7.5.2 Documentation Component

The primary purpose of the language documentation component is to address issues of loss, and to provide a foundation for future work in language development. The issue of loss has to do with language shift from a socially and linguistically vulnerable language community to a stronger more dominant language community. Before that shift occurs completely and there are no more speakers of that language, linguists seek to document these languages. Sometimes the shift is irreversible and there is little that can be done other than documentation. It is hoped, however, that in some cases efforts in documentation can help prevent the trend in language shift and provide a means by which the community can preserve their linguistic heritage.

In phase 1, one major objective will be to produce an annotated bibliography of the linguistic research that has been done in Nepal. Toba's bibliography of Nepalese languages is an important step in that direction, and this part of the project will seek to annotate some of the more important writings of this bibliography. This annotation will be published at the end of the first phase.

Based on the word list research in the sociolinguistic component, the research compiled for the annotated bibliography, communication with linguists, and the involvement of the language community itself, the survey will determine where to begin documentation projects. Unlike sociolinguistic research which entails visiting many communities, the documentation component will focus on just one speech variety in one of the languages of the cluster. Preference will be given to the documentation of languages that show signs of language shift and loss of language vitality.

Another objective of phase 1 will be to publish a typological overview of the Tibeto-Burman languages of Nepal. This last work is something that has been in progress for a number of years, and the author, David Watters, would like to contribute this as one of the first works of the project. This work will be particularly helpful, then, in providing researchers a linguistic and typological framework in which to write their own grammar sketches of the undocumented languages of Nepal.

In addition, the survey will compile as many texts of Kusunda as possible in a two year span. These oral texts will be recorded, accurately transcribed, interlinearized, and translated into Nepali and English. These texts will be archived according to international best practices.

The focus of the survey in phase 1 will be word list collection, so the man power needed for documentation in phase 1 is small. The essential need is to compile, read, and annotate the literature which is already existent. The man power needed for this will be one team: a team leader and three research assistants. This team should have good proficiency in English, and the ability to read academic books and articles. One team will also be assigned to the Kusunda text collection project.

### **7.5.3 Literacy Component**

Some may question the place of a literacy component in a linguistic survey. However, the purpose of the literacy component is not so much to engage in education, as to make short forays into developing mother-tongue materials for education use. The aims of this component are quite limited in scope.

Having said this, this project will conduct a survey of Literacy in Use. The purpose of this survey is to ascertain the extent to which different linguistic communities are practicing reading and writing, and what languages and scripts are being used for these purposes. The results of this survey will be published at the end of phase 1.

In addition to a Literacy in Use survey, the survey will compile what has already been done to implement the use of the mother-tongue in education, and other formalized settings such as radio. The purpose of this research will be to determine where the holes are in literacy, and to use this understanding in parallel with the Literacy in Use survey. The two combined studies should give an accurate assessment of the use of the mother-tongue in domains that require written or standardized language. The outcome of this will be an annotated bibliography of mother tongue materials that have been produced in the languages of Nepal.

The man power needed for the literacy component in phase one will be three teams: two to conduct the LiU survey and one to compile mother tongue materials.

The activities in sociolinguistics, documentation, and literacy are summarized in outline below.

PHASE ONE GOALS: (YEAR 1 AND 2)  
(underlined items are items that will be published in that phase)

- TRAINING
  - three scholars are trained at InField (this is really pre-project, but noted here to demonstrate that training has begun)
  - basic survey course at CDL
  - pilot survey of Newar dialects in Kathmandu valley
    - pilot survey is a training ground for research assistants
    - particular attention is given to:
      - word list elicitation techniques
      - making clear recordings
      - making clear transcriptions of these recordings
      - getting an accurate free translation of a recording
- SOCIOLINGUISTICS
  - initial contact with language communities
    - discussions with community leaders about survey
    - advocacy with community leaders for the involvement of the community in the survey
      - the ideal will be to have at least one member from the community to be one of the Research Assistants in the Research Team
    - fill out questionnaire with community leaders that helps to give a general profile of the community
  - For Chepang-Bhujeli, Kham Magar, Standard Magar, Newar, Tamangic, Austro-Asiatic/Dravidian, and Nepali and its dialects
    - word list collection from each distinct speech variety
    - analysis of the word lists in the above seven clusters
      - initial dialect determination based on lexical similarity
    - publish word list and lexical similarity analysis
  - For Kirat, Indo-Aryan, and Tibetan
    - partial word lists (partial in that we will not yet have collected a word list from each distinct speech variety in these clusters)
- DOCUMENTATION
  - publish an annotated bibliography of the linguistic research that has been done
    - Toba (1998) bibliography in particular will serve as a base reference for this part
    - use annotation in combination with the work in word lists to point to where potential documentation projects could be undertaken
  - publish a typological overview of the Tibeto-Burman languages of Nepal (contribution from David Watters).
  - Kusunda texts (as many possible in a two year span)
    - recorded, accurately transcribed, interlinearized, and translated into Nepali and English
    - archive the recorded and transcribed texts
- LITERACY
  - produce an annotated bibliography of mother tongue materials that have been produced in the languages of Nepal

- Literacy in Use Survey
  - this could be part of the questionnaire used to interview community leaders at the outset of the survey.
  - publish a summary of LiU survey

## 7.6 Phase 2 (Years 3, 4, 5, 6, and 7)

### 7.6.1 Sociolinguistics

In Phase 2, the research teams that have not yet completed getting the range of variation in word lists, will finish this task, and proceed with analysis to determine where to administer recorded text tests. Research teams that have already completed the word list phase of their work will proceed with testing dialect intelligibility through RTTs. This phase of the sociolinguistic survey can be quite time consuming, and take considerable patience on the part of the researchers.

Sociolinguistic interviews will be given to the individuals who take RTTs.

In addition to studying dialect variation, bilingualism will be studied through the use of Sentence Repetition Tests (SRTs).

As noted above dialect intelligibility testing will be conducted in the clusters that first complete word list work. These will likely be the Chepang-Bhujeli, Kham Magar, Magar, Newar, Tamangic, Austro-Asiatic/Dravidian, and Nepali and its dialects.

By the end of the first three years of Phase 2, the survey plans to have completed its work in five clusters and will publish its findings in several volumes. These volumes will also include maps of language and dialect boundaries. In the remaining two years of Phase 2, the survey will have completed its work in the remaining clusters, and the results of this research will be produced in several more volumes. These volumes will also include maps.

If the dialect studies in the larger clusters is not complete, an extension will be requested.

As a result of these studies in sociolinguistics, LinSuN will make a comprehensive set of recommendations to the Government of Nepal addressing issues of language policy and planning. It is hoped that these recommendations will serve as the foundation for future language development work in Nepal, and that they serve as the basis on which to include every distinct linguistic community into the national fabric that makes up Nepal.

The man power needed for this component for this phase will be 17 teams (of three members each).

