

In Preparation

A Rough Handbook for Field Work On Endangered Languages

**Udaya Narayana Singh
Gulab Chand**

**Amity Centre for Linguistic Studies [ACLiS],
Amity University Haryana
Dt. Gurgaon, Haryana (India)**

2021

@Mediating Multilingualism Project

Contents

1. Endangered Language : An Overview

- 1.1. Introducing Language Endangerment
- 1.2. Scales of Endangerment
- 1.3. Why Endangered Languages
- 1.4. The Tensions
- 1.5. The Plan of Action as Conceptualized
- 1.6. The Desired Action
- 1.7. Training of Teachers in Schools Using These Languages
- 1.8. Study of Endangered Languages: Aims and Objectives

2. Field Work : A Short Introduction

- 2.1. Aims and Objectives
- 2.2. The Issues

3. Mapping of the Field

- 3.1. The Tasks at Hand
- 3.2. Language Family Maps

4. Socio-linguistic profile of Informants

- 3.1. Domains of Use
- 3.2. Language Attitude
- 3.3. Language Prestige

5. Selection and Training of Enumerators

6. Template for Capturing Basic Data of Suveyors/Enumerators

7. Criteria of Informants

- 7.1. Age
- 7.2. Region
- 7.3. General Qualifications
- 7.4. Avoid
- 7.5. Good Informant
- 7.6. Personal Profile of the Informant
- 7.7. Community Name
- 7.8. Language Identification

8. Word and Sentence list

9. Specifications of Tools

10. Storage and Reporting of Output

References

Appendix

Survey Template for Presentation of Reports on Individual Languages

1. Endangered Languages : An Overview

1.1 Introducing Language Endangerment

One still lacks clarity as to what is exactly meant by ‘Endangerment’ of languages, and what its generally recognized levels are. The best yardstick for this is the behaviour of a language in various generations of a speech community, especially that of children. If it is the case that not more than 10–30 per cent of the children of a given speech community are able to speak what is supposed to be their “mother-tongue” or language of their tradition and culture, the language is potentially endangered. If there are only few speakers left in a community (with very few among whom are children), and if the youngest good speakers are young adults, the language is endangered. If their youngest good speakers are largely past middle age, the language is seriously endangered. If only a handful of mostly old speakers are left who can speak the language, the rest having shifted to a neighbouring tongue, the language is moribund. And if no speakers seem to be left, the language is believed to be extinct. Thus, language endangerment is a matter of scale, rather than belonging to an all-or-none category. Further, different authors and surveys categorize these stages slightly differently. Two of the most popular and widely referred surveys/resources are *Atlas of the World's Languages in Danger*¹ and the *Ethnologue*².

1.2 Scales of Endangerment

UNESCO adopts a 5-point scale (six degrees) of language endangerment based on “intergenerational transmission” on the recommendations made by an ad hoc team of linguists commissioned by UNESCO³:

Safe (Level 5): language spoken by all generations—uninterrupted

¹ Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*. 3rd Edition. Paris: UNESCO Publishing. ISBN 978-92—3-104095-5.

² Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the World*, Eighteenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.

³ UNESCO. 2003. *Language Vitality and Endangerment*. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Paris, 10-12 March 2003. <http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf> (accessed 11 February 2016)

intergenerational transmission.

Stable yet threatened (Level 5-): language spoken in most contexts by all generations with unbroken intergenerational transmission, but usurpation of certain important communication contexts by one or more dominant language(s).

Vulnerable (Level 4): language spoken by most but not all children or families of a particular community as their first language, but restricted to specific social domains (such as at home).

Definitely endangered (Level 3): language no longer learned as mother tongue by children at home—parental generation being the youngest speakers.

Severely endangered (Level 2): language spoken only by grandparents and older generations—no use of that particular language among parental generation.

Critically endangered (Level 1): language spoken by great-grandparental generation and not used for everyday interactions.

Extinct (Level 0): no one to speak or remember the language.

Prior to the UNESCO scale, an Eight-Point Framework was proposed by Joshua Fishman⁴ in 1991 on the “key role of intergenerational transmission in the maintenance of a language”⁵. This was later expanded by Lewis and Simons (2010) and adopted by *Ethnologue* for their description of language vitality. Following is a gist of Fishman’s description of language vitalization, known as “Graded Intergenerational Disruption Scale” (GIDS)⁶—

GIDS LEVEL	DESCRIPTION
---------------	-------------

⁴ Fishman, J. A., 1991, Reversing language shift, Clevedon, UK, Multilingual Matters Ltd.

⁵ Lewis, M. P. and G. F. Simons. 2010. Assessing endangerment: Expanding Fishman’s GIDS. *Revue Roumaine de Linguistique* 55(2):103–120. http://www.lingv.ro/RRL_2_2010_art01Lewis.pdf. Accessed 11 February 2016.

⁶ Reproduced from Lewis & Simons (2010).

1	The language is used in education, work, mass media, government at the nationwide level
2	The language is used for local and regional mass media and governmental services
3	The language is used for local and regional work by both insiders and outsiders
4	Literacy in the language is transmitted through education
5	The language is used orally by all generations and is effectively used in written form throughout the community
6	The language is used orally by all generations and is being learned by children as their first language
7	The child-bearing generation knows the language well enough to use it with their elders but is not transmitting it to their children
8	The only remaining speakers of the language are members of the grandparent generation

Fig. 1: Graded Intergenerational Disruption Scale

*Ethnologue*⁷, on the other hand, categorized a five-level scale (Fig. 2) for assessing language vitality—the main parameter for such categorization was the number of first language speakers. Consider this five-layered approach:

Category	Description
Living	Significant population of first-language speakers
Second Language Only	Used as second-language only. No first-language users, but may include emerging users
Nearly Extinct	Fewer than 50 speakers or a very small and decreasing fraction of an ethnic population
Dormant	No known remaining speakers, but a population links its ethnic identity to the language
Extinct	No remaining speakers and no population links its ethnic identity to the language

Fig. 2 Ethnologue: Language Vitality Scales

None of these three frameworks can adequately grasp the extent of language endan-

⁷ Gordon (2005), Grimes (2000) & Lewis (2009)

germent in its entirety. Lewis and Simons (2010) presented a critique of these three frameworks pointing to the inadequacies of all the three proposals. None of the three frameworks provides a sufficient differentiation between languages which are at a higher risk of endangerment. Hence Lewis and Simons (2010) proposed an Extended Graded Intergenerational Disruption Scale (EGIDS) harmonizing all the three schemes. The EGIDS has 13 levels, and a comprehensive list of EGIDS may be reproduced here from Lewis and Simons (2010)—

Expanded Graded Intergenerational Disruption Scale			
LEVEL	LABEL	DESCRIPTION	UNESCO
0	International	The language is used internationally for a broad range of functions.	Safe
1	National	The language is used in education, work, mass media, government at the nationwide level.	Safe
2	Regional	The language is used for local and regional mass media and governmental services.	Safe
3	Trade	The language is used for local and regional work by both insiders and outsiders.	Safe
4	Educational	Literacy in the language is being transmitted through a system of public education.	Safe
5	Written	The language is used orally by all generations and is effectively used in written form in parts of the community.	Safe
6a	Vigorous	The language is used orally by all generations and is being learned by children as their first language.	Safe
6b	Threatened	The language is used orally by all generations but only some of the child-bearing generation are transmitting it to their children.	Vulnerable
7	Shifting	The child-bearing generation knows the language well enough to use it among themselves but none are transmitting it to their children	Definitely Endangered
8a	Moribund	The only remaining active speakers of the language are members of the grandparent generation.	Severely Endangered
8b	Nearly Extinct	The only remaining speakers of the language are members of the grandparent generation or older who have little opportunity to use the language.	Critically Endangered
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community. No one has more than symbolic proficiency.	Extinct

10	Extinct	No one retains a sense of ethnic identity associated with the language, even for symbolic purposes.	Extinct
----	---------	---	---------

Fig 3: EGIDS Scales

1.3 Why Endangered Languages?

The operational definitions underscore the need for empirical studies on language use patterns in multilingual settings before categorizing languages as endangered. There are around 6000 languages spoken in the world today. There is a widespread agreement that by the end of this century, about 90% of world's languages will disappear or become extinct entirely, replaced by more widely used and economically "more viable" and "more productive" languages. This situation is generally referred to as language endangerment, a term used broadly for languages, which are threatened with absolute loss. A language is also considered endangered when it is used by few speakers and/or it is used in fewer situations or domains.

Nancy Dorian (1998: 3) opened her essay with the following observation, which is valuable in the context of endangered languages: "It might be said with certain metaphoric license that languages are seldom admired to death but are frequently despised to death."⁸ This was later elaborated in the essay, which illumines the process in which some languages slowly move towards critical positions. She says it is rare to expect a human language to achieve a position where the ordinary people will be in awe to speak it, or claim to speak it, and distance themselves from it. The language custodians preventing the ordinary people from using it is a rare thing. On the contrary, it is often seen that a language has become so exclusively associated with the masses of low-prestige people that the gentry and the upper class as well as potentially creative people begin to distance themselves from it and adopt another, a more prestigious language, for both creativity and identity. If this gains ground, it soon becomes a trend among the new generation of potential speakers looking for prestige markers. Of course, she does mention instances such as Sanskrit, Latin, or Coptic Egyptian that become so fossilized and specialized in use that the speaker bases are depleted. But where the work certainly made its mark was in mentioning how Nahuatl and Quechua – two major languages of the erstwhile great empires in the North America - the 'Aztec's, and South America – the 'Inka's – quickly became low-prestige speech forms as the better-armed Old World Spanish-people entered the region to dominate and subjugate them. In many instances, cultural erosion begins in this manner as is described in detail by Nicholas Ostler's (2005) *Empires of the Word: a language history of the world* (Harper-Collins).

⁸ Dorian, Nancy. 1998. 'Western language ideologies small-language prospects' in Grenoble, L.A & L.J. Whaley, eds. *Endangered Languages: Current Issues and Future Prospects*. Cambridge: Cambridge University Press.

With Language Endangerment, the possibility of Cultural Erosion is enhanced. One could of course partially revive the use of an extinct language for special purposes, which would be like re-enacting great Sanskrit plays in modern-day context relying heavily on the lexical commonalities that would obtain between the language of the play and the languages usually spoken by the audience that watch such performances. Using Biblical Hebrew or Latin for a religious-cultural event would be similar to that situation. However, for many in the positions of power, ecological and cultural diversity are of no importance. Therefore, no attempts are made or contemplated to retain this diversity. Those who wield power of governance and those who exert influence on decision-making systems in today's world often take 'Cultural diversity' to be a necessary and unavoidable evil. They often choose to overlook the fact that such diversities offer a rich tapestry of flora, fauna and cultural products that make the life worth living. Linguistic diversity or plurality – which is a part of this rich Cultural diversity, matches India's bio-diversity.

We look at the endangered species index in terms of a bio-diversity map, the threat to multilingualism is similar to the threat to biodiversity, not just because most languages are like disappearing "species", but also because there is an intrinsic and causal link between biological diversity and cultural diversity. Like plant and animal species, endangered languages are confined to small areas. As we know, more than 80% of countries with great biological diversity are also the places with the greatest number of endangered languages. This is because when people adapt to their environment, they create a special stock of knowledge about it, which is mirrored in their language, and is available only in such languages. Many of the world's endangered plant and animal species today are known only to certain people whose languages are dying out. As they die, they also take with them all the traditional knowledge about the environment.

There is no doubt that like natural disappearance of languages and civilizations, extinction of living organisms is a natural event happening in a routine manner. But that rate is not alarming as the average rate of extinction in the last 200 million years (MY) is 1-2 species per year, or 3-4 families per MY. There had been of course periods during which a large number of species have gone extinct in the same blink of geological time, but nothing is comparable to what has happened in the modern times, when the ecosystems are threatened with destruction due to human actions. Because of a massive growth of human population, roughly half of the world's forests have been transformed, degraded or destroyed by man, and almost half of the world's net primary productivity is appropriated for human use with most of available fresh water being disturbed by us, and virtually all of the available productivity of the oceans is affected. Over-hunting, and especially commercial hunting, or pollution due to chemical waste, and invasion of non-native space are some reasons but as Wilson (1992) had suggested⁹, among all human actions, perhaps habitat loss presents the

⁹ Wilson, E.O. 1992. *The Diversity of Life*. W.W. Norton & Co., New York.

single greatest threat to world biodiversity. Just to take only one example of ‘tropical forests’, according to an estimate by the University of Michigan - by “simply using the most conservative values of the slope, and assuming the true biodiversity of tropical forests is roughly 10 million species, the projected rate of loss of species is 27,000 per year, and three during this hour”¹⁰. Interestingly, like in the case of languages or mother-tongues, in the animal world too, the number of unknown species exceeds the count of the number of known species (about 1.4 million) that they have been classified by specialists. If take a serious look at the statistical details on languages in India as per Census 2001 records, we find similarly a large number of mother-tongues that have not yet been identified or classified. Consider the following Census table of returns of the 2001 Census:

S. No.	Item	Number
1.	Number of ‘Raw returns’ of Mother Tongues received from enumerators	6,661
2.	Number of Rationalized groups of Mother Tongues	3,592
3.	Classified Mother Tongues (out of 3,592)	1,635
4.	Unclassified Mother Tongues (out of 3,592)	1,957
5.	Identifiable Mother Tongues (that were returned by more than 10,000 speakers)	234
6.	Mother Tongues returned by 10,000 people but not published due to 1976 CCPA restrictions	68
7.	Clearly identified and Identifiable Mother Tongues (a total of 5 and 6 above)	302

Fig 4: Census Table of Returns (2001)

The Languages and Mother Tongues are usually presented by the Registrar General of India office by grouping them under 122 relevant languages (22 Part A languages and 100 Part B languages) in 2001 Census, as against 114 languages (and 216 Mother Tongues) as reported in 1991 Census. While the 22 Languages of the Eighth Schedule (listed in Part A of Census 2001) have several advantages in terms of recognition and support from State and Central Governments, the other languages and Mother Tongues are not so fortunate in this respect. While it is true that some would “die natural death” because of lack of speakers, or because of various factors that make the domains of use of these languages shrink, many of these languages and speech varieties face the danger of dying out due to lack of State support to protect and promote them. It is a matter of historical accident that some cultures and communities have given rise to authors, grammarians and language enthusiasts that have ensured early development of their languages. ‘Planned Language Development’ is viewed as a systematic by the state or such other authoritative bodies that decide to allocate resources to enhance the status of a language and expand its domains of use. When such investments are made, the yields could be seen in terms of sociolinguistic development (including literacy, literary and educational progress) of small and marginalized communities which have been at the fringes for long, losing out on the benefits of economic and social development of the country. In pluralistic society like ours, multilingual development is ab-

¹⁰ <http://www.globalchange.umich.edu/globalchange2/current/lectures/biodiversity/biodiversity.html>

solutely essential to take all sections of our people together. Linguistic minorities cannot be deprived of their rightful support from the state in this respect. Language planning is thus a key input to educational planning in particular and social engineering in general.

1.4 The Tensions:

One must sensitize all those well-meaning scholars and activists about the kind of tensions involved in working on these areas:

First, as we engage ourselves in the documentation, protection and promotion of endangered languages and their culture, we must understand that there will always be agencies that would like to color our perception and blur our vision. This is because any advocacy in favor of diversity may cause immense harm to several agencies that may have a stake in continuance of only a few languages as the media of trade, commerce, tourism, education and entertainment.

There is thus a constant tension between ‘being’ and ‘becoming.’ It can be explained how and why it is so. It is obvious that one is not only concerned with numerical strength (or weakness) of a speech group here, nor is this applicable for those societies that face some biological problem (even if it is triggered by their marital patterns and endogamous behavior) in terms of reproductivity. But even such societies that have a long list of membership can disappear after a while if they decide to change their natural ‘being’ of bilinguality (or multi-linguality) to ‘becoming’ members of a larger linguistic entity – for various reasons, ranging from higher education, better job prospects, or more social prestige. In a multi-cultural milieu, it is natural being bilingual. But in the process of becoming a bilingual, it is, as if, a large scale religious conversion at work, reminding the members of a smaller community of the limitation of its space (physical, cultural and social) and the opportunities it offers or cannot just because of their language tag.

Secondly, the danger comes also because of the ‘perception’ of one’s language universe. The members of a smaller linguistic group will often have to negotiate with some of these questions: How do you see yourself, and how do others look at you? Which of these two evaluations do you accept and why? How deep is your commitment to your ‘own’ tags of identity – the attachment for your own land, flora and fauna, dress and food, life and living, speech and writing system, or creative and re-creative texts? There is this second type of tension in smaller linguistic groups which was pointed out long ago by Dell Hymes in his paper ‘Two Types of Linguistic Relativity’. We are told how a relatively smaller-size speech group survives the onslaught of the forces of globalization, whereas how a larger group that accepts a negative evaluation of others force their own members to ‘flee’ from their own tags.