### 7.6.2 Documentation

In phase 2, the documentation component will focus on basic linguistic description and text collection. By basic description the survey means the following activities:

- Phonological analysis (to be used in preparing an ‘ad-hoc’ (working) orthography and transcription scheme in inter-linearized texts)
- A sampling of texts from different spoken genre i.e. folk tales, personal narrative, etc..
- Transcription, inter-linearizing, and free translations of these texts into Nepali and English
- Glossary based on these texts and other vocabulary elicitation techniques
- Some basic grammatical observations in a theory-neutral way based on the outline in the Encyclopedia of Nepalese languages

The survey plans to complete a basic description in at least ten languages (at least one in every cluster).

In some languages, we will follow a method of vocabulary collection that has been used with considerable success in African languages, and has been used with two language

communities in Nepal. This is an approach that uses a group of people from the language community in a workshop setting to generate large numbers of words based on a semantic domain based questionnaire. In Nepal, this approach has been successful in generating ten thousand words in a matter of two weeks.

As glossaries become available, the survey will publish them and make them available to the language communities.

A compilation of basic grammatical observations will also be published.

Some texts (particularly folk tales) will also be published as part of the literacy component to produce mother tongue reading materials.

The man power needed for this component for this phase will be 15 teams (of four members each). This includes the man power needed for the literacy activity for phase 2 mentioned below.

### 7.6.3 Literacy

The goals of the literacy are quite modest, and in many ways could be considered to be an extension of activities in documentation. In language communities that have an interest in writing down their language and publishing materials in their own language, the survey will make an effort to develop an 'ad-hoc' (working) orthography, and to use that orthography to put to writing some of the folk tales recorded in the documentation effort.

This effort will need to happen from among the ten languages chosen for basic documentation. Priority will be given to those languages that have not yet had any written material developed in their language and endangered languages.

The goals of phase two are summarized in the outlines below.

#### PHASE TWO (A) GOALS: (YEAR 3, 4, AND 5)

- TRAINING
  - Training in collaboration with InField is conducted in Kathmandu
  - Training in Recorded Text Tests (RTTs) and Sentence Repetition Tests (SRTs)
  - Professional meetings in South Asia and one International venue
    - Linguistics Conferences
    - Multi-lingual Education (MLE) Conferences and Workshops
- SOCIOLINGUISTICS
  - For Chepang-Bhujeli, Kham Magar, Standard Magar, Newar, Tamangic, Austro-Asiatic/Dravidian, and Nepali and its dialects
    - RTTs are developed for each of the dialect areas that show up as needing further study according to lexical similarity studies
      - (RTTs only need to be administered where lexical similarity shows marginal similarity i.e. dialects that are between 60 and 80 percent lexically similar.)
    - Dialect intelligibility and bilingualism are studied through RTTs and SRTs
    - Sociolinguistic questionnaires are also administered to those taking RTTs and SRTs
    - Publish Sociolinguistic Profile of languages in Chepang-Bhujeli, Standard Magar, Newar, Austro-Asiatic/Dravidian clusters
    - Publish Sociolinguistic findings of Nepali and its dialects
  - For Kirat, Indo-Aryan, and Tibetan
    - finish word list collection

- analysis of the word lists in the above three clusters
  - initial dialect determination based on lexical similarity
  - RTTs are developed for the dialect areas that show up as needing further study according to lexical similarity
  - Publish word list and lexical similarity analysis
- DOCUMENTATION
  - A couple of caveats:
    - Documentation will start in a language community after initial dialect intelligibility studies
    - It will not be feasible in this project to conduct documentation in each distinct dialect or language
    - A documentation project in a given dialect of a given language will start only after results from survey are discussed with community leaders, and consensus emerges with community leaders about the best place to begin documentation efforts
  - Depending on the needs expressed by community leaders, the aims of documentation described here will be quite modest:
    - Phonological analysis (to be used in preparing an ‘ad-hoc’ (working) orthography and transcription scheme in interlinearized texts)
    - A sampling of texts from different spoken genre i.e. folk tales, personal narrative, etc..
    - Transcription, interlinearizing, and free translations of these texts into Nepali and English
    - Glossary based on these texts and other vocabulary elicitation techniques
    - Some basic grammatical observations in a theory-neutral way based on the outline in the Encyclopedia of Nepalese languages
  - It is anticipated that a documentation project will begin in at least one language of every cluster (note: this is a more modest goal than saying a documentation project in every language).
  - Publish Phonological analysis along with working orthography in at least three languages
  - Publish Glossary in at least three languages
- LITERACY
  - In those language communities that have an interest (in consultation with community leaders):
    - develop an ‘ad-hoc’ (working) orthography
    - publish folk tales and other materials of interest to the community
  - Priority will be given to those languages that have not yet had any written material developed in their language and endangered languages

#### PHASE TWO (B) GOALS: (YEARS 6 and 7)

- TRAINING
  - Professional meetings in South Asia and one International venue
    - Linguistics Conferences
    - Multi-lingual Education (MLE) Conferences and Workshops
- SOCIOLINGUISTICS
  - For Kham Magar and Tamangic clusters:

- Publish Sociolinguistic Profile of each language in these two clusters
  - For Kirat, Indo-Aryan, and Tibetan
    - By the end of year 6, draw to a conclusion what research has been completed
    - Publish Sociolinguistic Profile with what can be said up to that point with the research completed.
    - If necessary, request an extension of the project to complete the dialect and intelligibility studies of the remaining languages.
- DOCUMENTATION
  - A continuation of Phase Two (A), working with community leaders on the results of the sociolinguistic survey to determine places where to begin documentation.
  - Depending on the needs expressed by community leaders, the aims of documentation described here will be quite modest:
    - Phonological analysis (to be used in preparing an ‘ad-hoc’ (working) orthography and transcription scheme in interlinearized texts)
    - A sampling of texts from different spoken genre i.e. folk tales, personal narrative, etc..
    - Transcription, interlinearizing, and free translations of these texts into Nepali and English
    - Glossary based on these texts and other vocabulary elicitation techniques
    - Some basic grammatical observations in a theory-neutral way based on the outline in the Encyclopedia of Nepalese languages
  - Publish Phonological analysis along with working orthography in at least seven languages
  - Publish Glossary in at least seven languages
- LITERACY
  - In those language communities that have an interest (in consultation with community leaders):
    - develop an ‘ad-hoc’ (working) orthography
    - Publish folk tales and other materials of interest to the community
  - Priority will be given to those languages that have not yet had any written material developed in their language and endangered languages

### 7.7 *Incomplete Documentation*

It is clear that this survey will only scratch the surface in documenting the linguistic diversity of the languages of Nepal. A great deal more will be known when the survey is completed, and no doubt this will spawn further interest in language documentation. While the survey is not requesting funds beyond the seven year plan, it should be acknowledged that much more will remain to be done, and that there will be need for further projects in documentation and literacy.

### 7.8 *Archiving Component*

It is well recognized in the linguistic community that the world’s languages are being lost at an alarming rate as younger generations stop speaking their language and choose instead to make another language their primary language. It is likely that some of the languages that LinSuN will work in will be languages that will cease to be spoken in a generation or two. It is also possible that many more languages will cease to be spoken sometime in the twenty-second century. As such, LinSuN will strive to keep all of its data in

such a way that it will be accessible as a record for centuries to come. This includes both the actual language data it collects, as well as the reports and write-ups that it generates. LinSuN will make its data available in ways that conform to ethical and legal best practices.

In the first phase, the main objective will be to build several LinSuN databases in such a way that they conform to international best practices on language archiving. The database needs to be built in such a way that the format is long-lived and can be easily migrated to improvements in computer technology. Unlike the prose of a word processing document, linguistic data such as interlinearized texts and glossaries are highly structured. The databases which house LinSuN data need to be appropriately structured to accommodate the unique needs of linguistic data. This should not be in a proprietary way, but in a way which conforms to best practices of other language data archives<sup>5</sup>.

In addition to building a database in phase 1, researchers will need to enter the data that is being collected through research. This will include information about the community, the data collected in sociolinguistic interviews, lexical data, and texts, as well as digital data files that contain recordings and pictures.

The man power needed for the archiving component will be constant throughout the survey. It will consist of one team with six persons: a team leader and archiving coordinator, a database designer, two database programmers, and two database managers.

Research teams will be the primary ones responsible for entering data into the databases.

### *7.9 Final Recommendations*

While the research and material that will be published by this Survey will be of great interest to a great many people in the academic community, the main purpose of this survey is to build a foundation for and facilitate the inclusion of linguistic communities into the national fabric of Nepal. The survey will result in a comprehensive set of recommendations to the Government of Nepal for language planning and policy. These recommendations will be presented to the Government of Nepal in the final year of the Survey.

### *7.10 Summary Of Work Plan*

In summary, then, the Survey will integrate research in sociolinguistics, documentation, and literacy to arrive at a fuller profile of the linguistic communities of Nepal. The research will be conducted on the basis of already known linguistic groupings, referred to here in this work plan as clusters. Research will be conducted concurrently in each cluster. The time it takes to complete the research and write-ups for each cluster will vary considerably

---

<sup>5</sup> For text data this usually means that the text is structured, and that each structure is marked with an appropriate 'handle'. There are a variety of ways of marking up text. One common ways of marking up a text is through XML.

according to the remoteness of the language communities, and the number of languages within each cluster. However, the overall time frame of the survey will be seven years.

At any given time, there may be multiple research teams working in a single cluster, each team focusing on one of the survey components. The clusters that are proposed for this survey are:

- Tibetan
- Tamangic
- Kham Magar
- Magar
- Chepang-Bhujel
- Newar
- Kirat (four sub-clusters)
- Terai (four sub-clusters)
- Nepali and its dialects

It will be important for these teams to work in a spirit of cooperation and collaboration to integrate the knowledge that is being gathered about each language community.

Finally, this work will be documented in a way that conforms to international best practices on archiving so that everything that is compiled is accessible for generations to come.

## **8 Human Resources and Training Requirements**

The success of the LinSuN Survey will depend largely on well-motivated, well-trained and fully committed human resources. The precise details of our manpower requirements have been worked out in the work plan. At least the following four different areas of training will be necessary for successful survey work under LinSuN:

### *8.1 A Phase by Phase requirement*

The manpower requirements for a long-term Survey like LinSuN will be fairly large, but these can be filled in several phases as follows:

7. The initial stage for the preparation of questionnaires, training in methods and techniques, and to develop a core work force to undertake field work.
8. The second stage will involve linguistic and sociolinguistic documentation of languages in select areas. These tasks will be accomplished by several groups working in coordination.
9. The third stage will be devoted to sifting, collating and transcribing the collected data of Texts, Lexicon, Audio and Video materials, and transferring the data on the computer for Editing and Analyses.
10. The fourth stage will undertake creation of Data Formats, Meta Data, and use of Software tools such as Toolbox and ELAN for Data Encoding. This phase will also prepare Archive quality materials by the use of Interlinearized Data, Audio and Video recordings.
11. The final stage will require editors, senior linguists, grammarians, Lexicon-builders and Consultants to write language descriptions.

There will obviously be some overlaps in certain tasks across various phases, but it is also clear that the manpower requirements will vary from stage to stage.

### *8.2 Special Training Programmes*

1. Recruitment of Research Assistants who have completed their M.A.'s in Linguistics with basic experience in field research and description of an undocumented language.

2. Training and upgrading of skills relevant to the Survey work through regular workshops and task-oriented meetings.
3. The development and support to specializations in Field Studies, including preparation of descriptive grammars and lexicon, and methods of Sociolinguistic Surveys.
4. Orientation and training to impart skills relevant to the survey work of language informants and other community participants, as well as students enrolled in University language departments such as English, Nepali, Maithili, Newar (Nepal Bhasa), Sanskrit, Hindi etc.

### 8.3 *Academic Training Programmes*

The main focus of such training programmes will be to provide thorough orientations in field work, data gathering, phonetic, phonological, morphological and syntactic description, as well as Sociolinguistics and Typological Studies. In addition, these training programmes may also include workshop sessions in Tagging, Database creation and Updation, Visual Phonetics and Transcription, and training in making Areal maps and Linguistic Atlases.

### 8.4 *Analytical and Technical Manpower*

We can also foresee the need for several Analytical Teams who will function as specialists of various aspects of language related to different levels of linguistic analyses. For example, we would require different levels of specialist manpower, as follows :

1. The language community will be regarded as having a prominent role in consulting in this project. It is their knowledge that we will be documenting and seeking to promote.
2. The expert services of Phoneticians, Phonologists and Morphophonologists, Morphologists and Syntacticians, Semanticists, Typologists, Lexicographers, Sociolinguists, Anthropologists, and Archivists, and so on.
3. We would also require the consultant services of experts of certain family of languages – Indo-Aryan (with the largest number of speakers), the Tibeto-Burman (with the largest number of languages), the Austro-Asiatic / Munda and Dravidian (consisting of a minority group of languages). On the basis of current evidence, we would also need to add Kusunda, a language isolate of Nepal.
4. Other specialists such as Statisticians and Management experts (e.g. MBA's) may be necessary to provide reliable statistics of speakers, language data, and efficient management of the Survey. The entire Survey will thus be coordinated so that the work of one team will often become the input for the other and all of them will meet regularly to exchange their findings.
5. The Survey work in progress will most certainly require the assistance of several professionals with technical background, such as experts from the Information Sciences, Software Developers etc to advise and improve the analytical tools and computational analyses of data, and also to solve certain unforeseen problems that may arise.

**Table 3: LinSuN Organizational Structure**

		Steering Committee (Executive Board)		
		International Advisory Board		
		Senior Management Committee		
		Director		
Consultants		Personal Assistant		
Support		Research Departments		
Operations Manager	IT Team Leader/Archiving Coordinator	Sociolinguistics Coordinator	Documentation Coordinator	Literacy Coordinator
Finance Manager	Database Designer	Research Teams	Research Teams	Research Teams
Book Keeper (2)	Database Programmer (2)	Team Leader	Team Leader	Team Leader
Logistical Support	Database Manager	Research Assistant	Research Assistant	Research Assistant
Office Staff (4)	Computer Support (2)	Research Assistant	Research Assistant	Research Assistant
Driver (2)		Research Assistant	Research Assistant	Research Assistant
(the number of research teams in each department will vary depending on the focus at the time of the overall project. The Sociolinguistic and Documentation Departments may have up to 15 teams at a time.)				