The third type of tension is how a speech group tries to ‘include’ or ‘exclude’ other

speech groups that may be genetically, genealogically, and culturally related to their own group, and whether the group in question tries to divorce from this larger identity to create a space for its own language or tries to forge an identity with the other to create a new and larger identity. ‘Split’ and ‘Merger’ are the political game societies often play, and this dynamics is often difficult for an observer from outside to understand but is nevertheless an interesting subject of speculation by the students of social and political sciences. But there again, one can only create only a speculative theory about the ‘thinking construct’ of the members of the group that play a prominent role in decision-making, and not be sure that the theory will work elsewhere, too, given another configuration of linguistic majority and minorities living and sharing a given geo-space.

However, this urge to be like the rest of ‘our’ type, or the itch to forge one’s own destiny quite often grows out of particular historical moments, or as counter-moves against a perceived move by others. It could also grow as a result of this tension of inclusion versus exclusion, and how a community decides to negotiate with it.

1.5. The Plan of Action As Conceptualized

The idea of preparing an action plan on endangered languages and cultures of India began with the University Grants Commission having set up an Empowered committee to identify areas of research and development with respect to indigenous languages. The Committee was also asked to suggest roadmaps for university departments of anthropology, linguistics, Indian languages, Cultural Studies, Comparative Literature, etc that must engage themselves in these areas, ensuring, however, that there were no overlaps and repetitions. The recommendations to ensure the continuity and transmission of the Indigenous Languages were considered by the Commission later.

It is also observed that while many developed countries (Germany, Canada, USA or UK) allocated substantial resources for the study of indigenous and endangered languages, something that the Governments of India at both Centre and State levels could begin to emulate and implement. It was hoped that with the new Census work, the issue of categorization of languages with special reference to the Indigenous languages and mother-tongues would receive due attention. That there has been a progressive decrease in enumeration of languages, and consequently, shrinking of a large number of language bases are well-known facts. The shrinking of bases has not been only due to internal displacements, but also due to majority-minority confrontation, or because of shift of identity by originally indigenous groups. India should positively encourage this plural ethos – through a set of organized state interventions.

It may be good to recall here that the awareness in India began with the new edition of

*UNESCO Atlas of the World's Languages in Danger*¹¹ appeared in 2010, and it had vastly improved the 1996 and 2001-version of the work both in terms of content and coverage, besides ensuring a state-of-the-art presentation online. A team of international scholars worked on the whole world scenario by dividing it into regions in the following manner, because the project was dealing with a large and geographically diverse space:

North Africa and Arab States: Salem Mezhoud, Yamina El Kirat, Bonnie Stalls
Sub-Saharan Africa: Matthias Brenzinger, Herman Batibo
Europe and the Caucasus: Tapani Salminen, Tjeerd de Graaf
Siberia and North-East Asia: Juha Janhunen
South-East Asia: David Bradley, Suwilai Premasirat, Gérard Diffloth
Pacific, Indonesia and Papua-New-Guinea: Darrell Tryon
Australia: Michael Walsh
North America: Lyle Campbell (with Naomi Fox and Chris Rogers), Mary Jane Norris
Mexico and Central America: Yolanda Lastra
Lowland South America: Marleen Haboud, Bruna Franchetto, Denny Moore
Highland South America and Caribbean: Willem Adelaar, Gustavo Solis Fonseca
West and Central Asia: Hakim Elnazarov
India and Himalayan Chain: Stuart Blackburn, Jean Robert Opgenort, Udaya Narayan Singh

Some specialists also accepted to provide complementary information on specific areas or languages, including Alexandre François (North Vanuatu), Sun Hongkai (China), Bruno Poizat (Aramaic), Tulio Rojas (Colombia), Jean Sibille (Aramaic) and Marie-Claude Simeone-Senelle (Southern Arabic Peninsula).

This whole group¹² decided to work in seven stages as follows:

1. The first stage was to check the existing maps (if available) of the regions in the 2nd edition (2001) of the *Atlas*, and note the locations and symbols used (a key to symbols was provided) to either confirm that they were still appropriate or alter them accordingly. This was usually done on a copy or photocopy of the map, clearly marking with a drawing pen. Where there were several editors for a region, the responsibility was defined at this stage.
2. In the second stage, the earlier texts were revised or new ones were created, in case no text relating to the maps of the region existed before. This was meant to be a general survey of the language endangerment situation in the region, with reference to

¹¹ See <http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/>

¹² The details of the group responsibilities could be seen in the Annexure at the end of this document.

specific languages if there were not too many of them. Remarks on the current causes of endangerment and threats to the languages' survival were considered relevant at this stage. Population figures were contrasted with speaker numbers and the known rate of decline, especially over the period since the last edition, if figures or adapt were available. At this point, the regional editors co-ordinated data from contributors with respect to number of speakers and classification etc, and attributed the information solely to them as many regions in the world were still much less known, except for efforts by individual researchers.

3. In the third stage, the revised or newly written text and maps were sent to the General Editor for checking of errors, and for harmonising the sections' texts and maps – so that there were no unnecessary overlap between the regions covered. All necessary cross-references were created at this stage.

4. In the fourth stage, the General Editor returned all comments and annotations to the Regional Editor(s), together with requests for clarifications from the contributors, especially to ensure that the sources were included in the general Bibliography at the end of the volume or web-portal.

5. The fifth stage was very crucial for the on-line edition of the work as it involved preparation for printing with professional cartography.

6. The sixth stage: General Editor wrote general introduction, citing history and genesis of the Project and Volume, the Red Book project, and detailing other organisations aimed at saving languages, such as Foundation of Endangered Languages (FEL). Meanwhile, the Regional editors received the proofs of the cartographic production of their data for approval and correction.

7. The seventh stage consisted of preparation of lists of languages mapped, classified by degree of endangerment, as in 2nd edition, and of alphabetical list of languages. The bibliography was also brought up to date, with finer details in consultation with UNESCO.

When *Times of India* reported on 21 February 2009 about UNESCO's list of endangered languages of India¹³, it was not anticipated that a lot of discussions and public debate would follow. The brief innocuous-looking news-item read as follows:

“With 196 of its languages listed as endangered, India, a nation with great linguistic diversity, tops the Unesco's list of countries having maximum number of dialects on the verge of extinction.

¹³ <http://timesofindia.indiatimes.com/india/Unesco-lists-196-Indian-languages-as-endangered/articleshow/4164064.cms>

India is closely followed by the US which stands to lose 192 languages and Indonesia, where 147 are in peril.

The facts were revealed in the latest Atlas of World's Languages in Danger of Disappearing unveiled by the UN's cultural agency Unesco on the eve of International Mother Language Day on February 21.

The atlas classifies around 2,500 of the 6,000 languages spoken worldwide as endangered."

The PTI had already reported this on 20th February 2009, and the item was carried out on its news portal by Rediff.com.¹⁴ It quoted the then UNESCO Director-General Koichiro Matsuura and reproduced his statement: "The death of a language leads to the disappearance of many forms of intangible cultural heritage, especially the invaluable heritage of traditions and oral expressions of the community that spoke it." This news item rather gave a detailed account of those 199 languages that are spoken by fewer than a dozen people, and 178 others with 10 to 50 speakers, including Karaim (in Ukrain) with six and Wichita (Oklahoma, USA) with only 10 speakers, and also mentions some revitalization efforts such as Cornish in the UK and Kanak¹⁵ in New Caledonia.

By March 2009, there was already a debate in the Upper House (Rajya-Sabha) of the Indian Parliament on the issue of language endangerment, and it expressed this concern that this list of Indian languages in danger did not even feature in the country's 2001 Census of India Report. During the debate, it was stated that the claim was that the Indian languages in the 'danger zone' are spread across the country, from the far north in Jammu & Kashmir to Tamil Nadu and Andhra Pradesh and the North-east in Assam and Manipur, and that five were already extinct including three — Ahom, Andro and Sengmai — in the North-east, and two others — Rangkas and Tolcha — in Uttarakhand.

When the initial discussions began, it was noted that there was at present no scheme exclusively for the Preservation and Development of those indigenous languages of India that were not included in the 8th schedule and had relatively less number of speakers. It was known that the Ministry of HRD was working on projects that would cover the endangered languages of India. Under the 11th Plan, the MHRD had proposed a Scheme for the Development and Promotion of minor Indian Languages, titled *Bharat Bhasha Vikas Yojana*, which was yet to take off. Later, by May 2013, the Central Institute of Indian Languages (CIIL), Mysore was asked to initiate a project titled *Protection and Preservation of Endangered Languages of India* (SPPEL). The CIIL's plan was to document nearly 500 endangered languages in the country, each spoken by less than 10,000 people, with an objective to bring out dictionaries and also

¹⁴ <http://www.rediff.com/news/2009/feb/20languages-are-in-peril-says-unesco.htm>

¹⁵ See: Chappell, David A. (2013) *The Kanak Awakening: The Rise of Nationalism in New Caledonia*. [Pacific Islands Monograph series, 27]. Honolulu: The University of Hawaii Press.

document and preserve the ethnic knowledge system enshrined in these languages, including their folklore. For this, it planned to start with a 2500-3000 word-list and a list of basic sentences. It also intends to frame grammar rules, and make efforts in revitalisation of these languages. It began by working on 70 such languages in the initial phase¹⁶.

Besides the above, the Ministry of Home Affairs, under its ‘Language Division’ attached to the Registrar General of India (RGI) initiated a move to study and describe the unclassified mother-tongues of India under another project called the *Mother-Tongue Survey of India* (MTSI). It is to be noted that once the MTSI’s work of re-categorization and classification of languages progresses further, it would be of great help for all in the field.

Although the new series of LSI activities had been going on since the 8th plan period, the MTSI or Mother-Tongue Survey of India activity was a new idea that began being operationalised by August 2010 when its first Technical Advisory Meeting was held to consider the methodological issues related to the Linguistic Survey of India (LSI) and the Mother Tongue Survey of India (MTSI) being conducted by the ORGI (Office of the Registrar General of India) through its ‘Languages Division’. It also discussed the issue of Training and supporting infrastructure needs for the surveys. During the 2001 Census data collection, a total of 6,661 raw mother-tongue and potentially language labels were received as returns which were scrutinized, edited and rationalized by comparing with all earlier Census records. The result was that 1635 mother-tongues were found attested and rationalised whereas another 1957 names were accepted as mother-tongues but were grouped under ‘Unclassified’ or ‘Other Mother-tongues’. In fact, the scrutiny is also based on a physical verification through exhaustive field surveys and also on the basis of existing knowledge base. It is to be mentioned that the original *Linguistic Survey of India* (LSI) conducted by the British government under the direction of Sir George Abraham Grierson between 1898 and 1928 had covered “the whole of India except Burma, Madras Presidency and Hyderabad”, where under the first division, the local authorities sent the list of labels, showing the languages and dialects spoken in each area, and under the second division, collection of specimens of each of these types were prepared by the local officers. Grierson had edited the specimens, publish selections from them and prepared a narrative identifying languages and dialects in terms of six broad language families, including the Indo-

¹⁶ Aiton/Khampti, Ardan, Atong, Bagi, Balastin, Bangro, Baradi, Bateri, Bawm, Bhadriliam, Beda, Bhala, Bharwad/Bharwadi, Bhunjia, Birhor Bodo Gadaba/Gutob, Bondo, Binjhia/Birjia/Brijia (Partly documented), Chinali, Dargari, Darlong, Darmiya, Dhimal, Eravallan (Partly documented), Gahri, Gorum (Partly documented), Hakkipikki, Handuri (en)/Handuri (fr)/Handuri (es), Holiya, Jad, Jangali, Jangshung, Jenu Kurumba (Partly documented), Kanashi, Khamba, Khamiyang, Khash, Komkar, Kutiya, Kushwahi, Lamongse (en)/Lamongse (fr)/Lamongse (es), Luro (en)/Luro (fr)/Teressa (es), Malasar, Malayan, Manda (Partly documented), Masidi, Meyor, Momba, Moopan, Mru, Muot (en)/Muot (fr)/Nicobares Central (es), Naa, Paliya, Pu (en)/Pu (fr)/Pur (es), Ralte, Ranglong, Rongpo, Sanenyo (en)/Sanenyo (fr)/Chaura (es), Sherdukpen, Simong, Singpho, Spiti, Takahanyilang (en)/Takahanyilang (fr)/Takahanyilang (es), Tangam, Tehgul, Tharua, Thoti, Uchai, Yobin/Yobin, Liju, and Zakhring

European (the Indo-Aryan and Iranian), Dravidian, Austric (Austro-Asiatic), Tibeto-Chinese families and the unclassified languages, etc. Since the geo-political boundaries of the South Asian region underwent many changes with new nation states emerging (Burma, Pakistan or Bangladesh being examples), and since Hyderabad and Madras presidencies included most of the southern states as on this day, the ORGI thought it was necessary to revisit the ground through fresh surveys. Also, Grierson's survey had attracted criticisms from professional circles of linguistics because of its overly simplified questionnaire and its dependence on untrained local officials. The new series of LSI activity was conceived of in the '80s by the ORGI, and it covered Orissa, Dadra & Nagar Haveli, Sikkim, Rajasthan, West Bengal, Bihar & Jharkhand and, Himachal Pradesh. The major languages and mother tongues of these states and union territories were covered in this fresh LSI, and some state-specific reports were also published or drafted. Yet, the biggest challenge was to take up a project to survey the unclassified mother tongue returns found in the Census to decide if these mother tongues were local names or dialect labels of already rationalised and classified mother-tongues or these need to be accepted as independent language labels. For this purpose, a total of 541 out of 1957 unclassified mother tongues were taken up to start with, and reports were written for about 230 of them. The remaining 1416 were left out for the present because of their meagre speaker strength. Although it appeared alright in terms of a working strategy, it must be mentioned that potentially these 1416 could qualify as mother-tongues or languages that were endangered should they turn out to be important additions to our languages list. Besides the in-house linguists of the ORGI, the MTSI survey was outsourced to various university departments after formulating a set of questionnaire and methodology. In all 1200 new recruits were used to collect or validate field data. The format of the new series LSI and MTSI required some modifications for better quality and greater use.

As a matter of fact, experts also suggested guidelines for future survey-work. For instance, they suggested training programme for two groups - [1] one that is a new recruit in the Census uninitiated in Linguistics, and [2] another which has had varying degrees of training in Linguistics/Field work/language analysis, because it was felt that they needed to have two different levels of intensity and perhaps a different time frame for each group.

For the purposes of data collection connected with languages, I guess the first group will need at least a six-months of intensive training - a kind of six hours per day into six days a week into six months, something like 1296 hrs, divided into the following:

- [i] 500 hrs spread over 125 days of Phonetics/Instrumentation/Data collection techniques @ 4 hrs each day (360 hrs Phonetics, and 140 hrs for the rest);
- [ii] 200 hrs of phonology;

- [iii] 296 hrs of morphology & word formation; and
- [iv] 300 hrs of syntax and data elicitation techniques with special reference to the instrument you would create.

If we leave out 2 hrs each day for library work or for language lab practices. For the second group, one may still need a three-months (600 hrs) training to brush up the team's

- [a] Morpho-Syntax (25% time);
- [b] Intro to the Field Technique with a practicum class....in the same slot, one could introduce the new instrument (30% time); and
- [c] a thorough training in phonetics and visual phonetics as well as the database architecture (45% time).

As a matter of strategy, the first group could be divided into sub-groups where in each sub-group you would include some from the second group. The rest of the second group could be at the data handling headquarters to systematize, arrange, do basic analysis and basic report writing work. The training would require about 14 (if not 20) faculty members for (i) to (iv) and (a) to (c) assuming one is training 50 persons in each batch with two batches under each group, i.e. a total of 200 trainees. And once the first set of 200 is ready, the pilot survey could begin with the next 200 getting trained in the next six months. The third batch of trainees could then have only the first type of personnel. It was thought that 600 ground level survey personnel would be enough to cover all India – perhaps region by region. Therefore, in all there would be 400 language data enumeration personnel and 200 plus linguistics graduates. A similar training regime could be thought about for other projects as well.

Notice that besides these projects supported by the Governmental agencies, there were other non-governmental initiatives such as *Peoples' Linguistic Survey of India* (PLSI), spearheaded by the Bhasha/Tribal Research Academy, Tejgarh/Vadodara. Similarly, the *Ethnologue* supported by the SIL International¹⁷ - had already given a list of languages on the basis of 10 point scale of Vulnerability. The PLSI and Ethnologue models were different, and would not be discussed here.