### 8.5 Training Allocations

The LinSuN Work Plan envisages four components for carrying out the Survey, namely Sociolinguistics, Language Documentation, Literacy and Archiving. The total period of seven years will consist of Phase 1 (Years 1,2), Phase 2a. (Years 3,4,5), and Phase 2b. (Years 6,7). Each of these phases provides for salaries, training and development of human resources required for accomplishing the tasks under the four components. The following is a summary of the phase-wise provisions for trainings and budgetary allocations to achieve the set targets in the four areas of the Survey :

Sociolinguistic Component :

Phase 1. Training in Survey Course : Years 1-2.

Phase 2a. Professional Development (South Asia Conferences) : Years 3-5.

Phase 2b. Professional Development (South Asia Conferences) : Years 6-7.

Language Documentation :

Phase 1. Infield Course (UCSB,USA) :Years 1-2.

Phase 2a. Infield Course (Kathmandu?) :Year 3

Phase 2b. Training in Documentation : Years 4-5  
(including Archiving)

Literacy :

Phase 1. Training in Orthography Development : Years 1-2.

Phase 2a. Training in Material Development : Years 3-5

Phase 2b. Training in Material Development : Years 6-7

Archiving :

Phase 1. Training : Database Management Workshops : Years 1-2.

## 9 Expected results

The primary outcome of LinSuN is to lay a foundation that provides for the linguistic rights of the citizens of Nepal so that all her people, regardless of linguistic background, will be included in the overall fabric of the nation. By better understanding of the linguistic and ethnic diversity of the Nepalese people, it is hoped that no community will be excluded from the activities of the nation because of language, and that each language community will be given a foundation for long term language development.

On the completion of the Survey the following results can be expected:

1. This survey will contribute to a comprehensive database with information about each language of Nepal. Each language will have at least a sociolinguistic profile in the database. Many languages, however, will have a fuller picture with an annotation of previous research, grammatical information, and a glossary size lexicon.
2. Policy makers at the government and political levels, linguists, as well as language communities will have access to this database to help them plan and formulate language policies. It is hoped that this will help interested parties make informed decision about including marginalized languages in nation building efforts.
3. The Survey will foster discussion about what it takes to make a spoken-only language to a standardized written language that can be used in written and other more formalized domains.
4. Through its activities in documentation in a limited set of languages, previously spoken-only languages will have a written form that can be used in limited ways in education, media and local administration. It is believed that extending its use in this way will help language communities keep their languages.
5. The host institution and other academic institutions will benefit from the findings of this study as they incorporate the information from this research into their curricula in teaching linguistics and languages.

6. This survey will help GOs and INGOs who are interested in mother tongue education through the information it gleans about orthographies and materials that have been developed in the languages of Nepal, as well as better understanding what remains to be done.
7. In the initial phase of the Survey, a series of trainings will be conducted to orient the researchers in various methods and techniques (including language technology) required for field work and analysis of data. This will develop capacity and expertise for further linguistic studies in Nepal.
8. A series of high quality publications on the languages of Nepal will be made available to the language communities and other interested parties. This material will be published electronically as well as in book form.

## **10 Ethical Issues**

### *10.1 The fieldwork*

1. This phase of fieldwork is the most important time in the survey period. As soon as the relationship between the researchers and the native speakers is established, it can help to further the task of survey.
2. During the whole period and especially at the fieldwork situation the researchers should maintain high respect to the individuals, communities and cultures they work with.
3. Oral consent must be sought at the time of recording for each and every item. The researchers should start every recording of the items such as word list, sentence patterns or a text with consent of the native speaker. The native speakers should be informed about how the manual and mechanical recording is done.
4. The speaker will have the right to immediately delete any portion of the recording they are uncomfortable with, and will have access to the data they have supplied.
5. The individuals and communities should be made clear that the researchers will not snatch away their cultural heritage like the language and items of folklore. They have come to collect linguistic materials to store in the archives, write grammar and compile dictionaries, use in preparing literacy materials and disseminate their unique features by analyzing them. But at the same time they should be explained that the researchers will not assert their ownership on the language and oral traditions they possess.
6. The individuals and communities will be paid reasonable amount of money for long hour's time they spend with the researchers while recording the linguistic items and also for the travel and other services. But in no price ownership of the indigenous knowledge can be transferred. The ownership of the indigenous people on intellectual property must be recognized.
7. The research team will follow the recommendations of the World Intellectual Property Organization (WIPO) and the Copyright Act (2001) and Copyright Regulations (2003) of Nepal Government while acquiring linguistic and folklore materials and publishing them. It is mentioned in the Act that the ownership of the creative works and folklore materials remains with the individual creators and the sponsoring organization will have the copyright of their recorded items.

### 10.2 Archiving

1. Archiving is another important phase in the survey. All the materials (questionnaire, field notes, recorded tapes, videos, photographs etc.) collected during the fieldwork will be archived systematically.
2. The archived materials should be accessible to the individuals and communities as well as to the researchers and students.
3. If the native speakers and researchers feel it appropriate to put certain restrictions on the access of some of the archived items, they can be inaccessible for certain period only.

### 10.3 Publication

1. Publication is the last phase of the survey. For wider dissemination of the findings of the survey in a form of the reports will be published in different volumes in English language. The international language is chosen to make the findings of the survey accessible for the users worldwide. But brief information on the survey and findings and literacy materials etc. will also be published in the national languages with the consent of the concerning individuals and institutions of the linguistic groups.
2. Due acknowledgements will be given to the, authorities, advisors, researchers, report writers, fieldworkers, consultants, tradition bearers and others helping for the success of the survey program.
3. Consents of the native speakers or communities should again be sought to publish their interviews, images, conclusions of the joint works etc. The items or part of the items that the native speakers want to make confidential, will not be published.
4. Academic honesty and high level of academic standards must be maintained in writing, reporting and publishing the survey results as well as other items.
5. The LinSuN will try its best to disseminate the findings of the survey by publishing them as soon as they are ready for the benefit of the native speakers, the country and academic world at large.
6. LinSuN can recommend, permit and help the researchers, students and native speakers to use and publish the materials with appropriate acknowledgements. Detailed guidelines will also be prepared for this purpose.

## 11 Reporting requirements

1. **Annual Reports:** A progress report comprising the technical and financial positions of the survey will be submitted to the National Planning Commission after the end of each year of the survey. The NPC Steering Committee will evaluate these reports and provide general guidance to the Survey activities reported.
2. **Final Report:** A final report will be submitted no later than four months after the completion of the survey. This will include an evaluation of the extent to which the original aims and objectives have been fulfilled and the survey's outcomes together with a detailed breakdown of annual expenditure.
3. All the reports will be submitted in hard copy, soft copy, and via e-mail.