The typical problem that came up in all these studies was that there were different names used for the same language in the *Ethnologue* list (at <https://www.ethnologue.com/>)¹⁸, and in other reports. One way out seems to be to opt for the official name(s) – where at times optional names were also suggested, as declared by the Government of India, although the other alternate names would also be

¹⁷ SIL is a U.S.-based, worldwide, Christian non-profit organization, whose main purpose is to study, develop and document languages; See: <http://www.sil.org/>

¹⁸ Of which Gary F. Simons was the Executive Editor, M. Paul Lewis the General Editor, and Charles D. Fennig the Managing Editor (with a research team of Colleen Ahland, Michael Ahland, J. Albert Bickford, Lorna Priest Evans, and Linda Simons), and the Lead Cartographer being Irene Tucker.

included during the description. This is because at times the alternate names given by the outsiders or the researchers and the names given by the community themselves would throw more light on the given mother-tongue.

Having noted all these, with respect to conceptualizing the CFEL programme, it was agreed that the UGC would be in touch with the Ministries of HRD, Culture, and Home Affairs, its other organizations such as ASI, CIIL, NLM or SSS, or NCERT etc which may help in beginning this new initiative. All agree that the linguistic diversity in India is a unique heritage which must be preserved and nurtured. We are aware that even among the 8th Schedule languages, some languages are lagging behind in facilities to produce literary works of great merit, whereas many other newly emerging literary languages need encouragement to publish their creative works. The other priorities include preparation of necessary pedagogical instruments and texts to support teaching-learning of these languages in our elementary and middle schools. Lastly, one also needs to encourage publishing activity in these languages before any further depletion of their rich heritage takes place. It was noted with some satisfaction that the Tribal Languages Development Board of Sahitya Akademi under the Ministry of Culture got the approval for a special-purpose project titled 'NEIL-Srijana' (North Eastern Indian Languages-Srijana), as proposed and designed by Udaya Narayana Singh for the Akademi. A similar project of importance, namely, 'Indian Literature Abroad' or ILA, was also assigned by the Ministry of Culture to Sahitya Akademi to take it forward.

The general agreement in the UGC initiative was that the paradoxical issues in the society which neglected the indigenous communities by "*otherization*" of these communities as well as their languages and culture should be carefully considered and the world view of the "Other" should also be recognized. It is sure that "these others" are relevant to institutionalized forms of language or religion. There is a need to look in to the paradigm of Man-Nature relationship. Indigenous people are relevant in the contemporary scenario as we are dealing with living human beings. It is obvious that the impersonal classification of people on the basis of physical and anthropometric features should be modified about which Dr Kapila Vatsyayan had spoken in several places. the Government may take immediate steps to introduce as many of these Indigenous languages as possible into the school system as nothing else can arrest the high dropout rates among tribal children from schools where they are forced to speak in "mainstream" languages. Further, it was noted that many a times a speech community's demands are put off on the argument that it lacked a script of its own.

Ideally, any policy that may lead to alienation of the so-called "tribals" from their own indigenous language, culture and traditions should be modified. There is a need to devise new educational models for indigenous and minority children, because the questions of Indigenous languages and endangered languages are closely related. For revitalization of languages, especially those under threat, language technology can be of great help. In addition to creating structures, it is also necessary, from a long term

perspective, to take steps to create awareness and opinion among the people speaking major languages in our country that they must support the development of indigenous languages, cultures, and knowledge systems. Support may be given to the production of awareness literature - informative and readable knowledge literature - on the subject, among similar activities. Only then could we do our bit to help preserve these smaller and marginalized speech varieties – or mother-tongues. It is useless to rake up or join in the endless debate on whether a particular mother-tongue could be or should be given the status of a language, or should it be forced to accept the status and nomenclature of a dialect.

At the *21st ECMSAS, 2010*, held at the Universität Bonn's Institut für Orient und sienwissenschaften Abteilung für Indologie, I had made a comment which is still relevant, and I quote from there:

“Let us try and understand how these numerous languages are structured within this space. My conjecture is that they are organised in layers, and that each Indian ‘language’ presents itself as a ‘peelable’ entity, because each speech community has a number of languages in its repertoire – used for different functions. The base language(s) would be in the inner fringe – and so are the basic elements of grammar. First, the grammatical layer then the cultural layer, and then the moral layer. Even within grammar – there are outer and inner layers. Such was the conceptualisation within our linguistic tradition. The more you peel off, and the deeper you go, and it gets darker and darker; some think we know a lot, but the saner elements realise that we know less and less about the structure of inner elements. The situation becomes often more complex here with different languages occupying same or different spaces, because different languages in the same speech communities perform different functions, defining our linguistic landscape. Since each speech community has so many speech varieties, what is acquired through acculturation and socialisation are in the relatively outer fringes. The sheer number of languages, mothertongues, and speech varieties in active use has been very frightening...A lot of this plurality had been reflected in the number of languages and writing systems in which we find manuscripts and inscriptions from ancient and medieval India”. (Singh 2010)

It is this plurality of language that needs to be given description of through certain programmes, which would also suggest to the policy making and implementation agencies as to what ought to be done to arrest this erosion of our rich cultural bases.

1.6. Desired Action

Many have a general feeling that at this moment, the nascent discipline of 'Language Planning & Management' may not have a mechanism in place to protect and promote minor and endangered languages. There have been moves and attempts by different governments to engage with this subject though. The trouble is that many often point to small countries where Constitutional provisions and in Universal Education documents are specifically mentioned. In case of India, the sheer size of the country and complexity of the administration are such that her situation cannot be easily compared with other contexts. This is the first problem researchers in this area face as a challenge.

'Structure' of Endangered Languages has also developed as a theme of serious research in informed linguistic circles for about two decades now. This is of course of purely linguistic and typological interest. But it is the 'Function' of endangered languages, and their 'Consequences' that have drawn public attention thanks to a media coverage the UNESCO Atlas has received under which India was found to be on the top of countries with 197 endangered languages.

Traditionally India is viewed as a pluralistic society that is supportive of all languages-big or small. Our constitution too is committed to the language rights of all, including the right to mother tongue education. However, the education system has encouraged the growth of dominant languages more, and in practice most of the smaller languages are not included. This has resulted in marginalization of diverse linguistic communities and enhanced threat perceptions to their languages. To safeguard these languages we also need to formulate clear cut plans for the empowerment of these languages and their speakers. This would involve linking languages with literacy and education, with technology and with economic opportunities.

A clear assessment and understanding of this sense of danger is possible where one finds that Language Planning is used not only as an extension of governance over a plural linguistic space, which is usually done with proclamations, orders and legal instruments, but they often become a mechanism of state repression or dominance of the majority linguistic group(s) that rule the state. If Language Planning is an extension of struggle to build civil societies, those who would study language endangerment to suggest any remedy, would have to have their rightful place in decision making as they have an understanding of the total nature of the language endangerment inherent in a given situation. The three kinds of tension as described earlier, added to the laws of majority-minority configurations (Samuel P. Huntington, 1993, 'The Clash of Civilizations?' *Foreign Affairs* 72: 22–49, and 1996, *The Clash of Civilizations and the Remaking of the World Order*, New York: Simon and Schuster) as well as the forces at play in the game of managing 'scarce opportunities' – all of them together create a 'total' sense of danger, the metaphysics of which must be understood before one recommends any cure.

The issue has already reached the Indian Parliament where the consequences and remedies were thoroughly discussed by the members. Therefore, the academic bodies and institutions are now asking if many of our smaller languages, several of which belong to tribal and marginalized communities, are indeed seriously threatened, and what steps are to be taken in order to safeguard our most precious linguistic heritage, and document the same. Here lies the genesis of the attempt of the University Grants Commission (UGC) to set up a few Centers for Endangered Languages (CFEL) in both Central and State Universities.

One of the major tasks would be documentation of endangered languages and mother-tongues in India. Documentation of languages and literature in oral traditions is recommended as an activity for support, including digital and photographic archiving and documentation of socio-cultural contexts important for our understanding of the nature, production processes and creative aspects behind emergence and sustenance as well as social function of such literature. The documentation activity that could be supported should also include compilation of scholarly/historical/generic/style-based and period-based/dialectal/diachronic/ thematic anthologies of available written literature in such languages. The activity will also include re-production of works of earlier ethnographers, linguists and freelance researchers having bearing on literature. It will further include preparation, publication and dissemination of such literature aimed at wider readership. The dissemination can be in electronic, digital, print or mimeograph medium. Such documentation should also include collection and conservation of previously documented literature in their 'original first editions' or 'original manuscripts' as a record of orthographic conventions in a given language. It is also expected that a metadata would be created for ease of accession of all such documented literature by the Institute or by an agency (for the Institute), and such tasks could also be outsourced under this category of support. The expected output would be:

- i) books and/or edited collections,
- ii) multi-lingual or multi-cultural anthologies, or anthologies of genre- or literature-types,
- iii) information databases and search engines,
- iv) programming codes and database structure and/or metadata designs,
- v) research papers,
- vi) broadcast quality visual documentation, and
- vii) Still photographic archives.

The second possible task could be formulating or designating scripts and typography codes. The activity under this type will include projects in historical and sociological research related to scripts and orthographic conventions of smaller languages (where there is a demand), typographic and design projects for creation of scripts and writing systems for smaller languages although under strict quality control and parity with demands of modern-day key-board and UNICODE requirements. The focus of these

projects will be towards designing typographic codes/scripts for producing textbooks that will be user-friendly for school students in these languages. Grants could also be made available for linguistic and lexicographic research projects using these writing (and printing) systems. The expected output would be:

- i) Creation of a de novo or modified writing system for a given language,
- ii) Documentation of their existing or old variety of scripts,
- iii) Designing of typographic conventions, and preparation of case sheets for the UNICODE and national consortia
- iv) Production of text-books or on-line teaching materials with a focus on writing systems.
- v) Reports, books and research papers on single scripts and multi-scriptal situations.
- vi) Standardization of Graphemic and/or Spelling Conventions.

The third set of activities could include preparation of Pictorial Glossaries. This set of activities envisages conceptualization, preparation, production and distribution of word lists/glossaries (with a standard format to be created by the Institute) of these languages in state script or community's preferred script(s), or in multiple scripts (especially for the use of teachers whose mother tongues are not these languages). These language glossaries shall be prepared for specific regions or for specific locations. The glossaries could also be pictorial or illustrative, so that they form a bridge between the unlettered pupils, and the teachers for whom meaning concepts in these languages are not easily accessible. The languages that already have such glossaries, which may also be in circulation, can also produce specialized (or, revised/elaborated/abridged) pictorial glossaries for use in economic transactions or else related to areas such as housing, costumes, jewellery, festivals, rituals, craft, medicinal practices, law, the IPC, various developmental schemes, local history, etc.

The expected output would be:

- i) Books and/or edited collections on lexicographical issues,
- ii) Multi-lingual or multi-cultural glossaries, or collection/expansion of such glossaries for special or general purposes,
- iii) Web-version of such glossaries, and
- iv) Creation of software/databases/designs of such glossaries.

Fourthly, one could also prepare dictionaries and grammar books of these languages. Such specialized research projects will include graded grammar books for use in schools as well as general-purpose grammar books for use among linguists and scholars. Similarly, the dictionaries to be prepared under the activity will be assorted dictionaries including the following:

- i) General and/or special-purpose dictionaries for language use to facilitate standardization of languages,

- ii) Mono-lingual dictionaries for use of the mother-tongue speakers;
- iii) Intra-dialect dictionaries for standardizing languages;
- iv) Bi-lingual dictionaries for language learners and for translation purposes,
- v) School/College Grammars, including Contrastive Grammars; and
- vi) Full-fledged/Comprehensive Grammars or Linguistic Descriptions of the entire language or of certain components or comparative grammatical works involving two or more languages.

The fifth and next task is with respect to preparation of textbooks of primary and secondary education, where support of school-teachers, pedagogists, linguists, and scholars are needed for preparing the same language and bi-lingual school textbooks, at par with the state language textbooks for all subjects in primary schools. The activity should extend support to curricular experimentation of such books in established formal schools. The school textbooks will make optimum use of locationally relevant illustrations and concepts in explaining the contents conveyed through the books. In preparation of secondary school textbooks, priority will be given to preparation of books of scientific subjects, particularly computer literacy books. The textbooks will be expected to remain sensitive to the theological/religious practices rooted in the language, as well as the ecological and gender sensitive issues valued by the language community. The expected output would be:

- i) Text-books and/or edited collections, or Supplementary Readings for language teaching
- ii) Text-books on different disciplines using the languages under question
- iii) Literacy Books and Materials; and
- iv) Practice-books and Testing materials.

Sixthly, one needs to help these threatened languages so that they could begin to use their language to bring out little magazines and periodicals. Specialized journals devoted to study of these languages or produced in these languages would also fall under such tasks. One should encourage and help or advise individual periodicals, ‘little’ magazines, occasional publications, weeklies, fortnightlies, folios, etc published in these languages by way of annual financial grants – in the pattern of GIA for Little Magazines (published in 8th Schedule languages) already operated by CIIL. Similarly, magazines of serious nature promoting the study of these languages but published in Hindi, English or other Indian languages, could also be included in this activity for support. The extent of support could vary.

1.7. Training of Teachers in Schools Using these Languages

Teacher training is proposed for those school teachers who could be engaged under this project – preferably from among the community or near-by places/languages to

teach in local schools or literacy centres using these languages. It is expected that this structured training – to be handled by the 13 regional outfits jointly with one or the other university would sensitize them in all aspects of teaching/learning of these languages and in comparing them with other languages as well as in work on oral literature, lexicography, translation practices, and material production in order to relate them with state language(s), English and/or Hindi, as the case may be. This will also ensure effective use of textbooks and other teaching materials produced in these languages. This part of the scheme does not pertain exclusively to training of ‘language-teachers’, though they are not excluded. The scheme will be applicable to teachers of all subjects who wish to use the local language(s). The output would be certain number of trained manpower resources for these smaller languages, and also provide them with opportunities to teach in their regions, and offer them appropriate post-training support for this purpose.

1.8. Study of Endangered Languages: Aims & Objectives

- To undertake inter-departmental and inter-disciplinary research related to endangered languages.
- To undertake fieldwork, research, analysis, archiving and documentation of smaller indigenous/endangered languages using state-of-the art speech and language technologies, in formats that are universally acceptable viz. digital textual, audio and video formats.
- To produce and publish monographs, grammars, grammatical sketches, dictionaries and lexicon, ethno-linguistic and theoretical descriptions, collection of oral and folk literature and scholarly book on endangered languages.
- To produce language and dialect atlases with special reference to minority and endangered languages.
- To organize workshops and seminars aimed towards promoting advanced research related to endangered languages.
- To train teachers and students from other departments/centres in Field Linguistics, Lexicography and in techniques for data management and documentation. Field Linguistics should constitute an indispensable part of the centre.
- Each unit should serve the indigenous and endangered language communities by making accessible the products of the research of the Centre i.e. digital and analogue archives of linguistic data, language teaching material, and language artefacts.
- To promote and foster various domains of endangered languages so as to ensure minority/endangered language communities in maintaining and preserving language vitality, including the development of orthographical resources like scripts, book of letters primers.
- To digitize data collected in the course of the research in the Centre and make

it available to public by internet.

- In the initial phases one may have to draw resources (such as manpower, labs, books, students etc.) from other centers of languages, linguistics, folklore, anthropology, and literature in the university but eventually should conceive of forming an independent centers purely devoted to the issues of endangered and indigenous languages.

2. Endangered Languages Field Work : A Short Introduction

2.1 Aims & Objectives

The aim of any fieldwork is to collect data on certain aspect of human enquiry in natural environment. In linguistic fieldworks linguists are supposed to study a language in its natural environment, i.e. in the place where it is spoken by the people who usually speak it. In other words, it is the collection of accurate data in an ethical manner. It involves producing a result which both the community and the linguist approve of. That is, the ‘community’ should know why and how the fieldwork is being done. They should also be comfortable with the methodology and the outcome. Another component involves the linguist interacting with a community of speakers at some level. That is, fieldwork involves doing research in a place where the language is actually spoken. Linguistic fieldwork is about working on a language in culturally, socially and ethically appropriate ways in a context where the language is being used.