## 12 Institutional Profile

The Central Department of Linguistics (CDL) is located on the Kirtipur Campus of Tribhuvan University under the Faculty of Humanities and Social Sciences. Established in May 1996, it has been engaged for more than a decade in promoting studies in linguistics leading to degrees in all the academic levels such as I.A., B.A., M.A. and Ph.D. Teaching and research programs seek to impart in-depth knowledge of linguistics and its practical

applications in the Nepalese context including a recently introduced course in IT-based natural language processing.

Manned by a competent and qualified faculty comprising a dozen permanent and another dozen part-time faculty members, and six administrative staff this department has been involved in funded-research in lexicography, grammar writing, sociolinguistic studies and, above all, natural language processing. It has gained experience and expertise in conducting the following international projects:

1. Chintang and Puma Documentation Project (DOBES, Germany 2004-2008, in collaboration with Leipzig University, Germany), with the aim to provide a rich linguistic and ethnographic documentation of two highly endangered but almost totally undocumented languages in eastern Nepal, Chhintang and Puma.
2. Nepali Language Resources and Localization for Education and Communication (NeLRaLEC) Project, funded by the EU Asia IT&C programme, (reference number ASIE/2004/091-777) for the period of 2005-2007 and undertaken by a consortium of Nepali and European partners including the Open University, UK; Madan Puraskar Pustakalaya, Kathmandu; and Central Department of Linguistics, Tribhuvan University, Kathmandu; ELRA; the University of Göteborg, Sweden; and Lancaster University, UK. A variety of Nepali language technology support outcome were achieved within Nelralec, including software localisation and font development, the development of a *Nepali National Corpus* (NNC), the online Nepali dictionary, text-to-speech and so on.
3. Linguistic and Ethnographic Documentation of Baram Language (Endangered Language Documentation Project, SOAS, University of London, 2007-2010) This project aims to document Baram, an extremely endangered language of Nepal, for its preservation and promotion. For this purpose, we will develop a corpus of the various texts of the Bara□ m language and thereby prepare resources such as its sketch grammar, lexicon and ethnographic profile. To revitalize the language, we will develop Bara□ m orthography and prepare a Baram primer to facilitate its use in basic education and literacy programmes. The audio-video recordings of the Baram database will be archived with its annotation in a website for its accessibility to linguists, ethnographers and the Baram speech community.
4. The department has completed the basic dictionaries and grammars of 25 endangered indigenous languages with support from National Foundation for Development of Indigenous Nationalities.

CDL has been providing consultancy services in language-related issues to Central Bureau of Statistics, Nepal Academy, Curriculum Development Centre, and so on. It has signed MoUs with several international universities, viz. University of Oregon (USA), University of Leipzig, University of Mainz (Germany), and Arts and Humanities Research Council (UK). It has been provided consultancy and affiliation to foreign linguists in carrying out research on various languages of Nepal.

In addition, the CDL has the most comprehensive collection of materials on linguistics and languages in Nepal. This includes over three thousand volumes and five international academic journals. This collection of materials has been catalogued digitally and can be accessed by an electronic database.

Recently, it collaborated with UNESCO, Ministry of Education and Sports (MoE), CNAS, TU and SIL Int'l in organizing an international Multilingual Symposium in Kathmandu. We have also completed the Nepali translation of *Advocacy Kit for Multilingual Education* (UNESCO, Bangkok, 2006) with adaptations to Nepalese context in association with MoE and UNESCO Office in Kathmandu and also the English version of *National Language Policy Recommendation Commission* (2008).

CDL has been serving as the Permanent Secretariat of Linguistic Society of Nepal. It brings regular publications of periodicals, newsletter, journals, dictionaries and grammars.

It is also equipped with a computer laboratory supported by recent language softwares.

CDL has a well-furnished office space and classrooms for administration and teaching and research.

### **13 Additional sources of funding**

LinSuN is commissioned by the National Planning Commission (NPC) of the Government of Nepal (GoN). As such, GoN will take responsibility for partial funding of this project. This amount is yet to be determined, but it is anticipated that the amount that will be contributed will be about thirty percent of the overall budget. This leaves the remaining amount to be sought from other sources.

We believe that there will be other interested stake holders in funding this survey. The survey addresses many of the issues which the development world seeks to address. Three issues come to mind immediately: Social Inclusion, Education for All, and the Documentation of Languages.

This survey will speak to the issue of Social Inclusion in that identification of linguistic variation is the foundation from which linguistic inclusion can happen. It is clear that inclusion is not possible until the extent of variation is known. This survey will speak to the issue of Education for All in that it seeks to develop writing systems and develop written materials in at least one language in each cluster. The development of written materials in the mother tongue is one of the foundations for MLE programs. This survey speaks to the issue of Documentation in that it begins to document some languages, and develops a database in which each of the languages of Nepal will be included.

There are other issues that are of benefit as well. One is the idea of capacity building and knowledge development. This survey clearly builds capacity among Nepalese University students, training and giving employment to many young scholars. In addition, it builds capacity in the language community to take ownership of its linguistic destiny.

In summary, this survey addresses many of the major issues of interest to the development world, and as such, we would suggest it qualifies for funding from a number of different sources.

It may not be possible to get funding for the Survey in its entirety from a single donor, in which case different parts of the survey could be divided up for funding from different sources. The preference, though, would be for it to get funding from one source (in addition to NPC).

At the least, the survey team needs to begin exploring funding sources from donors. To do this, it will make contact with its already established networks to begin to determine interest. Some possible donors are the European Union, World Bank, Asian Development Bank, and certain UN agencies. Some foreign university professors may also be interested in helping to find funding for certain parts, or in sending students from their departments to do research.

At the very outset we have already received quite a few encouraging responses from foreign academic institutions. Puma and Chintang documentation has been continuing as a part of LinSuN with the financial assistance of the Volkswagen Foundation in Germany. Endangered Languages Documentation Project at University of London has started assisting the Linguistic and Ethnographic Documentation of Baram Language since 2007.

SIL Intl have contributed seed money to help in some early stages of the proposal writing, as well sponsoring three students from CDL to attend InField at UCSB. They will continue to contribute small amounts (in comparison to the overall budget) to certain parts of

the survey. Steve Watters and Carl Grove have also been active participants of the proposal writing phase.

Carol Genetti of University of California at Santa Barabara has indicated a desire to help in applying for funds for training from the US National Science Foundation.

Similarly, scholars from many Universities have been involved in collaborative research in Nepalese languages with CDL, and they have been encouraging in their support of this project.

Finally, the National Foundation for Development of Indigenous Nationalities (NFDIN) has been working in collaboration with CDL in developing dictionaries and grammars of marginalized languages of Nepal, and the survey looks forward to its continued cooperation with them.

## 14 BUDGET

The overall budget for the seven year span is NRS 44,58,49,200 (forty-four crore, fifty eight lakh, forty nine thousand, two hundred, only). The budget for each phase is as follows:

phase 1: 11,84,46,000  
 phase 2a: 19,48,55,000  
 phase 2b: 13,25,48,200

The details for the budget can be found in the following pages. However, for clarity sake, we have included the pay scales and number of employees below to give a clear picture of the scale of the project. In total, the project will require 232 persons of which 200 people will be in research. For reference to where the different positions fit into the project administrative structure, please the chart on page X. All figures are in NRS.

**Senior Management Committee:** 8 Persons @ 30,000.00/month

**Director:** 1 @ 65,000.00/month

Director's Personal Assistant: 1 @ 25,000.00/month

**Middle Management:** 5 Persons @ 30,000.00/month

Operations Manager, IT Team Leader/Archiving Coordinator, Sociolinguistics  
 Coordinator, Documentation Coordinator, Literacy Coordinator

**Consultants:** @ 40,000/month

See budget for number of consultants that can be hired for each phase

This is intended for Nepalese consultants. No provision for expatriate consultants is included in the budget.

**Administrative Staff:** 11 Persons

Operations Manager (included above under middle management)

Finance Manager : 1 Person @ 30000/month

Book Keeper: 2 Persons @ 20000/month

Logistical Support: 1 Person @ 20000/month

Staff : 6 Persons @ 10,000/month

Research Teams (vary in number according to Phase): max is 114 Persons, 3 Persons on a sociolinguistic team and 4 persons on a literacy and documentation team.

Team Leaders: @ 25,000/month

Research Assistant: @ 20,000/month

Field Allowance: Team Leaders and Research Assistants: @ 7,000/month

**IT Department:** 6 Persons

IT Team Leader / Archiving Coordinator (included above under middle management)

Database Designer: 1 Persons @ 25000/month

Database Programmer: 2 Persons @ 25000/month

Computer Support: 2 Persons @ 20000/month

FINANCIAL DETAILS FOR PHASE I (Year 1 and 2)				
	Description	NRS (per month)	Units (2 years)	Total
Salaries	Senior Management Committee (30,000/8 persons)	240000	24	5760000
	Director + PA	90000	24	2160000
	Middle Management (30,000/5 persons)	150000	24	3600000
	Consultants (40,000/month)	40000	10	400000
	Team Leaders (25,000/41 persons)	1025000	24	24600000
	Administrative Staff (30,000 + (25,000*2)+20,000)	100000	24	2400000
	IT Department (30,000+(25,000*3)+(20,000*3)	145000	24	3480000
	Research Assistant (20,000/123 persons)	2460000	24	59040000
	Staff (10,000/6 persons)	60000	24	1440000
	CDL Institutional Overhead	431000	24	10344000
	Health Insurance	50000	2	100000
<b>Sub Total (A)</b>				<b>113324000</b>
Infrastructure (Office and Logistics)				
	Kathmandu Office	50000	24	1200000
	Utilities	15000	24	360000
	Internet / Communication	20000	24	480000
	Operations	15000	24	360000
	One Time Setup	120000	1	120000
	Field Offices (10,000/9)	90000	24	2160000
	Field Office Operations (2000/month/9 offices)	18000	24	432000
<b>Sub Total (B)</b>				<b>5112000</b>
Travel (for Sociolinguistics, Documentation, and Literacy)				
	Sociolinguistic Teams (30,000/25 teams)	750000	24	18000000
	Documentation Teams (20,000/6 teams)	120000	24	2880000
	Literacy Teams (20,000/10 teams)	200000	6	1200000
	Field Allowance (7,000/164 persons)	1148000	24	27552000
	Language Consultant Salaries	160000	24	3840000

	Conference Presentations South Asia (10,000/10 persons)	100000	1	100000
	Conference Presentations International (120,000/2 persons)	240000	1	240000
	Travel for International Consultants (120,000/4 persons)	480000	1	480000
	InField at UCSB(3 Persons for varying costs)	900000	1	900000
<b>Sub Total ( C )</b>				<b>55192000</b>
<b>Equipment</b>	Car	2500000	2	5000000
	Car Insurance (5000/2 cars)	10000	24	240000
	Field Equipment (120000/41 teams)	4920000	1	4920000
	Office Computers	50000	4	200000
	Photocopy	50000	1	50000
	IT Computers	50000	7	350000
	Networking Infrastructure	120000	1	120000
<b>Sub Total (D)</b>				<b>10880000</b>
<b>Publishing</b>	Text Collection of Kusunda	100000	1	100000
	Annotated Bibliography of Documentation	100000	1	100000
	Typological Overview	200000	1	200000
	Literacy in Use Survey	100000	1	100000
	Annotated Bibliography of Literacy Materials	100000	1	100000
<b>Sub Total (E)</b>				<b>600000</b>
<b>Grand Total (A+B+C+D+E)</b>				<b>185108000</b>
<b>In Words</b>	eighteen crore, fifty one lakh, eight thousand			

FINANCIAL DETAILS FOR PHASE IIa (Year 3, 4, 5)				
	Description	NRS (per month)	Units (3 years)	Total
Salaries	Senior Management Committee (30,000/8 persons)	240000	36	8640000
	Director + PA	90000	36	3240000
	Middle Management (30,000/5 persons)	150000	36	5400000
	Consultants (40,000/month)	40000	15	600000
	Team Leaders (25,000/50 persons)	1250000	36	45000000
	Administrative Staff (30,000 + (25,000*2)+20,000)	100000	36	3600000
	IT Department (30,000+(25,000*3)+(20,000*3)	145000	36	5220000
	Research Assistant (20,000/150 persons)	3000000	36	108000000
	Staff (10,000/6 persons)	60000	36	2160000
	CDL Institutional Overhead	507500	24	12180000
	Health Insurance	50000	3	150000
<b>Sub Total (A)</b>				<b>194190000</b>
Infrastructure (Office and Logistics)	Kathmandu Office	50000	36	1800000
	Utilities	15000	36	540000
	Internet / Communication	20000	36	720000
	Operations	15000	36	540000
	Field Offices (10,000/9)	90000	36	3240000
	Field Office Operations (2000/month/9 offices)	18000	36	648000
<b>Sub Total (B)</b>				<b>7488000</b>
Travel (for Sociolinguistics, Documentation, and Literacy)	Sociolinguistic Teams (30,000/25 teams)	750000	36	27000000
	Documentation Teams (20,000/15 teams)	300000	36	10800000
	Literacy Teams (20,000/10 teams)	200000	36	7200000
	Field Allowance (7,000/200 persons)	1400000	36	50400000
	Language Consultant Salaries	160000	36	5760000

	Conference Presentations South Asia (10,000/10 persons)	100000	3	300000
	Conference Presentations International (120,000/2 persons)	240000	3	720000
	Travel for International Consultants (120,000/4 persons)	480000	2	960000
	InField at UCSB(3 Persons for varying costs)	900000	1	900000
<b>Sub Total ( C )</b>				<b>104040000</b>
<b>Equipment</b>	<b>Car Insurance (5000/2 cars)</b>	<b>10000</b>	<b>24</b>	<b>240000</b>
	Field Equipment (120000/9 teams)	1080000	1	1080000
	IT Computers	50000	4	200000
	Networking Infrastructure	120000	1	120000
<b>Sub Total (D)</b>				<b>1640000</b>
<b>Publishing</b>	<b>Sociolinguistic Findings</b>	<b>200000</b>	<b>1</b>	<b>200000</b>
	Sketch Grammars	200000	2	400000
	Glossary / Dictionaries	100000	3	300000
	Literacy Materials	150000	3	450000
<b>Sub Total (E)</b>				<b>1350000</b>
<b>Grand Total (A+B+C+D+E)</b>				<b>308708000</b>
<b>In Words</b>	<b>thirty crore, eighty seven lakh, eight thousand</b>			