2.2 The Issues

The following issues need to be addressed in a linguistic fieldwork—

- **Data collection**—there are established techniques for obtaining linguistic data. The fieldworker doesn’t only collect data. There is more to data gathering than just asking questions. Decisions need to be made as to what to record, what to collect, and what to write down. Data should be naturalistic and spontaneous speech data.
- **Interpretation of data**—data must be interpreted. This is where our previous linguistic training comes in. We also need some way to organize the data effectively, we shall need some method of categorizing, coding and storing the information we collect – that is, we need a structured and searchable database.
- **Administration & liaison**—community-linguist interaction issues tend to consume a large proportion of a fieldworker’s energy. We need to pay our

consultants for their time, need housing and food at the field site, need to administer grant monies and keep appropriate records and need to arrange appropriate dissemination of our research results within our field community.

- **Handling technical issues**—fieldworkers will be making audio (and maybe video) recordings of their consultants, and they need to be able to operate their recording equipment effectively.
- **The ethical issue**—Fieldworkers should get the data ethically, i.e., without violating local customs. The process of going to a community to work on a previously undescribed language has non-linguistic implications. Does the community approve the writing of their language? Do speakers mind being recorded? Perhaps one is working with the last few fluent speakers of a language; does he/she have an obligation to provide teaching materials, learner's guides and dictionaries, even if they might not be used and younger members of the community are not interested?
- **Anthropological issue**—it's impossible to do fieldwork of any length without also (consciously or unconsciously) observing human interaction and cultural practices. Learning about the culture of the speakers, whose language is being studied, is vital, not only as a key to the language but also as a key to better fieldwork. For example, we are unlikely to get good data in a field session involving both men and women if the culture has strong prohibitions against men and women interacting. Religion is another issue which need to be addressed carefully so that it does not hurt anybody's sentiment.
- **Behavioural issue**—field workers need to be aware of their own behaviour in the field and how it reflects on them and their culture. They are also required to fit in with a new society and learn a new language, while retaining contact with their other lives as academics. Fieldworkers don't leave behind their own identities and culture when they go to the field. This is why there is much more to linguistic fieldwork than just turning up to record someone!
- **Literature survey**—fieldwork is not done in a vacuum. While it is good practice to rely only on our elicitation in a field methods class, in the field we need as much information about the language and culture as we can find. We should make the most of available resources so that we are not duplicating the efforts of others. Many fieldworkers also have an epigrapher's hat too.

3. Mapping of the Field

3.1 The Tasks at Hand

The primary need for field surveying is to choose the area and languages which need to be surveyed. To fulfil this task following methods could be applied:

- Follow Census Report (2001), Language Atlas by UNESCO (2010), or *Ethnologue* (2015) to find out the areas where most of the endangered language speakers are surviving.
- Find out the areas where the numbers of tribal people are more.
- Locate the border areas where due to socio-political reasons migrated people are more in number and the community loses its original language. Thus the community becomes multi-lingual and people start following the dominant language/s.
- Another linguistic character of the border areas and the big cities is that the migrated people (particularly the young generation) leave their native language and co-adapt the dominant language of that place. These areas should be included for endangered language survey.
- According to recent Census report the gradability of endangerment of language becomes more where the younger generation doesn't use their native language anymore due to socio-politico-economic reasons. Following Census report those areas need to be defined and surveyed.

3.2. Language Family Maps

In what follows, a set of four language family maps are presented to give an overall picture. It may be noted that in the 10,000 plus list released by Census of India, where 122 languages figure, the distribution of these is reflected in these maps. They will look obviously different if one trusts the *Ethnologue* or the *Peoples' Linguistic Survey of India* (PLSI), or the *People of India* data prepared by the ASI is taken as source material. For example, under the encyclopaedic *People of India* (henceforth, POI) series of reports, the Anthropological Survey of India had identified 75 "major languages" out of a total of 325 languages used in Indian households. *Ethnologue*, too reports India as a home for 398 languages, including 387 living and 11 extinct languages.

There are four maps belonging to four major language-families of India:

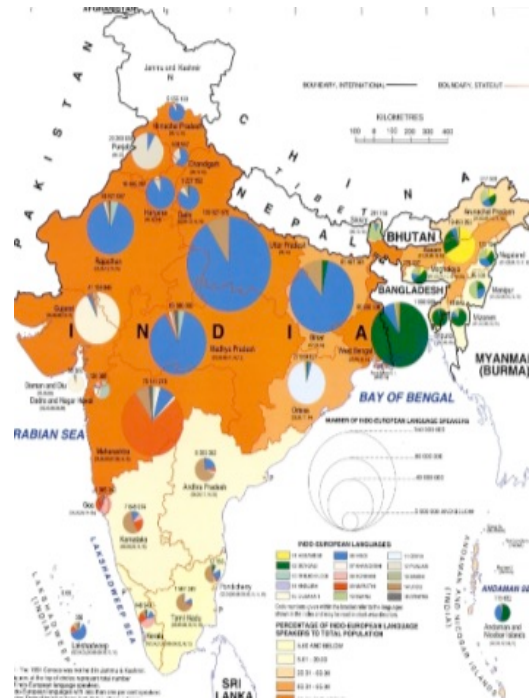
Indo Aryan Language Family

The Largest Speaking Population in the Country

Indo Aryan

21 languages with 106 Mother Tongues - 76.86% of Indian Population

1	Assamese (Scheduled)	1
2	Bengali (Scheduled)	4
3	Bhili/Bhilodi	17
4	Bishnupuriya	1
5	Dogri (Scheduled)	1
6	Gujarati (Scheduled)	3
7	Halabi	1
8	Hindi (Scheduled)	49
9	Kashmiri (Scheduled)	3
10	Khandeshi	4
11	Konkani (Scheduled)	3
12	Lahnda	3
13	Maithili (Scheduled)	1
14	Marathi (Scheduled)	1
15	Nepali (Scheduled)	1
16	Oriya (Scheduled)	5
17	Punjabi (Scheduled)	3
18	Sanskrit (Scheduled)	1
19	Shina	1
20	Sindhi (Scheduled)	2
21	Urdu (Scheduled)	1



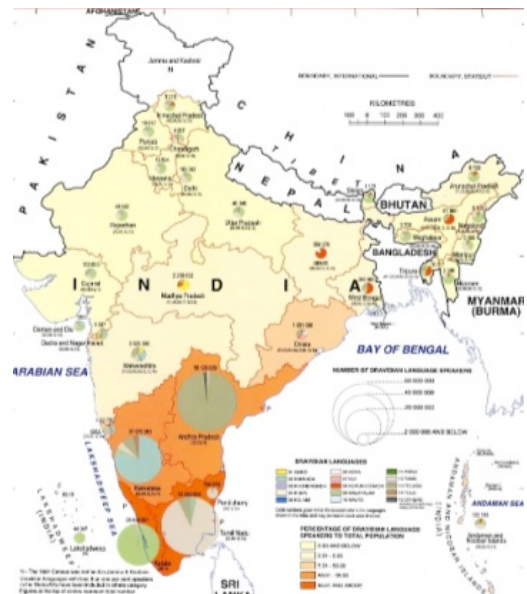
The Dravidian Language Family

Next to Indo Aryan as the Largest Group

Dravidian

17 languages with 27 Mother Tongues - 20.82% of Indian Population

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



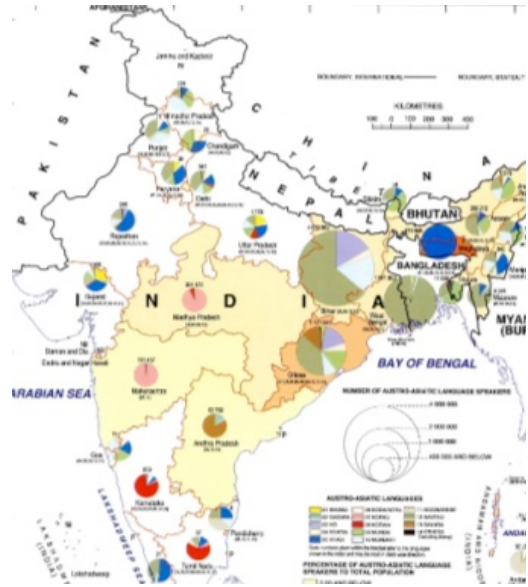
Austro – Asiatic Language Family

1 Percent but Extremely Diverse

Austro Asiatic

14 Languages with 21 Mother Tongues - 1.11% of Indian Population

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



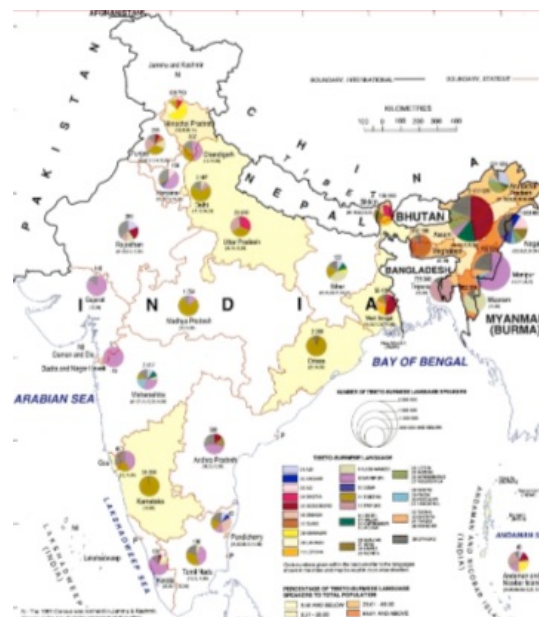
Tibeto Burmese Family

The Tiniest but the Most Diverse

Tibeto – Burmese

66 languages with 76 Mother Tongues - 1.00% of Indian Population

1	Adi	3
2	Anal	1
3	Angami	1
4	Ao	1
5	Balti	1
6	Bhotia/ zanskari	2
7	Bodo (Scheduled)	1
8	Chakesang	1
9	Chakru / Chokri	1
10	Chang	1
11	Deori	1
12	Dimasa	1
13	Gangte	1
14	Garo	1
15	Halam	1
16	Hmar	1
17	Kanui	2
18	Karbi / Mikir	1
19	Khezha	1
20	Khiemnungan	1
21	Kinnauri	1



Tibeto Burmese Family

The Most Number of Endangered Languages..

22	Koch	1
23	Kom	1
24	Konyak	1
25	Kuki	1
26	Ladakhi	1
27	Lahauli	1
28	Lakher	1
29	Lalung	1
30	Lepcha	1
31	Lianmei	1
32	Limbu	1
33	Lotha	1
34	Lushai / Mizo	1
35	Manipuri (Scheduled)	1
36	Maram	1
37	Maring	1
38	Miri / Mishing	1
39	Mishmi	1
40	Mogh	1
41	Monpa	1
42	Nissi / Dafla	4
43	Nocte	1
44	Paite	1

45	Pawi	1
46	Phom	1
47	Pochury	1
48	Rabha	1
49	Rai	1
50	Rengma	1
51	Sangtam	1
52	Sema	1
53	Sherpa	1
54	Simte	1
55	Tamang	1
56	Tangkhul	1
57	Tangsa	1
58	Thado	1
59	Tibetan	1
60	Tripuri	3
61	Vaiphei	1
62	Wancho	1
63	Yimchungre	2
64	Zeliang	1
65	Zemi	1
66	Zou	1

4. Socio-linguistic Profile of Informants

4.1. Domain of Use

What language(s) do you speak at home?

- With grandparents?
- With spouse?
- With children?
- With grand children?
- With siblings?
- With pets and livestock?

What language(s) do you speak at school?

- With the teacher in the classroom?
- With the teacher outside the classroom?
- With friends of the same community in the classroom?
- With friends of other community in the classroom?
- With friends of the same community outside the classroom?
- With friends of other community outside the classroom?

What language(s) do you speak at the market?

- With the merchant of the same tribe/community?
- With the merchant of the other tribe/community?
- With an acquaintance in the market?
- What language(s) do you speak at a place of worship?
- While praying to god?
- While reciting or performing rituals?

- While singing religious songs?

What language do you speak at your community meetings?

(Mother Tongue/Dominant Language/Other language)

Do you know English? Yes/No

Do you know Hindi? Yes/No

What language do you speak with a stranger? Mother tongue/Dominant language:

What language do you speak to the healer/native doctor?

Is your language used in the mass media?

- Radio
- TV
- News paper, Journals
- Others (specify)

What language is used for professional activities such as agriculture, hunting, fishing, elephant herding, honey/food gathering etc?

What language(s) do you use mostly for the following purpose:

- Thinking/planning
- Dreaming (at sleep)
- Self-talking
- Counting
- Abusing
- Joking
- Story telling
- Riddle/lullaby
- Flattering
- When get possessed?

- While talking with other worshippers at the worship place?
- For religious discussions at the place of worship with the priest?

4.2. Language Attitude

- Is your language/mother tongue easy to learn/speak: Yes/No
 - If yes, compared to which other language (s)
 - If no, compared to which other language (s)
 - Can you express your thoughts in your mother tongue : Yes/No
 - Do you accept mixing other language words in your mother tongue: Yes/No
 - If yes, with which other language (s)
 - If no, with which other language (s)
- When you are given opportunities will learn to write in your mother tongue: Yes/No
 - If no, which other language(s) will you prefer?
- Do you think that your language is better suited for performing various activities such as business, employment, schooling etc.: Yes/No
 - If yes, compared to which other language(s)
 - If no, compared to which other language(s)
 - Do you think that other languages are easy to learn : Yes/No
 - If yes, which other language(s)
 - If no, which other language(s)
- Do you think that one language is enough for all types of usages : Yes/No
- If no, which other language(s) should be included:
- Do you think that learning other languages improve your knowledge level: Yes/No
- What language(s) do you want your children to know well?
- What language(s) will you like to listen to scriptures in?

- Do you think that your language will still be used 20 years from now?
Yes/No
- What language(s) do you think your children will learn when they grow up? (If not your language, what do you think of this?)
- When your children become adults what language(s) do you think that they will speak with their children?
- Do you like your language to be spoken by your children?
- Do you speak your mother tongue as same as the way your parents speak your mother tongue?
- Do you think that your language is essential for day to day activities?
- Do you prefer to marry a person who does not know your language?

4.3. Language Prestige

- The language of which village is most prestigious? Why?
- Which of the dialects of your language is considered most prestigious? Why?
- Do people from neighbouring languages learn your language? Yes/No?
If yes, which dialects?
- Do you think your language is a hindrance to social economic mobility?
Yes/No
If yes, why?
- Do you ever feel that your mother tongue is dying?
- How do you react when the speakers of your language/mother tongue speak other languages?
 - Get annoyed
 - Encourage them
 - Will feel jealous of them
 - Avoid them
 - No feeling

5. Selection and Training of Enumerators

The following points need to be considered for the selection and training of the Enumerators:

Qualification of enumerators:

- **Essential**—Ideally at least Master's degree holders (failing which Bachelor's with Honours degree) in Linguistics, Language Technology, Anthropology, Geography, Culture Studies (including Indigenous Studies), Language & Literature (any Indian language or English), Sociology, Education, Folklore, Tribal Studies, Mass Communication, Area Studies.
- **Desirable**—Experience in fieldwork, a native speaker/ proficiency in the major language spoken in a particular area with minimum essential qualification and training in computer operations.

Appointment Procedure

- i. An advertisement is to be published in leading national dailies, website, notice board of the respective university/centre,
- ii. Personal interview. The PI may be the Chairman of the appointment committee—permission to be sought from the Vice-Chancellors of respective university.
- iii. Appointment may be made as per the UGC or Government of India rules, and the reservation policy is expected to be followed.

Training of enumerators –

An 8-10-day training on linguistic data collection, basic structural linguistics, i.e.- Phonetics, Phonology, Morphology, Syntax etc. (considering all the enumerators are not academically from linguistics background), introducing with IPA chart and how to use it, training on use of devices, etc. for enumerators.

Training of enumerators in mobile telephony –

- i. At first, procurement of GPRS-enabled telephonic instruments and recording cards for data collection would be made; and then
- ii. Training of enumerators in data gathering while training in mobile telephony data capturing and data transference will have to be arranged.

6. Template for capturing basic data of the field surveyors/ enumerators

1. Name:
2. Contact Details:
 - a. Address:
 - b. Contact no.
 - c. Email Id:
3. Sex: Male/Female/Transgender
4. Date of birth & age:
5. Name of the parents
 - a. Father:
 - b. Mother:
6. Marital Status:
7. Religion:
8. SC/ST/Minority Status:
9. Mother Tongue:
10. Languages known:
 - a. Read:
 - b. Write:
 - c. Speak:
11. Highest educational level attained:
12. Medium of instruction in primary and secondary level:
13. Occupation:
14. Place of birth:
15. Residential status: Rural/Urban/Semi-Urban (Mufassal):

Recent Colour
Photograph
(Passport size)

7. Criteria of Informants

7.1. Age:

- Age should not be below 12 years.
- The safe policy to interview among the younger generation should start from their teen onwards. The speech of the teenagers seems to be highly converged if the society is bilingual.
- People aged above 70 should be put in a special group.
- They seem to retain some of the indigenous grammatical structures and lexicons which have become extinct in the younger generation.
- Further, the investigator must as far as possible look for an informant who has all the teeth intact in order to elicit actual sound features of the language. Absence of teeth can distort the actual pronunciation.