FINANCIAL DETAILS FOR PHASE IIb (Year 6 and 7)				
(Increased by 10% to cover inflation)				
	Description	NRS (per month)	Units (2 years)	Total
Salaries	Senior Management Committee (30,000/8 persons)	264000	24	6336000
	Director + PA	99000	24	2376000
	Middle Management (30,000/5 persons)	165000	24	3960000
	Consultants (40,000/month)	44000	10	440000
	Team Leaders (25,000/50 persons)	1375000	24	33000000
	Administrative Staff (30,000 + (25,000*2)+20,000)	110000	24	2640000
	IT Department (30,000+(25,000*3)+(20,000*3)	159500	24	3828000
	Research Assistant (20,000/150 persons)	3300000	24	79200000
	Staff (10,000/6 persons)	66000	24	1584000
	CDL Institutional Overhead	558250	24	13398000
	Health Insurance	50000	2	100000
<b>Sub Total (A)</b>				<b>146862000</b>
Infrastructure (Office and Logistics)	Kathmandu Office	55000	24	1320000
	Utilities	16500	24	396000
	Internet / Communication	22000	24	528000
	Operations	16500	24	396000
	Field Offices (10,000/9)	99000	24	2376000
	Field Office Operations (2000/month/9 offices)	19800	24	475200
<b>Sub Total (B)</b>				<b>5491200</b>
Travel (for Sociolinguistics, Documentation, and Literacy)	Sociolinguistic Teams (30,000/25 teams)	825000	24	19800000
	Documentation Teams (20,000/15 teams)	330000	24	7920000
	Literacy Teams (20,000/10 teams)	220000	6	1320000
	Field Allowance (7,000/200 persons)	1540000	24	36960000

	Language Consultant Salaries	176000	24	4224000
	Conference Presentations South Asia (10,000/10 persons)	110000	1	110000
	Conference Presentations International (120,000/2 persons)	264000	1	264000
	Travel for International Consultants (120,000/4 persons)	528000	1	528000
	<b>Sub Total ( C )</b>			<b>71126000</b>
<b>Equipment</b>	Car Insurance (5000/2 cars)	11000	24	264000
	Replacing Field Equipment (120000/5 teams)	600000	1	600000
	Replacing Office Computers	55000	2	110000
	Replacing IT Computers	55000	3	165000
	<b>Sub Total (D)</b>			<b>1139000</b>
<b>Publishing</b>	Sociolinguistic Findings	225000	1	225000
	Sketch Grammars	225000	2	450000
	Glossary / Dictionaries	150000	3	450000
	Literacy Materials	200000	3	600000
	<b>Sub Total (E)</b>			<b>1725000</b>
	<b>Grand Total (A+B+C+D+E)</b>			<b>226343200</b>
<b>In Words</b>	twenty two crore, sixty three lakh, forty three thousand two hundred			

## REFERENCES

- Acharya, Bhagavat (1997). *Dialects of Nepali Based on Pronunciation [in Nepali]*. MA dissertation, Central Department of Nepali, TU.
- Bandhu, Chura Mani (1968). *Report of the preliminary linguistic survey in Bagmati and Janakpur zones of Nepal [in Nepali]*. Kathmandu: TU, CNAS.
- Bielmeier, Roland (To Appear). *Historical-Comparative Lexicon of Tibetan Dialects*. Berne: University of Berne.
- Blair, Frank. 1990. *Survey on a Shoestring: A Manual for Small-Scale Language Surveys*. SIL and The University of Texas at Arlington.
- Casad, Eugene. 1974. *Dialect Intelligibility Testing*. Norman: SIL International.
- CBS. 1911. *Population Census*. Kathmandu.
- CBS. 1952/54. *Population Census*. Kathmandu.
- CBS. 1961. *Population Census*. Kathmandu.
- CBS. 1971. *Population Census*. Kathmandu.
- CBS. 1981. *Population Census*. Kathmandu.
- CBS. 1991. *Population Census*. Kathmandu.
- CBS. 2001. *Population Census*. Kathmandu.
- Cho, Kwang-ju, John William Epele and Padam Kumar Rai (2003). *Bantawa: A sociolinguistic survey* (mimeo).
- Dhimal, Chandra; Holly J Larsen and Rebecca G Williams (2001). *The Dhimal: A linguistic Survey* (mimeo).
- Dhungana, Lavany Prasad (1997). *Nepali Dialects Based on Basic Vocabulary [in Nepali]*. MA dissertation, Central Department of Nepali, TU.
- Epele, John; Jonathan D Grimes and Tanka Rajbamshi (2001). *The Rajbamshi of Nepal: A linguistic study* (mimeo).
- Glover, Warren W (1971). "Swadesh list calculations on thirty Tibeto-Burman languages". Kathmandu: SIL (mimeo).
- Glover, Warren W and John K Landon (1980). "Gurung dialects". In: S A Wurm (ed). *Papers in South-East Asian Linguistics No. 7, Pacific Linguistics series A, No. 53*. Canberra: Australian National University, 29-77.
- Glover, Warren W and Ratna Bahadur Gurung (2003). *Gurung-Nepali-English Dictionary*. Kathmandu: Tamu Bauddha Seba Samiti.
- GoN 2063 VS (2007 AD). *Interim Constitution of Nepal*. Kathmandu.
- GoN 2063 VS (2007AD). *National Curriculum Framework for School Education in Nepal*. Kathmandu.
- Gordon, Kent 1976. *The Phonology of Dhangar-Kurux*. Dallas: SIL International.
- Gordon, Raymond G. Jr., (ed.). 2005. *Ethnologue: Languages of the World, 15<sup>th</sup> edition*. Dallas, Texas: SIL International.
- Government of Nepal. *Interim Constitution of Nepal, 2007*. Kathmandu.