7.2. Region:

Informants from all sections of a Stratified Society

- From both rural and urban areas.
- From different social and economic background.
- From different educational background.
- From different caste and clan.
- From different occupational background.
- Informants should be both male and female

7.3. General Qualifications:

The informant chosen for data elicitation is expected to have knowledge of the following areas:

- History of their community.
- Socio cultural information about the community.
- Folk literature pertinent in their community.
- Knowledge about the native language and cultural practices.
- Apart from this, it is ideal that the language consultant should have been a person who was born, brought up and continues to live in that ethnic settlement. It would be ideal if the informant is actively involved in various cultural and religious activities of the community.

7.4. Avoid:

- A multilingual informant in the beginning of data elicitation.
- One unwilling informant who does not want to give interview or relevant information.
- Working with a single informant (unless he/she is the terminal speaker of an almost extinct language).
- A language teacher

7.5. Good Informant (Kibrik-1977)

- Good knowledge of the target language.
- Good knowledge of the contact language.
- Translating capacity.
- Clear pronunciation.
- Associative mobility of the thought.
- Patience.
- Honesty and a lack of feeling of 'linguistic prestige'
- Strictness.
- Experience in working with the Investigator.
- Linguistically not sophisticated.

7.6. Personal Profile of the Informant

Personal Details

- Name of the Informant:
- Sex :
- Date of birth & Age :
- Religion:
- SC/ST/Minority Status:
- Education :
- Medium of instruction at primary level:
- Residence/Name of the Settlement:
- Alternate name, if any :
- Occupation :
- Place of Work/Distance travelled :

(Informant may be living in a certain village but his work place may be in city or elsewhere where his linguistic skill is being influenced)

- Mode of transport used for going to the place of work—on foot/bicycle/two wheelers/four wheelers/hired vehicles/public transport/no travel/others
- Work Status—main/marginal/non-worker
- Category of work—cultivator/agricultural worker/worker in household industries/other worker
- Nature of industry—trade/service
- Class of worker—employer/employee/single worker/family worker
- Informant's birth place :
- Name of the place where the informant was brought up:
- Parents Native Village :
- Parents First Language :
- Marital status :
- Age of marriage:
- If married, spouse's name :
- Whether the Spouse belong to the same community or the other community:
- Spouse's Education :
- Spouse's occupation :
- No. of children born:
- No. of children surviving:
- Informant's Parents' name
 Father :
- Mother :
- Informant's Parents' education:
 Father :
- Mother :
- Informant's Parents' occupation:
 Father :
- Mother :
- Number of brothers/sisters of the Informant :

7.7.Community Name:

- Name by which the community/mother tongue is identified:
- Indigenous name of the community/mother tongue:
- Name of the community/mother tongue given by outsiders/ government:
 (The investigator should inquire if the community and language/mother tongue have separate names.)
- How long have you been living in this place (duration of stay since migration)?
- Where did you live before moving into this place?
- Is this the original settlement of the community or did the community migrate to this place?
- If the community has migrated, Does (some of) your people still live in the

original settlement?

- Why and when did you migrate to this place (if migrated):
- Mention the names and number of settlements where the community lives:
- Do the community people live in towns/cities?
- If yes, mention the names of towns/cities:
- Name the other communities that live in your settlement:
- If so, what is the relationship with other communities- cordial/friendly/non-friendly/hostile?
- How many clans are there in your community: Name them.
- What are the kinds of marriage systems that community follows?
(Monogamy, Endogamy, Exogamy, Polygamy).
- Is marriage permitted within the same clan?
- What is the naming pattern (proper name: first name–middle name–last name (surname/title) of your community?

7.8. Language Identification

Primary Language spoken at home:

- If the primary language spoken at home is same with the mother-tongue or not:
- Other language spoken at home:
- Other language known to you:
- Speak:
- Read:
- Write:
- Understand:
- Does your language have a written form?

8. Word and Sentence lists

This word list includes words from Swadesh's list (1955), Gudschinsky's (1956) word list, Lingua's (1977) word list, Abbi's (2001) list and CIIL's (2014) word lists with appropriate modifications.

This data collection is possible to complete within 8-14 days/language/enumerator. A group of two persons for each language may be appointed. However, depending on the situation the number of enumerator, number of days, etc may also change.

WORDLIST

8.1.1 Human Body Parts

This domain will contain terms for both external and internal organs of the body. Along with types of the different human body parts, its sub parts such as types of fingers, lips, hair, leg, hand, tooth, bone etc should be included.

Sr.No	Word (in English)	Word (in)
•	Abdomen	
•	Abdomen (lower)	
•	Abdomen (upper)	
•	Alveoli	
•	Ankle	
•	Arm	
•	Artery	
•	Baby finger	
•	Baby toe	
•	Back	
•	Bald	
•	Beard	
•	Belly	
•	Bladder	
•	Blood	
•	Bone	
•	Bone marrow	
•	Braid	
•	Brain	
•	Breast	
•	Bun	
•	Cartilage	
•	Cheek	
•	Chest	

•	Chin	
•	Collarbone	
•	Curly hair	
•	Dandruff	
•	Dimple	
•	Ear	
•	Ear lobe	
•	Elbow	
•	Eye	
•	Eyebrow	
•	Eyelash	
•	Eyelid	
•	Face	
•	Feet	
•	Finger (Hand)	
•	Finger (Toe)	
•	Flesh	
•	Foot	
•	Forearm	
•	Forehead	
•	Gum	
•	Hair (Black)	
•	Hair (White)	
•	Hallux	
•	Hand (Left)	
•	Hand (Right)	
•	Head	
•	Heart	
•	Heel	
•	Hip	
•	Index Finger	
•	Intestine	
•	Jaw	
•	Kidney	
•	Knee	
•	Leg	
•	Ligament	
•	Limb	
•	Lip	
•	Liver	
•	Long toe	
•	Lower lip	
•	Lung	
•	Middle finger	
•	Middle toe	

•	Mole	
•	Moustache	
•	Muscle	
•	Nail (hand)	
•	Neck	
•	Nerve	
•	Nose	
•	Nose hair	
•	Nose hole	
•	Organ	
•	Ovary	
•	Palate	
•	Palm	
•	Partition of hair	
•	Pupil	
•	Rectum	
•	Rib	
•	Ring finger	
•	Ring toe	
•	Shoulder	
•	Side bun	
•	Skeleton	
•	Skin	
•	Skull	
•	Spine	
•	Stomach	
•	Straight hair	
•	Taste bud	
•	Teeth	
•	Temple	
•	Thigh	
•	Throat	
•	Thumb	
•	Toe	
•	Tongue	
•	Tooth	
•	Upper lip	
•	Uvula	
•	Vein	
•	Waist	
•	Wart	
•	Wrinkle	
•	Wrist	

8.1.2 Kinship Terms, Address and References

- While collecting kinship terms, investigator should remember to collect at least three generations preceding and following one's ego (self). For example, one's parents, grandparents, great grand-parents and so on, and one's child, grand-child, great grand-child and so on. It also must be noted if there is any different terms between, for example, maternal and paternal parents.
- The Investigator must check for the differences in address and reference terms.
- Is the same kinship term used when you address someone? For example, how a stranger or a wife/husband/superior would be addressed or referred keeping in view the formal and informal sense. Also, how it should be referred to in the presence or absence of the referee.
- What are the different terms of endearment, if any? For example, how do you address a lovely baby, your lover etc such as darling, dear, honey etc. in English culture.
- What are the terms or reference name used while abusing.

Sr No	Word (in English)	Word (in)
•	Ancestor	
•	Bachelor	
•	Brother (Born 2 nd or 3 rd or any middle no, not the eldest or youngest)	
•	Brother (elder)	
•	Brother (younger)	
•	Brother's daughter	
•	Brother's daughter (from female ego's side)	
•	Brother's daughter (from male ego's side)	
•	Brother's father in law	
•	Brother's mother in law	
•	Brother's son	
•	Brother's son (from female ego's side)	
•	Brother's son (from male ego's side)	
•	Daughter	
•	Daughter (Born 2 nd or 3 rd or any middle no, not the eldest or youngest)	
•	Daughter's father in law	
•	Daughter's husband	
•	Daughter's husband's brother	
•	Daughter's husband's sister	
•	Daughter's mother in law	
•	Descendants	

•	Divorcee (female)	
•	Divorcee (male)	
•	Elder brother's wife	
•	Elder sister's husband	
•	Eldest Daughter	
•	Eldest Son	
•	Father	
•	Father in law	
•	Father's brother	
•	Father's brother's daughter	
•	Father's brother's son	
•	Father's brother's wife	
•	Father's Elder brother	
•	Father's elder brother's wife	
•	Father's elder sister	
•	Father's elder sister's husband	
•	Father's sister	
•	Father's sister's daughter	
•	Father's sister's husband	
•	Father's sister's son	
•	Father's younger brother	
•	Father's younger brother's wife	
•	Father's younger sister	
•	Father's younger sister's husband	
•	Generation	
•	Grand Daughter	
•	Grand Son	
•	Grandfather (Maternal)	
•	Grandfather (Paternal)	
•	Grandmother(Maternal)	
•	Grandmother(Paternal)	
•	Great Grand daughter	
•	Great Grandfather (Maternal)	
•	Great Grandfather (Paternal)	
•	Great Grandmother(Maternal)	
•	Great Grandmother(Paternal)	
•	Great Grandson	
•	Great great Grand daughter	
•	Great great Grand son	
•	Husband	
•	Husband's brother	
•	Husband's brother's daughter	
•	Husband's brother's son	
•	Husband's brother's wife	
•	Husband's elder brother	

•	Husband's elder brother's wife	
•	Husband's elder sister	
•	Husband's elder sister's husband	
•	Husband's sister	
•	Husband's sister's daughter	
•	Husband's sister's husband	
•	Husband's sister's son	
•	Husband's wife (in case of two wives)	
•	Husband's younger brother	
•	Husband's younger brother's wife	
•	Husband's younger sister	
•	Husband's younger sister's husband	
•	Lineage	
•	Married (female)	
•	Married (male)	
•	Mother	
•	Mother in law	
•	Mother's brother	
•	Mother's brother's daughter	
•	Mother's brother's son	
•	Mother's brother's wife	
•	Mother's elder brother	
•	Mother's elder brother's wife	
•	Mother's elder sister	
•	Mother's elder sister's husband	
•	Mother's sister	
•	Mother's sister's daughter	
•	Mother's sister's husband	
•	Mother's sister's son	
•	Mother's younger brother	
•	Mother's younger brother's wife	
•	Mother's younger sister	
•	Mother's younger sister's husband	
•	Sister (Born 2 nd or 3 rd or any middle no, not the eldest or youngest)	
•	Sister (elder)	
•	Sister (younger)	
•	Sister's daughter	
•	Sister's daughter (from female ego's side)	
•	Sister's daughter (from male ego's side)	
•	Sister's father in law	
•	Sister's mother in law	
•	Sister's son	
•	Sister's son (from male ego's side)	
•	Sister's son (from female ego's side)	

•	Son	
•	Son (Born 2 nd or 3 rd or any middle no, not the eldest or youngest)	
•	Son's father in law	
•	Son's mother in law	
•	Son's wife	
•	Son's wife's brother	
•	Son's wife's sister	
•	Spinster	
•	Step brother	
•	Step daughter	
•	Step father	
•	Step mother	
•	Step mother	
•	Step son	
•	Widow (female)	
•	Widow (male)	
•	Wife	
•	Wife's brother	
•	Wife's brother's daughter	
•	Wife's brother's son	
•	Wife's brother's wife	
•	Wife's elder brother	
•	Wife's elder brother's wife	
•	Wife's elder sister	
•	Wife's elder sister's husband	
•	Wife's husband (in case of two husbands)	
•	Wife's sister	
•	Wife's sister's daughter	
•	Wife's sister's husband	
•	Wife's sister's son	
•	Wife's younger brother	
•	Wife's younger brother's wife	
•	Wife's younger sister	
•	Wife's younger sister's husband	
•	Younger brother's wife	
•	Younger sister's husband	
•	Youngest Daughter	
•	Youngest Son	

8.1.3 Housing and Related

- The investigator should include terms for housing, related materials

used for housing or surrounding of houses such as backyard, bathroom, drain, sanitation etc.

- Kinds of houses, wall, thatch, pillar and pole, rooms, veranda, backyard etc should be collected.
- Collect at least words for following objects from daily life and use each of them in sentences

Sr. No.	Word (in English)	Word (in)
•	Balcony	
•	Bamboo	
•	Basement	
•	Bathroom	
•	Bedroom	
•	Bolt	
•	Brick	
•	Building	
•	Ceiling	
•	Cement	
•	Chimney	
•	Cistern	
•	Commode	
•	Courtyard	
•	Cowshed	
•	Dining room	
•	Door	
•	Door bell	
•	Drainage	
•	Drawing room	
•	Dressing room	
•	Fence	
•	Floor	
•	Foundation	
•	Frame of door/window	
•	Garage	
•	Garden	
•	Gate	
•	Grill	
•	Guest room	
•	Home	
•	House	
•	Hut	
•	Iron	
•	Jamb	
•	Kitchen	
•	Ladder	

•	Latch	
•	Lavatory	
•	Lawn	
•	Mezzanine floor	
•	Mud	
•	Nail	
•	Out house	
•	Pillar	
•	Pipe	
•	Pole	
•	Portico	
•	Roof	
•	Room	
•	Rope	
•	Sand	
•	Staircase	
•	Stone	
•	Store room	
•	Storey	
•	Study room	
•	Tank	
•	Tap	
•	Tent	
•	Terrace	
•	Thatch	
•	Tile	
•	Toilet	
•	Upper floor	
•	Ventilator	
•	Verandah	
•	Wall	
•	Wash basin	
•	Window	
•	Worship room	

8.1.4 Artefacts and Items of Daily Use

- Under this domain the researcher will collect names of cultural artifacts such as bamboo or cane basket, types of wooden hammer, winnowing tray, types of handicrafts etc.
- It also will include daily used items in the kitchen, bed rooms, kinds of seats, bed, pot etc or items outside the house.

Sr. No.	Word (in English)	Word (in)
---------	-------------------	-----------------

•	Almirah	
•	Arm chair	
•	Axe	
•	Bag	
•	Basket	
•	Bed	
•	Bench	
•	Blanket	
•	Book	
•	Box	
•	Broom	
•	Bucket	
•	Calendar	
•	Candle	
•	Chair	
•	Clock	
•	Colour pencil	
•	Comb	
•	Computer	
•	Cot	
•	Couch	
•	Cupboard	
•	Curtain	
•	Desk	
•	Diary	
•	Divan	
•	Door mate	
•	Doormat	
•	Dressing table	
•	Dustbin	
•	Duster	
•	Fan	
•	Furniture	
•	Glue	
•	Hair oil	
•	Hammer	
•	Hanger	
•	Husk	
•	Incense stick	
•	Key	
•	Lamp-stand	
•	Lantern	
•	Light	
•	Lock	
•	Mat	

•	Matchstick	
•	Mattress	
•	Mirror	
•	Mobile phone	
•	Mop	
•	Mug	
•	Needle	
•	News paper	
•	Note book	
•	Paper	
•	Pen	
•	Pencil	
•	Pestle	
•	Picture	
•	Pillow	
•	Pressing iron	
•	Rack	
•	Radio	
•	Rolling board	
•	Rolling Pin	
•	Rope	
•	Sandalwood	
•	Saw	
•	Scissor	
•	Shelf	
•	Show case	
•	Side Pillow	
•	Soap	
•	Sofa	
•	Study chair	
•	Study table	
•	Swing	
•	Table	
•	Telephone	
•	Television	
•	Thread	
•	Toothbrush	
•	Toothpaste	
•	Torch	
•	Torch	
•	Towel	
•	Umbrella	
•	Wardrobe	
•	Washing machine	
•	Watch	

Kitchen		
•	Bowl	
•	Chopper	
•	Container	
•	Cup	
•	Frying Pan	
•	Glass	
•	Greater	
•	Holder	
•	Jar	
•	Kettle	
•	Knife	
•	Ladle	
•	Lid	
•	Oven	
•	Peeler	
•	Plate	
•	Pot	
•	Pressure cooker	
•	Refrigerator	
•	Saucepan	
•	Saucer	
•	Spoon	
•	Stove	
•	Strainer	
•	Tea Pot	
•	Tiffin box	
•	Tray	
•	Tumbler	
•	Utensils	

8.1.5 Adornments and Costumes

- The investigator will collect terms of dress and ornaments natively used by them.
- The investigator should further note the difference between male and female dress.
- Always remember that names of attire and ornaments given below may not exist in the community.

Sr. No.	Word (in English)	Word (in)
•	Anklet	
•	Armlet	
•	Bangle	

•	Blazer	
•	Blouse	
•	Border	
•	Button	
•	Button hole	
•	Calf link	
•	Cap	
•	Casket	
•	Cloth	
•	Coat	
•	Cotton	
•	Dress	
•	Ear ring	
•	Fabric	
•	Finger ring	
•	Frock	
•	Gloves	
•	Gown	
•	Handkerchief	
•	Hat	
•	Ingot	
•	Lace	
•	Lingerie	
•	Locket	
•	Muffler	
•	Necklace	
•	Night suit/night dress	
•	Nose ring	
•	Nose-pin	
•	Ornament	
•	Pant	
•	Perfume	
•	Petticoat	
•	Pouch	
•	Sandal	
•	Saree	
•	Scarf	
•	Sew	
•	Shaving set	
•	Shawl	
•	Shirt	
•	Shocks	
•	Shoe	
•	Silk	
•	Skirt	

•	Slipper	
•	Spectacles	
•	Suit	
•	Sweater	
•	Tie	
•	Toe ring	
•	Towel	
•	Trouser	
•	Turban	
•	Wool	
•	Wrapper	
•	Wrist watch	
•	Wristlet	

8.1.6 Food and Related

- Under this domain the investigator will collect edible food items and related items that are known locally.
- He/she should collect the words for different types of milk products, oil, spices and the flour of different grains, kinds of pickle etc.

Sr. No.	Word (in English)	Word (in)
•	Alcohol	
•	Arrack	
•	Asafoetida	
•	Bamboo shoot	
•	Beef	
•	Betel leaf	
•	Betel nut	
•	Beverage	
•	Bitter	
•	Black pepper	
•	Bread	
•	Breakfast	
•	Butter	
•	Buttermilk	
•	Cake	
•	Candy	
•	Cardamom	
•	Chapati	
•	Cheese	
•	Chicken	

•	Chickpeas	
•	Chilly	
•	Cinnamon	
•	Clove	
•	Coconut oil	
•	Coffee	
•	Cooked rice	
•	Cooked vegetable	
•	Coriander	
•	Curd	
•	Curry	
•	Cutlet/Pakora	
•	Daliya	
•	Dessert	
•	Diner	
•	Dry fish	
•	Egg	
•	Fenugreek	
•	Fish	
•	Flattened rice (cura)	
•	Flour	
•	Flour gram	
•	Fried (potato/rice/fish)	
•	Ghee	
•	Grams	
•	Granules	
•	Gravy	
•	Honey	
•	Ilachi	
•	Jaggery	
•	Juice	
•	Khichdi	
•	Lentils	
•	Lunch	
•	Meal	
•	Meat	
•	Milk	
•	Milk-rice (kheer/payesh)	
•	Mustard	
•	Mustard oil	
•	Mutton	
•	Nutrela	
•	Oat	
•	Olive oil	
•	Paddy	

•	Paneer	
•	Papad	
•	Pickle	
•	Poppy seed	
•	Pork	
•	Prawn	
•	Previous night's rice (soaked in water)	
•	Puffed rice(muri)	
•	Pulses	
•	Raw vegetable	
•	Rice (uncooked)	
•	Roasted paddy(khoi)	
•	Salad	
•	Salt	
•	Sattoo	
•	Snacks	
•	Snauff	
•	Sooji (Rava)	
•	Soup	
•	Sour	
•	Sugar	
•	Sweet	
•	Taste (of food)	
•	Tea	
•	Turmeric	
•	Water	
•	Wheat	
•	White oil	

8.1.7 Health Ailments and Remedies

- Under this domain the researcher will collect mere words of health, ailments (physically and mentally) and words of traditional treatment.
- The researcher may however simultaneously, according to his or her convenience, inquire about the treatment process for each ailment although this will be dealt in ethno-linguistic section as well.

Sr. No.	Word (in English)	Word (in)
•	Ache	
•	Acidity	
•	Acne	

•	Allergy	
•	Blister	
•	Blood clotting	
•	Boil	
•	Cancer	
•	Cataract	
•	Chicken pox	
•	Cholera	
•	Cold	
•	Cough	
•	Diarrhea	
•	Disease	
•	Epilepsy	
•	Fever	
•	Flue	
•	Fracture	
•	Gout	
•	Headache	
•	Hiccup	
•	Hospital	
•	Illness	
•	Indigestion	
•	Injection	
•	Itching	
•	Jaundice	
•	Joint pain	
•	Limping	
•	Lunatic	
•	Malaria	
•	Measles	
•	Medicine	
•	Migraine	
•	Nursing home	
•	Pain	
•	Pimple	
•	Plaster	
•	Pregnancy	
•	Psychic	
•	Pus	
•	Rashes	
•	Saline	
•	Shoe-bite	
•	Sinusitis	
•	Stomach ache	
•	Stomach upset	

•	Surgery	
•	Swelling	
•	Tooth ache	
•	Tuberculosis	
•	Typhoid	
•	Vomit	

8.1.8 Religious and Rituals terms

- Under this domain, the investigator will collect terms that are related to religion and ritual activities.
- Investigator should always remember that terms which are mentioned here may not be available in certain community. For example, there will be a ceremony or a prayer time for the first meal (rice) of the harvest season.
- The investigator should note that these terms can be sensitive in certain context. Therefore, he/she should be aware of things around him before collecting them.

Sr. No.	Word (in English)	Word (in)
•	Alter	
•	Arti	
•	Auspicious	
•	Bell	
•	Bhajan	
•	Birth ritual	
•	Blessing	
•	Camphor	
•	Candle	
•	Church	
•	Cremation	
•	Death ritual	
•	Deity	
•	Demon	
•	Ember	
•	Fasting	
•	Festival	
•	First rice ceremony	
•	Flower offering	
•	Funeral pyre	
•	God	
•	Goddess	
•	Guggle	

•	Heaven	
•	Hell	
•	Holy Place	
•	Holy threading ceremony	
•	House warming ceremony	
•	Hymn	
•	Idol	
•	Immersion	
•	Incense stick	
•	Marriage	
•	Mosque	
•	Pedestal	
•	Penance	
•	Prasadam	
•	Prayer	
•	Procession	
•	Puja	
•	Religion	
•	Temple	
•	Tomb	
•	Worship	
•	Yajna	

8.1.9 Festivals and Related

People and communities individualism and ethnicity reflects through their practise and celebration of festivals. The investigator will include all types of festivals that are traditionally known to the community.

Sr. No.	Word (in English)	Word (in)
•	Festival	
•	Harvest festival	
•	Cattle festival	
•	Kite flying festival	
•	Holi	
•	Spring festival	
•	Fishing festival	
•	Forest festival	
•	Seed sowing festival	
•	Flower festival	

8.1.10 Music and its Instruments

Every communities show their individuality through their culture and music and the instruments they use for musical performance. The investigator should collect words of music and musical instruments and its types.

Sr. No.	Word (in English)	Word (in)
•	Beat	
•	Cassette	
•	Clarion	
•	Cymbals	
•	Dance	
•	Drum	
•	Flute	
•	Guitar	
•	Harmonium	
•	Hum	
•	Instrument	
•	Music	
•	Music player	
•	One-string Instruments	
•	Record-player	
•	Rhythm	
•	Song	
•	String Instruments	
•	Tabla	
•	Two-string Instruments	
•	Veena	
•	Violin	

8.1.11 Occupation and Related

- Under this domain the investigator will collect all the occupational terms existing in the community along with the gender distinction (if available) according to occupation.
- The investigator will also ask tools related to their occupation.

Sr. No.	Word (in English)	Word (in)
•	Actor	
•	Advocate	
•	Barber	
•	Basket maker	

•	Beggar	
•	Blacksmith	
•	Boatman	
•	Broker	
•	Businessman	
•	Businessman of betel leaves	
•	Businessman of metal pot	
•	Carpenter	
•	Clerk	
•	Cobbler	
•	Cook	
•	Coppersmith	
•	Dancer	
•	Director	
•	Doctor	
•	Driver	
•	Engineer	
•	Farmer	
•	Fisherman	
•	Gardener	
•	Gardener	
•	Goldsmith	
•	Hawker	
•	Hunter	
•	Industrialist	
•	Journalist	
•	Judge	
•	Labour	
•	Manager	
•	Milkmaid	
•	Milkman	
•	Money lender	
•	News reporter	
•	Nurse	
•	Officer	
•	Painter	
•	Pathologist	
•	Pharmacist	
•	Photographer	
•	Player	
•	Postman	
•	Potter	
•	Priest	
•	Professor	
•	Receptionist	

•	Recitor	
•	Restaurateur	
•	Sailor	
•	Sculpture	
•	Shopkeeper	
•	Singer	
•	Snake charmer	
•	Soldier	
•	Sorcerer	
•	Sweeper	
•	Swimmer	
•	Tailor	
•	Teacher	
•	Vendor	
•	Waiter	
•	Washer man	
•	Watchman	
•	Water distributor	
•	Weaver	
•	Writer	

8.1.12 Transport, Directions and Related

- Under this domain the investigator will collect terms and types of transport system such as in road, water etc.
- The terms people use for pointing out the direction will also be collected under this domain.

Sr. No.	Word (in English)	Word (in)
•	Aeroplane	
•	Airport	
•	Auto rickshaw	
•	Avenue	
•	Axel	
•	Axel pin	
•	Berth	
•	Bicycle	
•	Blind lane	
•	Boat	
•	Break	
•	Bridge	
•	Bullock cart	
•	Bus	

•	Car	
•	Cartage	
•	Chowk	
•	Coach	
•	Driver	
•	Engine	
•	Footpath	
•	Guard	
•	Highway	
•	Lane	
•	Level crossing	
•	Lorry	
•	Main road	
•	Metro rail	
•	Motor cycle	
•	Over bridge	
•	Passenger	
•	Path	
•	Pedestrian	
•	Platform	
•	Port	
•	Push cart	
•	Rader	
•	Rail line	
•	Rickshaw	
•	Road	
•	Road divider	
•	Seat	
•	Serpentine lane	
•	Ship	
•	Short cut	
•	Sledge	
•	Station	
•	Stoppage	
•	Taxi	
•	Ticket	
•	Tractor	
•	Traffic	
•	Traffic jam	
•	Traffic Signal	
•	Train	
•	Tram	
•	Vehicle	
•	Wheel	
Directions		

•	Back	
•	Direction	
•	Down	
•	East	
•	Front	
•	Left	
•	North	
•	North east	
•	North west	
•	Right	
•	Side	
•	South	
•	South east	
•	South west	
•	Up	
•	West	

8.1.13 Earth, Water and Celestial Body Related

- The investigator should collect words for Earth and its types of heavenly objects such as clouds, stars, moon and its stages, words for non living objects or items that are related to earth, mud, stone in this section.
- Terms of air and wind, fire and its related words etc, terms related to water resources both natural and artificial water resources will also be collected in this domain.

Sr. No.	Word (in English)	Word (in)
Celestial Bodies and Related		
•	Cloud	
•	Galaxy	
•	Lunar eclipse	
•	Milky way	
•	Moon	
•	Planet	
•	Rain	
•	Rainbow	
•	Satellite	
•	Seven star	
•	Sky	
•	Solar eclipse	
•	Star	
•	Sun	
•	Sunshine	

Earth and Related		
•	Bush	
•	Cave	
•	Clay	
•	Cliff	
•	Climate	
•	Desert	
•	Dust	
•	Earth	
•	Forest	
•	Glacier	
•	Ground	
•	Hill	
•	Island	
•	Monolith	
•	Mountain	
•	Mud	
•	Peak	
•	Plateau	
•	Rock	
•	Sand	
•	Slope	
•	Snow	
•	Soil	
•	Stone	
•	Valley	
•	World	
Water and Related		
•	Brook	
•	Bubble	
•	Canal	
•	Dew	
•	Drainage	
•	Drizzling	
•	Fountain	
•	High tide	
•	Lake	
•	Low tide	
•	Moisture	
•	Ocean	
•	Pond	
•	River	
•	Sea	
•	Stream	
•	Water	

•	Waterfall	
•	Wave	
•	Whirlpool	
Air and Related		
•	Air	
•	Fog	
•	Gas	
•	Wind	
Fire and Related		
•	Ass	
•	Camphor	
•	Coal	
•	Cracker	
•	Ember	
•	Fire	
•	Firewood	
•	Flame	
•	Forest fire	
•	Fuel	
•	Furnace	
•	Smoke	
Natural Calamities		
•	Cyclone	
•	Drought	
•	Earthquake	
•	Flood	
•	Hailstorm	
•	Landslide	
•	Lightening	
•	Snowstorm	
•	Storm	
•	Thunder	
•	Weather	
Metal Mineral and Related		
	Aluminium	
	Amalgamated	
	Brass	
	Bronze	
	Coal	
	Copper	
	Diamond	
	Glass	
	Gold	
	Iron	
	Lead	

Lime	
Magnet	
Marble	
Metal	
Mine	
Mineral	
Ore	
Pearl	
Silver	
Steel	
Stone	
Agriculture and Related	
Acre	
Axe	
Crop	
Farmer	
Harvest	
Land	
Plough	
Seasonal crop	
Seed	
Sickle	
Sow	
Hunting Fishing and Tools	
Arrow	
Bow	
Catapult	
Javelin	
Mud pellet	
Net	
Quiver	
Sling	
Spear	
Trap	
Watch tower	

8.1.14 Animals, Reptiles, Insects and related

- Under this domain the investigator should include all kinds of domestic animal names available in the native language.
- The investigator will note to see if there are different names for the same animal on account of its gender/sexual category and its stages of life.
- Include the names of young ones if any in this section itself.
- Cultural/professional significance of animal and their sounds

(presence or absence) should be noted for ethno-linguistic study.