- Government of Nepal. *Interim Plan, 2007*. National Planning Commission, Kathmandu.
- Grierson, GA (1927). *Linguistic Survey of India* volume 1 part 1. Delhi: Motilal Banarasidass [Reprint 1967].
- Grierson, George and Sten Konow. 1898-1927. *Linguistic Survey of India*, 19 volumes, Calcutta: Superintendent of Government Printing, India. Reprinted 1967. Delhi: Motilal Banarasidas.
- Hale, Austin (1993). Sociolinguistic Survey of Nepal. A proposal for a joint project to be conducted under the auspices of the Centre for Nepal and Asian Studies with the sponsorship of the Summer Institute of Linguistics. Draft Three, 15 April 1993.
- Hansson, Gerd (1991). *The Rai of Eastern Nepal: Ethnic and Linguistic Grouping. Findings of the Linguistic Survey of Nepal*. Kathmandu: Linguistic Survey of Nepal and CNAS.
- HMG/N 2047 VS (1990 AD). *Constitution of the Kingdom of Nepal*. VS 2047 (1990 AD). Kathmandu: HMG. English translation of Nepali document available online at <http://www.nepalhmg.gov.np/constitution.html>
- His Majesty's Government of /Nepal's National Planning Commission Secretariat. 2003. *Population Monograph of Nepal, volumes I and II*. Kathmandu: Central Bureau of Statistics.
- HMG/N. *Copyright Regulations*. 2006. Kathmandu.
- HMG/N. Rastriya Bhasha Niti Sujhav Aayogko Prativedan (The Report of National Languages Policy Recommendation Commission), 2050 VS (1994 AD) (English version in 2008 published by Central Department of Linguistics). Kathmandu.
- HMG/N. *The Copyright Act*. 2003. Kathmandu.
- HMG/N. *The Local Self-Governance Act, 1999*. Kathmandu.
- Hodgson, B H (1834). Classification of Newars or aborigines of Nepal proper, preceded by the most authoritative legend relative to the origin and early history of the race. *Journal of the Asiatic Society of Bengal* 3.29:215-21.
- Hodgson, B H (1847-48). On the aborigines of the sub-Himalayas. *Journal of the Asiatic Society of Bengal* 16:1235-44, 17:73-78.
- Hodgson, B H (1857-8). Bahing vocabulary. *Journal of the Asiatic Society of Bengal* 26.5:486-522, 27.5:393-442.
- Hodgson, B H (1857a). Comparative grammar of the languages of the broken tribes of Nepal. *Journal of the Asiatic Society of Bengal* 26.5:317-332.
- Hodgson, B H (1857b). "Comparative vocabulary of the several languages (dialects) of celebrated people called Kirantis, now occupying the easternmost province of the kingdom of Nepal, or Basin of River Arun". *Journal of the Asiatic Society of Bengal* 26.5:333-71.
- Hodgson, B H (1857c). Vayu vocabulary. *Journal of the Asiatic Society of Bengal* 26.5-6:372-485.
- Hodgson, B H (1880). "Vocabulary of the dialects of the Kiranti languages". *Miscellaneous Essays Relating to Indian Subjects*. Volume 1: London, pp 178-215, Section II: On Himalayan ethnology.

- Hoernle, A F Rudolf (1880). *A Comparative Grammar of the Gaudian Languages with Special Reference to the Eastern Hindi*. London: Rubner.
- Lee, Maureen B., Bagdevi Rai, Buddhi Kumar Rai and Carin Boone (2005). *Bayung Rai: A Sociolinguistic Survey*. Kathmandu: TU, CNAS.
- Malla, K. P. 1989. 'Language and Society in Nepal' in Malla, ed., *Nepalese Perspectives on Continuity and Change*. Kathmandu: CNAS.
- Michailovsly, Boyd (1988). Phonological typology of Nepal languages. *Linguistics of the Tibeto-Burman Area* 11.2: 25-50.
- Niraula, Yagyanshwar (1993). *Dialects of Nepali Based on Verb Forms* [in Nepali]. MA dissertation, Central Department of Nepali, TU.
- Noonan, M. 2005. *Language Documentation and Language Endangerment*, University of Iceland.
- Oakes, Leigh (2001) *Language and national identity. Comparing France and Sweden*, Amsterdam: John Benjamins.
- O'Leary, Clare F. (ed.). 1992. Sociolinguistic Survey of Northern Pakistan, Volume II in *Languages of Northern Areas*. NIPS and SIL.
- Pokharel, Bal Krishna (2022 VS). *Rastra bhasa* [National language]. Kathmandu: Sajha Prakashan.
- Pokharel, Bal Krishna (mimeo). *Nepali bhaasaa ko naapi* [A survey of Nepali language].
- Pradhan, Rajendra and Ava Shrestha 2005. *Ethnic and Caste Diversity: Implications for Development*. Kathmandu: Asian Development Bank.
- Radolf, Carla. 1989. *Sentence Repetition for Studies of Community Bilingualism*. Prepublication draft, unpublished manuscript.
- Simons, Gary F. 1984. *Word List Analysis in the Field with a Notebook Computer*. Occasional papers in Academic Computing, 5. Dallas: SIL International.
- Subedi, Sakhi Sharan (1994). *A Survey of Nepali Dialects Based on Syntactic Variation* [in Nepali]. MA dissertation, Central Department of Nepali, TU.
- Timilsina, Yamanath (1993). *Dialects of Nepali Based on Numerals and Pronominals* [in Nepali]. MA dissertation, Central Department of Nepali, TU.
- Toba, Sueyoshi (1973). *Preliminary exploratory language survey report of some Rai languages (dialects)* [mimeo]
- \_\_\_\_\_ (1991). *A Bibliography of Nepalese Languages and Linguistics*. Kathmandu: Linguistic Society of Nepal (2nd edition. 1998. Kathmandu: Central Department of Linguistics, TU).
- Toba, Sueyoshi, Ingrid Toba and Novel Kishore Rai (2002). *UNESCO Language Survey Report*. Kathmandu: UNESCO.
- Tuladhar, Nirmal M. 1999, "Sociolinguistic Survey in the Jirel Community." *Gipan*, 1:2 (The Journal of Central Department of Linguistics, Tribhuvan University ).
- UNESCO. 2006. *Advocacy Kit for Multilingual Education*. Bangkok. (Adapted and translated into Nepali in 2008)

- Varenkamp, Bryan K (1996). *Tamang Tam*. A Sociolinguistic Study of Eastern Tamang Dialects (in Nepal). Kathmandu: TU, CNAS.
- Watters, David and YP Yadava (2003). An Encyclopedia of Nepal's Languages: a proposed format. ms.
- Watters, Steve (1988). A sociolinguistic profile of the Kham Magar. ms.
- \_\_\_\_\_ (1998). A Sociolinguistic Profile of Mugu. Paper presented at the XVIII Annual Conference of the Linguistic Society of Nepal, 26-27 November 1998.
- Webster, Jeff (1992). The sociolinguistic survey of Tibeto-Burman dialects: North Gorkha District (mimeo).
- Webster, Jeff (2001). "The sociolinguistic study of Limbu". *Journal of Nepalese Literature, Art and Culture*. Kathmandu: Royal Nepal Academy, 4.1:51-84.
- Winter, Werner (1984). Distribution of languages and sublanguages in Eastern Nepal. TU CNAS (mimeo)
- Winter, Werner (1986). "The linguistic survey of Nepal". In: Bernahard Kolver: Aspects of Nepalese Traditions, 171-179. Stuttgart: Franz Steiner.
- Yadava, Yogendra P. 2002. A Sketch Grammar of Dhangar. ms.
- Yadava, Yogendra P. 2003. Language. *Population Monograph*, volume I. Kathmandu: CBS.
- Yadava, Yogendra P. and Turin, Mark. 2005. Indigenous Languages of Nepal: A Critical Analysis of the Linguistic Situation and Contemporary Issues in *The Indigenous Languages of Nepal (ILN): Situation, Policy Planning and Coordination*. Lalitpur: NFDIN.