Sr. No.	Word (in English)	Word (in)
Animals		
•	Animal	
•	Bear	
•	Bitch	
•	Buffalo	
•	Bull	
•	Calf	
•	Camel	
•	Cat	
•	Chick	
•	Cow	
•	Cub	
•	Deer	
•	Den	
•	Dog	
•	Elephant	
•	Foal	
•	Fox	
•	Goat	
•	Hen	
•	Hippopotamus	
•	Horse	
•	Jackal	
•	Kid	
•	Kitten	
•	Lion	
•	Lioness	
•	Mongoose	
•	Monkey	
•	Pig	
•	Piglet	
•	Puppy	
•	Rhinoceros	
•	Ship	
•	Tiger	
•	Tigress	
Reptiles Rodent		
•	Caterpillar	
•	Coral	
•	Crab	
•	Crocodile	

•	Frog	
•	Lizard	
•	Mouse	
•	Porcupine	
•	Rabbit	
•	Rat	
•	Reptile	
•	Seal	
•	Shell	
•	Snail	
•	Snake	
•	Squirrel	
•	Toad	
•	Tortoise	
Flies and Insects		
•	Ant	
•	Bee	
•	Butterfly	
•	Cockroach	
•	Conch	
•	Earthworm	
•	Fly	
•	Grasshopper	
•	Honey bee	
•	Housefly	
•	Insect	
•	Leech	
•	Louse	
•	Mosquito	
•	Scorpion	
•	Spider	
•	Wasp	
•	Worm	
Fish and Related		
•	Dolphin	
•	Fish	
•	Hilsa	
•	Katla	
•	Prawn	
•	Sea fish	
•	Sea goat	
•	Small Fish	
•	Whale	
Birds and Related		
•	Beak	

•	Bird	
•	Bulbul	
•	Chick	
•	Crest	
•	Crow	
•	Cuckoos	
•	Duck	
•	Feather	
•	Goose	
•	Heron	
•	Kingfisher	
•	Nest	
•	Owl	
•	Parrot	
•	Peacock	
•	Pigeon	
•	Sparrow	
•	Wing	
•	Woodpecker	
Animal's Body part		
•	Claw	
•	Horn	
•	Paw	
•	Tail	
•	Trunk	
•	Tusk	
Name of the sounds produced by Birds and Animals		
•	Chirp	
•	Crow	
•	Hiss	
•	Hoot	
•	Meow	
•	Roar	
Trees		
•	Bamboo	
•	Banana	
•	Banyan	
•	Betel Nut	
•	Coconut	
•	Deodar	
•	Eucalyptus	
•	Evergreen tree	
•	Jackfruit	
•	Mango	
•	Neem	

•	Oak	
•	Palm	
•	Peepal	
•	Pine	
•	Plant	
•	Tamarind	
•	Teak	
•	Tree	
•		
Climbers and Creepers		
•	Beans	
•	Bitter gourd	
•	Cane	
•	Cucumber	
•	Gourd	
•	Grapes	
•	Money plant	
•	Pea plant	
•	Pumpkin	
•	Snake gourd	
•	Watermelon	
Herbs Grasses and Weeds		
•	Basils	
•	Coriander leaf	
•	Grasses	
•	Herb	
•	Lavender	
•	Lemongrass	
•	Mint	
•	Shrubs	
•	Spinach	
•	Touch me not	
•	Weeds	
Flowers and Related		
•	Abloom	
•	Bely	
•	Bud	
•	Cosmos	
•	Daisy	
•	Flower	
•	Flower bush	
•	Hibiscus	
•	Jasmine	
•	Lily	
•	Lotus	

•	Marigold	
•	Nightqueen	
•	Petals	
•	Pollen	
•	Rose	
•	Sunflower	
•	Tube rose	
Vegetables		
•	Beans	
•	Beetroot	
•	Bell Paper	
•	Brinjal	
•	Cabbage	
•	Capsicum	
•	Carrot	
•	Cauliflower	
•	Corn	
•	Fig	
•	Garlic	
•	Ginger	
•	Green Chilly	
•	Green Papaya	
•	Lady's finger	
•	Lettuce	
•	Maze	
•	Mushrooms	
•	Onion	
•	Potato	
•	Radish	
•	Red Chilly	
•	Spinach	
•	Spring onion	
•	Sweet Potato	
•	Tomato	
•	Tuber	
•	Turnip	
•	Unripe Banana	
•	Unripe Jackfruit	
•	Vegetable	
Fruits and Seeds		
•	Almond	
•	Apple	
•	Berry	
•	Betel nut	
•	Cashew	

•	Chaya seed	
•	Coconut	
•	Date	
•	Fruit	
•	Grape	
•	Ground nut	
•	Guava	
•	Jackfruit	
•	Lemon	
•	Lime	
•	Litchi	
•	Mango	
•	Melon	
•	Nut	
•	Orange	
•	Palm	
•	Papaya	
•	Peer	
•	Pineapple	
•	Pomegranate	
•	Raisin	
•	Sapota	
•	Sprout	
•	Tender coconut	
•	Water melon	
Parts of Tree Fruit and Vegetable		
•	Bark	
•	Branch	
•	Leaf	
•	Node	
•	Root	
•	Stem	
•	Thorn	

8.1.17 Adjectives and Adverbs

- In this domain the investigator will find out the adjectival and adverbial words.
- This list will also find out if the language has different lexical items for a particular adjectival form.

Sr. No.	Word (in English)	Word (in)
•	Young	
•	Old	
•	Tall	

•	Short	
•	Fat	
•	Thin	
•	Thick	
•	New	
•	Fresh	
•	Ripe	
•	Rotten	
•	Curly	
•	Straight	
•	Curve	
•	Flat	
•	Narrow	
•	Huge	
•	Broad	
•	Good	
•	Bad	
•	Better	
•	Best	
•	Worse	
•	Worst	
•	Ugly	
•	Beautiful	
•	Strong	
•	Weak	
•	High	
•	Low	
•	Many	
•	More	
•	Much	
•	Little	
•	Small	
•	Big	
•	Rich	
•	Poor	
•	Expensive	
•	Costly	
•	Cheap	
•	Chief	
•	Blunt	
•	Sharp	
•	Light	
•	Smooth	
•	Soft	
•	Sweet	

•	Bitter	
•	Salty	
•	Sour	
•	Native	
•	Foreign	
•	Lively	
Adverb		
•	Fast	
•	Slow	
•	Soon	
•	Quickly	
•	Slowly	
•	Adversely	
•	Repeatedly	
•	Quietly	
•	Lately	
•	Often	
•	Now	
•	Finely	
•	Finally	
•	Extremely	
•	Severe	
•	Recently	
•	Solemnly	
•	Lovely	
•	Frankly	
•	Truly	
•	Perhaps	
•	Bravely	
•	Severe	
•	Nicely	
•	Humbly	

8.1.17 Stages of life

In this section various terms for stages of human life will be collected.

Sr. No.	Word (in English)	Word (in)
•	Adult	
•	Baby	
•	Baby (Female)	
•	Baby(Male)	
•	Boy	
•	Child	

•	Children	
•	Dying Person	
•	Dying Person (Female)	
•	Dying Person (Male)	
•	Female	
•	Foetus	
•	Girl	
•	Infant	
•	Infant (Female)	
•	Infant (Male)	
•	Male	
•	Man	
•	Middle Aged	
•	Middle Aged (Man)	
•	Middle Aged (Woman)	
•	Old Person	
•	Old Person (Female)	
•	Old Person (Male)	
•	Twin	
•	Twin (Female only)	
•	Twin (Male only)	
•	Virgin	
•	Woman	
•	Young Man	
•	Young Men	
•	Young Woman	
•	Young Women	

8.1.18 Colour Terms, Geometrical Shapes and Measurements

- In this section the investigator will collect the colour terms, geometrical shapes and measurement units known to the local people.
- The colour terms would be best collected with the sample colours.
- The geometrical shapes also would be best collected with the picture presentation of the shape.
- A particular culture might have different ways of measurement (i.e.-finger length, hand length, land's length.) The investigator should note these best in contextual application.

Sr. No.	Word (in English)	Word (in)
Colour Terms		
•	Black	

•	Blue	
•	Brown	
•	Green	
•	Grey	
•	Indigo	
•	Orange	
•	Pink	
•	Purple	
•	Red	
•	Violate	
•	White	
•	Yellow	
Geometrical Shapes and Sizes		
•	Angle	
•	Arc	
•	Centre	
•	Circle	
•	Cube	
•	Cylindrical	
•	Diagonal	
•	Diameter	
•	Half circle	
•	Hexagonal	
•	Octagonal	
•	Oval	
•	Pentagonal	
•	Periphery	
•	Rectangle	
•	Rombas	
•	Semi circle	
•	Square	
•	Triangle	
•	Volume	
Measurements		
•	Area	
•	Depth	
•	Distance	
•	Feet	
•	Fraction	
•	Grain	
•	Half	
•	Height	
•	Hour	
•	Inch	
•	Length	

•	Liquid	
•	Minute	
•	Moment	
•	Powder	
•	Quantity	
•	Quarter	
•	Second	
•	Solid	
•	Three-fourth	
•	Time	
•	Volume	

8.1.19 Numbers

- In this section the investigator will collect all numerical nouns of the language.
- The investigator should collect the pattern of counting numbers also while collecting the data.

Sr. No.	Word (in English)	Word (in)
Cardinal Numbers		
•	0	
•	1	
•	2	
•	3	
•	4	
•	5	
•	6	
•	7	
•	8	
•	9	
•	10	
•	11	
•	12	
•	13	
•	14	
•	15	
•	16	
•	17	
•	18	
•	19	
•	20	
•	21	
•	22	

•	23	
•	24	
•	25	
•	26	
•	27	
•	28	
•	29	
•	30	
•	31	
•	32	
•	33	
•	34	
•	35	
•	36	
•	37	
•	38	
•	39	
•	40	
•	41	
•	42	
•	43	
•	44	
•	45	
•	46	
•	47	
•	48	
•	49	
•	50	
•	51	
•	52	
•	53	
•	54	
•	55	
•	56	
•	57	
•	58	
•	59	
•	60	
•	61	
•	62	
•	63	
•	64	
•	65	
•	66	
•	67	

•	68	
•	69	
•	70	
•	71	
•	72	
•	73	
•	74	
•	75	
•	76	
•	77	
•	78	
•	79	
•	80	
•	81	
•	82	
•	83	
•	84	
•	85	
•	86	
•	87	
•	88	
•	89	
•	90	
•	91	
•	92	
•	93	
•	94	
•	95	
•	96	
•	97	
•	98	
•	99	
•	100	
•	200	
•	300	
•	400	
•	500	
•	600	
•	700	
•	800	
•	900	
•	1000	
•	10000	
•	100000	
•	1000000	

•	1000000	
•	10000000	
•	100000000	
•	1000000000	
Ordinal Numbers		
•	100 th	
•	10 th	
•	1 st	
•	20 th	
•	2 nd	
•	30 th	
•	3 rd	
•	40 th	
•	4 th	
•	50 th	
•	5 th	
•	60 th	
•	6 th	
•	70 th	
•	7 th	
•	80 th	
•	8 th	
•	90 th	
•	9 th	
•	Crores	
•	Lakhs	
•	Once	
•	Thrice	
•	Twice	

8.1.20 Time Season and Weather

- Every language has their individual terms for time, season, month and days.
- Each language has different terms for timing (i.e.-time according to their work).
- Each language determines their time and month following their own calendar and the movement of Sun, Moon, Star, Tide etc.
- The investigator should keep in mind the aforementioned points while collecting these data.

Sr. No.	Word (in English)	Word (in)
Time Season and Weather		

•	Afternoon	
•	Age	
•	April	
•	August	
•	Autumn	
•	Blue moon	
•	Century	
•	Chilly	
•	Cloudy	
•	Cloudy	
•	Cold	
•	Date	
•	Dawn	
•	Day	
•	Day	
•	Day after tomorrow	
•	Day before yesterday	
•	Decade	
•	December	
•	Dusk	
•	Evening	
•	Fall	
•	February	
•	Fortnight	
•	Friday	
•	Hot	
•	Hour	
•	Humid	
•	January	
•	July	
•	June	
•	Last month	
•	Last two years	
•	Last week	
•	Last year	
•	March	
•	May	
•	Mid day	
•	Mid night	
•	Minute	
•	Moment	
•	Monday	
•	Monsoon	
•	Month	
•	Month	

•	Morning	
•	Next month	
•	Next two years	
•	Next week	
•	Next year	
•	Night	
•	No moon	
•	Noon	
•	November	
•	October	
•	Rainy	
•	Saturday	
•	Season	
•	Second	
•	September	
•	Spring	
•	Summer	
•	Sunday	
•	Sunny	
•	Temperature	
•	Thursday	
•	Time	
•	Today	
•	Tomorrow	
•	Tuesday	
•	Two days after tomorrow	
•	Two days before yesterday	
•	Wednesday	
•	Week	
•	Week days	
•	Week end	
•	Windy	
•	Winter	
•	Year	
•	Yesterday	

8.1.21 Financial

In this section the investigator will collect the terms of money, finance, banking and its related words.

Sr. No.	Word (in English)	Word (in)
Money and Related		

•	Bank	
•	Bonus	
•	Borrow	
•	Budget	
•	Cash	
•	Coin	
•	Collection	
•	Credit	
•	Currency	
•	Debt	
•	Expenditure	
•	Fund	
•	Income	
•	Interest	
•	Loan	
•	Money	
•	Payment	
•	Rupee	
•	Saving	
•	Scheme	
•	Tax	
•	Transaction	
•	Transfer	

8.1.22 Sport, Entertainment and related

The terms for different kind of traditional and indigenous sports, games and entertainment of the community would be collected in this section.

Sr. No.	Word (in English)	Word (in)
Sports Games and Related		
•	Arrow	
•	Athlete	
•	Basket ball	
•	Bow	
•	Boxer	
•	Boxing	
•	Bull fight	
•	Chess	
•	Coach	
•	Cock fight	
•	Cricket	
•	Cycle	

•	Cyclist	
•	Dice	
•	Fighter	
•	Football	
•	Game	
•	Goli-danda	
•	Ground	
•	Hide and Seek	
•	High Jump	
•	Javelin	
•	Judo	
•	Kabaddi	
•	Karate	
•	Kite	
•	Lawn Tennis	
•	Long Jump	
•	Player	
•	Referee	
•	Shot-put	
•	Spinning	
•	Sports	
•	Stadium	
•	Swimmer	
•	Swimming	
•	Swimming Pool	
•	Table Tennis	
•	Toys	
•	Trainer	
•	Umpire	
Entertainment		
•	Arena	
•	Auditorium	
•	Camera	
•	Cinema	
•	Cinematography	
•	Costume	
•	Cultural function	
•	Culture	
•	Dance	
•	Drama	
•	Film	
•	Microphone	
•	Movie	
•	Movie hall	
•	Narration	

•	Performance	
•	Picture	
•	Play	
•	Puppet show	
•	Recitation	
•	Screen	
•	Song	
•	Sound	
•	Theatre	

8.1.23 Education and related

- In this section education and related terms would be collected.
- The investigator may find out the borrowed terms also while collecting the data.

Sr. No.	Word (in English)	Word (in)
•	Anthropology	
•	Archive	
•	Ball pen	
•	Bench	
•	Biology	
•	Black board	
•	Book	
•	Botany	
•	Calculator	
•	Certificate	
•	Chair	
•	Chemistry	
•	Class	
•	Class room	
•	College	
•	Computer	
•	Course	
•	Degree	
•	Desk	
•	Dissertation	
•	Document	
•	Duster	
•	Education	
•	Entrance Examination	
•	Eraser	
•	Examination	
•	Fountain pen	

•	Geography	
•	Highlighter	
•	History	
•	Holiday	
•	Illiterate	
•	Ink	
•	Inkpot	
•	Internet	
•	Journal	
•	Laboratory	
•	Language	
•	Library	
•	Life Science	
•	Linguistics	
•	Literate	
•	Magazine	
•	Marker	
•	Mathematics	
•	Note Book	
•	Paper	
•	Pen	
•	Pencil	
•	Period bell	
•	Physical Science	
•	Physics	
•	Psychology	
•	Recess	
•	Research	
•	Researcher	
•	Result	
•	Scholarship	
•	School	
•	School uniform	
•	Slate chalk	
•	Social Science	
•	Sociology	
•	Student	
•	Subject	
•	Supervisor	
•	Syllabus	
•	Table	
•	Teacher	
•	Thesis	
•	University	
•	Vacation	

•	Viva-Voce	
•	Vocational Course	
•	Website	
•	Zoology	

8.1.24 Classifier-Numeral- Determiner

The classifier-numeral-determiner list will try to find out the relation and agreement between classifier-numeral-determiner and the quantifier/ noun.

Sr. No.	Word (in English)	Word (in)
•	One boy	
•	Two Boys	
•	Three boys	
•	Four boys	
•	More than five boys	
•	Group of ten boys	
•	The boy	
•	All boys	
•	Many boys	
•	A number of boys	
•	Every boy	
•	Each boy	
•	Few boys	
•	A few boy	
•	Some boys	
•	Some boy	
•	No boy	
•	Almost every boy	
•	This boy	
•	That boy	
•	These boys	
•	Those boys	
•	Boy	
•	Boys	
•	The boy in the room	
•	Boy in the room	
•	Boys in the room	
•	The fourth boy arrived	
•	The eighth boy arrived	
•	One girl	
•	Two girls	

•	Three Girls	
•	Four girls	
•	More than five girls	
•	Group of ten girls	
•	Girl	
•	Girls	
•	The girl	
•	All girls	
•	Each girl	
•	Every girl	
•	Many girls	
•	Few girls	
•	A few girl	
•	Some girls	
•	Some girl	
•	No girl	
•	Almost every girl	
•	This girl	
•	That girl	
•	These girls	
•	Those girls	
•	Girl in the room	
•	The girl in the room	
•	Girls in the room	
•	The seventh girl arrived	
•	The tenth girl arrived	
•	One chair	
•	Two chairs	
•	Three chairs	
•	Four chairs	
•	More than five chairs	
•	Bunch of ten chairs	
•	The chair	
•	All chairs	
•	Every chair	
•	Each chair	
•	Many chairs	
•	Few chairs	
•	A few chair	
•	Some chair	
•	Some chairs	
•	No chair	
•	Almost every chair	
•	This chair	
•	That chair	

•	These chairs	
•	Those chairs	
•	Chair	
•	Chairs	
•	Chair in the room	
•	The chair in the room	
•	Chairs in the room	
•	I saw the third chair	
•	I saw the sixth chair	
•	One cat	
•	Two cats	
•	Three cats	
•	Four cats	
•	More than five cats	
•	Group of ten cats	
•	The cat	
•	All cats	
•	Every cat	
•	Each cat	
•	Many cats	
•	Few cats	
•	A few cat	
•	Some cat	
•	Some cats	
•	No cat	
•	Almost every cat	
•	This cat	
•	That cat	
•	These cats	
•	Those cats	
•	Cat	
•	Cats	
•	Cat the room	
•	The Cat in the room	
•	Cats in the room	
•	I saw the second cat	
•	I saw the ninth cat	
•	One tree	
•	Two trees	
•	Three trees	
•	Four trees	
•	More than five trees	
•	Rows of tree	
•	The tree	
•	All trees	

•	Every tree	
•	Each tree	
•	Many trees	
•	Few trees	
•	A few tree	
•	Some tree	
•	Some trees	
•	No tree	
•	Almost every tree	
•	This tree	
•	That tree	
•	These trees	
•	Those trees	
•	Tree	
•	Trees	
•	The tree beside my room	
•	Tree beside my room	
•	Trees beside my room	
•	I saw the first tree	
•	I saw the seventh tree	
•	One book	
•	Two books	
•	Three books	
•	Four books	
•	More than five books	
•	Bunch of ten books	
•	Book	
•	The book	
•	All books	
•	Every book	
•	Each book	
•	Many books	
•	Few books	
•	A few book	
•	Some book	
•	Some books	
•	No book	
•	Almost every book	
•	This book	
•	That book	
•	These books	
•	Those books	
•	Books	
•	The book on the table	
•	Book on the table	

•	Books on the table	
•	I read 1/5 th of the book	
•	I read the last book of the series	
•	One day	
•	Two days	
•	Three days	
•	Four days	
•	More than five days	
•	Couple of days	
•	Day	
•	The day	
•	All days	
•	Every day	
•	Each day	
	Many days	
	Few days	
	A few day	
	Some day	
	Some days	
	No day	
	Almost every day	
	This day	
	That day	
	These days	
	Those days	
	Days	
	Day before Sunday	
	The day before Sunday	
	Days before war	
	I came on the third day	
	I came on the last day	
	Sugar	
	One spoon sugar	
	Some sugar	
	A lot of sugar	
	1kg sugar	
	1/3 rd of the sugar	
	Approximately 5kg of sugar	
	Oil	
	1litre oil	
	Some oil	
	A lot of oil	
	7litre oil	
	1/8 th of the oil	
	One spoon oil	

	Almost 9 litre oil	
--	--------------------	--

8.1.25 Verb

The Verb list will contain the most daily used verbal items as well as to check how compound verb formation and causative verb formation take place in that particular language.

Sr. No.	Word (in English)	Word (in)
	Abuse	
	Accept	
	Act	
	Add	
	Adopt	
	Advise	
	Agree	
	Alert	
	Answer	
	Appear	
	Arrange	
	Arrive	
	Ask	
	Assure	
	Attract	
	Bake	
	Bark	
	Bathe	
	Beat	
	Belch	
	Believe	
	Bend down	
	Bind	
	Bite	
	Blast	
	Bleed	
	Bloom	
	Bluff	
	Blush	
	Board	
	Boil	
	Book (a ticket)	
	Born	
	Bounce	

	Break	
	Breathe	
	Bring	
	Build	
	Burp	
	Burst	
	Buy	
	Call	
	Catch	
	Chant	
	Cheat	
	Check	
	Chew	
	Clap	
	Click	
	Climb	
	Close	
	Collect	
	Collect	
	Come	
	Command	
	Compose	
	Compute	
	Confirm	
	Confront	
	Console	
	Cook	
	Cough	
	Count	
	Cross	
	Cry	
	Cut	
	Dance	
	Deal	
	Decide	
	Decorate	
	Deduct	
	Delete	
	Demand	
	Demonstrate	
	Departure	
	Dictate	
	Die	
	Differentiate	
	Dig	

	Disagree	
	Disappear	
	Discriminate	
	Disown	
	Dissolve	
	Distinguish	
	Distract	
	Ditch	
	Divide	
	Do	
	Doubt	
	Draw	
	Dream	
	Drink	
	Drive	
	Drop	
	Eat	
	Elevate	
	Erase	
	Escape	
	Evaporate	
	Excrete	
	Exercise	
	Exhale	
	Express	
	Fall	
	Fart	
	Fast	
	Feed	
	Feel	
	Felicitate	
	Fetch (water)	
	Fight	
	Fill	
	Find	
	Float	
	Fly	
	Forget	
	Get	
	Give	
	Go	
	Gossip	
	Greet	
	Grind	
	Hammer	

	Hang	
	Hate	
	Hear	
	Hesitate	
	Hiccup	
	Hide	
	Ignore	
	Illuminate	
	Imagine	
	Impress	
	Include	
	Inhale	
	Insult	
	Invite	
	Irritate	
	Joke	
	Jolt	
	Jump	
	Kill	
	Kiss	
	Kneel down	
	Knit	
	Know	
	Laugh	
	Leak	
	Lean	
	Leap	
	Learn	
	Lie	
	Lie down	
	Like	
	Listen	
	Lock	
	Look	
	Love	
	Make	
	Mark	
	Marry	
	Meet	
	Melt	
	Mix	
	Mop (the floor)	
	Move	
	Multiply	
	Murmur	

	Nail	
	Narrate	
	Notch	
	Note down	
	Notice	
	Observe	
	Open	
	Operate	
	Order	
	Own	
	Pack (bag)	
	Paint	
	Pant	
	Paste	
	Paw	
	Peep	
	Perceive	
	Pick	
	Pinch	
	Plant (a tree)	
	Play	
	Plough (the land)	
	Pluck	
	Point out	
	Pour	
	Practise	
	Pray	
	Prefer	
	Pretend	
	Prey	
	Promise	
	Pull	
	Punch	
	Push	
	Rain	
	Reach	
	Read	
	Recite	
	Refuse	
	Regret	
	Reject	
	Remember	
	Remind	
	Remove	
	Repeat	

	Reply	
	Resign	
	Respond	
	Ride	
	Ring	
	Ripe (as a fruit)	
	Rise	
	Rob	
	Roll	
	Rub	
	Run	
	Sack	
	Save	
	Say	
	Scold	
	Score	
	Scratch	
	Scream	
	Scribble	
	Scribe	
	Search	
	See	
	Seek	
	Sell	
	Send	
	Serve	
	Shake	
	Shave	
	Shiver	
	Shoot	
	Shout	
	Show	
	Shut	
	Sing	
	Sip	
	Sit	
	Slap	
	Sleep	
	Slip	
	Smash	
	Smell	
	Smile	
	Smoke	
	Sneeze	
	Snore	

	Soak	
	Speak	
	Stammer	
	Stand	
	Steal	
	Step	
	Stitch	
	Stop	
	Stumble	
	Submit	
	Subtract	
	Suck	
	Summon	
	Surrender	
	Survey	
	Sweat	
	Swell	
	Swim	
	Take	
	Talk	
	Taste	
	Teach	
	Tell	
	Think	
	Thrash	
	Throw	
	Tie (a knot)	
	Touch	
	Travel	
	Trim	
	Tuck in	
	Tweak	
	Type	
	Understand	
	Unite	
	Urinate	
	Utter	
	Vomit	
	Wait	
	Waive	
	Walk	
	Wander	
	Warn	
	Wash (clothes)	
	Wash (dishes)	

	Wear	
	Weave	
	Wed	
	Weep	
	Whistle	
	Wither	
	Work	
	Worship	
	Wrap	
	Write	
	Yawn	
Causative Verb		
	Make to feed	
	Make to sleep	
	Make to run	
	Make to laugh	
	Make to cry	
	Make to Cook	
	Make to cut	
	Make to dance	
	Make to sing	
	Make to bathe	
	Make to speak	
	Make to say	
	Make to fight	
	Make to beat	
	Make to read	
	Make to write	
	Make to like	
	Make to sit	
	Make to stand	
	Make to open	
	Make to close	
Compound Verb		
	Arrive (pauch gya)	
	Build (bana diya)	
	Come (aa cuka)	
	Decorate(saja diya/saja liya)	
	Demonstrate (karke dikhaya)	
	Departure (choR diya)	
	Do (kar cuka)	
	Eat (kha cuka)	
	Fly away (urr gya)	
	Give (de diya)	
	Go (cal paRe)	

	Go away (bhag cuka)	
	Kill (mar diya)	
	Laugh (has paRe)	
	Look (dekh liya)	
	Read (paR liya)	
	Say (bol baiThe)	
	Send (bhej diya)	
	Sit (baiTh cuka)	
	Stop (bandh kiya/ruk gya)	
	Take (le liya)	
	Throw (fek diya)	
	Write (likh diya)	

SENTENCE LIST

8.2.1. Small Sentences:

A small sentence has a very simple structure, consisting of only one clause with one finite verb, single subject and the object is optional. These types of sentences are basic syntactic units of any language.

Sr. no.	Sentence (in English)	Sentence (in....)
1	I do	
2	You eat rice	
3	You (non honorific) go home	
4	You (honorific) cook rice	
5	He does	
6	She does	
7	They do	
8	Ram saw a tiger	
9	Mary saw a tiger	
10	They saw a tiger	
11	I ate	
12	I came yesterday	

8.2.2. Compound Sentences

These types of sentences are basically more than one sentence conjoined together with and/or/but etc.

Sr.	Sentence (in English)	Sentence (in....)
-----	-----------------------	-------------------

no.		
1	I came back and she started	
2	Mina and Nina are two dear friend	
3	Either I will go or my father	
4	Hima and Shashi will come together	
5	I will go if he will come along	
6	Either Ira will drive or we will call a cab	
7	Neither my parents nor my friends like the movie	
8	My mother will never agree nor my father	
9	Sima will come but won't bring her friends	
10	Mangoes are not sweet this time because there is not enough rain	

8.2.3. Complex Sentences

These sentences contain one main clause and one or more subordinate clause/s.

Sr. no.	Sentence (in English)	Sentence (in....)
1	I don't know what Mita told Rima	
2	Rita does not know how to fix the problem	
3	Sovana does not know when will Sipra come	
4	Roni has no idea where Anup lives	
5	The girl who is wearing a pink salwar is my sister	
6	The boy whom she gave a bag is her little brother	
7	The man who has long beard is the Principal	
8	That woman near the chair is my aunt	
9	The book that I lost is costly	
10	Men who work are happier than man who don't work	

8.2.4. Cases

Cases are grammatical categories which reflect the semantic functions performed by subject, object and take inflected forms according to their functions in sentence.

Nominative

1	I ate rice	
2	You ate an apple	
3	He eats a mango	
4	She ate the whole packet of biscuit	
5	They ate a lot	
6	Today is Monday	

Accusative

1	Mita loves sita	
2	He killed the rat yesterday	
3	Ram reads book	
4	The dog bite the man	
5	Rekha wrote a letter to Keka	

Instrumental

1	He cut the onion with a knife	
2	Sekhar painted the wall with a brush	
3	Nita cut her nail with a nail cutter	
4	The picture was painted by Mr. Alam	
5	The building collapsed due to earthquake	

Locative Case

1	My home is near the river	
2	They live in California	
3	There is a haunted house near the big market	
4	Mika loves to dance in rain	
5	Tabsum kept the book on the table	
6	There is a river near my home	

Dative

1	Go to my home	
2	Students have gone to school	
3	They will come to my home	
4	The tourists have gone to jungle	
5	Give rice to Deepak	

Ablative

1	The baby fell down from the bed	
2	The boy jumped from the tree	
3	Leaves fall from the tree	
4	I fell down from the bike	
5	Koli will take out the bucket from the well	

Possessive

1	Rimi has a pet cat	
2	She kept her notebook on the round top table	
3	Keka gave Rekha her ear ring	
4	Grandma has given her all money for charity	
5	The colour of my bag is red	

8.2.5. Passive Construction

A passive construction is sentential construction where the subject takes the role of an object and the object becomes the subject and the verb form changes as needed.

Sr. no.	Sentence (in English)	Sentence (in....)
1	I was told to leave the place before sunset	
2	The room was cleaned by Hirak	
3	The crops were smashed by group of elephants	
4	The board was cleaned by a duster	
5	The Moon was covered by cloud	
6	We have been told by the man that the village is 1km away from the bus stand	
7	The driver was stopped by a group of robber	
8	Rita was beaten by her mother	
9	All the houses are decorated by earthen lamps	
10	The city will be covered by hoardings	

8.2.6. Reflexive and Reciprocals

Sr. no.	Sentence (in English)	Sentence (in....)
1	I myself finished the food	
2	We saw ourselves in the mirror	
3	He himself washes his car	
4	Keka talked to herself	
5	Jonaki's father's picture of himself is very attractive	
6	She herself has bought a car	
7	The door opened by itself	
8	They talked a lot about themselves	
9	They met each other at fair ground	
10	We respect each other	

8.2.7. Quantifiers and Intensifiers

A quantifier is word which expresses the amount or quantity of a noun while the intensifier modifies the role of the noun.

Sr. no.	Sentence (in English)	Sentence (in....)
1	No one is dishonest here	
2	Nothing comes free	
3	Everyone is honest here	
4	Everything is very expensive	
5	Everybody has their own way of living	
6	All is well	
7	No one has reached on time	
8	Someone will come to help her	
9	Nobody came to help her	
10	Somebody will come to pick you up	
11	No one would support him	
12	Something wrong will happen tonight	
13	I was very hungry in the morning	
14	You are too late to reach on time	
15	I truly trusted her	
16	Bring sweets as many as you can	
17	I don't like much sugar in tea	
18	Hidam does not like this place anymore	
19	Sikha really worked hard to pass the exam	
20	The child is holding his father's hand firmly	

8.2.8. Tense/Aspect/Mood

Tense is a category which expresses time reference of an event or an action while aspect expresses how an action or an event relates to the time frame. Mood, in a sentence conveys the speaker's attitude while uttering the sentence.

8.2.8.1. Tenses

Let us look into various forms of tenses here:

Present Tense (Singular)

1	I love dogs	
2	Mani cooks rice	
3	You see him	
4	Roma sleeps	
5	He wears blue jeans	
6	She dances well	
7	Raka has a pet cat	
8	The boy runs after kite	
9	He gives me a rose everyday	
10	The cat eats fish	

Present Tense (Plural)

1	We read books	
2	They love playing cricket	
3	We play football	
4	Old men walk in the morning	
5	Young ladies watch movies	
6	Cows give milk	
7	Birds fly	
8	Boys go to school	

Present Tense (With Negative)

1	I do not eat fish	
2	He does not like me	
3	She does not go to market	
4	We do not come home daily	
5	They do not read newspaper	
6	Kids do not love milk	
7	Fishes can't fly	
8	You do not work	
9	Rita can't write Bangla	
10	Bina does not sleep early	

Past Tense (Singular)

1	I cooked meat yesterday	
2	You went to Puri last summer	
3	He liked the movie	
4	She played guitar very well in the concert	
5	The meeting was over in an hour	
6	The boy knocked the door	
7	The girl fell down	
8	The cat chased after the mouse	
9	The tiger ate the rabbit	
10	The man caught the fish	

Past Tense (Plural Tense)

1	We went for a field trip in last winter	
2	You demonstrated the work in detail	
3	They danced very well	
4	Kids had a long vacation	
5	Delegates reached late for the meeting	
6	Dogs chased after the thief	
7	The robbers managed to escape	
8	The protestors took out a rally	
9	The journalist took Bacchan's interview for an hour	
10	The linguists archived an endangered language	

Past Tense (With Negative)

1	I did not come late	
2	You could not finish the work on time	
3	He did not eat the whole biriyani	
4	She did not painted the wall	
5	Mother did not make fish yesterday	
6	Father did not scold the young boy	
7	We did not catch the train	
8	They could not play chess	
9	Dogs could not chase the cat	
10	Monkeys did not attack the rat	

Future Tense (Singular)

1	I will buy a book tomorrow	
---	----------------------------	--

2	He will give me the ticket	
3	She will return from home on next monday	
4	You will finish the work in an hour	
5	Teacher will not help the student to solve the problem	
6	The little girl will throw a party on her birthday	
7	Train will start from New Delhi station	
8	My sister will come tomorrow	
9	It will rain today	
10	The meeting will start at 10.30am	

Future Tense (Plural)

1	We will attend the meeting on 22 nd February	
2	Three of us will go there together	
3	They will book the ticket tomorrow	
4	You will start for Delhi on 25 th February	
5	My friends will visit the university for the seminar	
6	Their parents will come here for a short trip	
7	We shall be there to help them	
8	The team will assist us to finish the project	

Future Tense (Negative)

1	I will not be able to attend the party	
2	They will not help us to find the bird	
3	My brother will not come home in this vacation	
4	The train will not start on time	
5	We will not buy anything for him	

8.2.8.2. Aspect:

Aspect is a grammatical category that expresses how a verb relates an action, event or state to the flow of time.

Aspect (Present Progressive)

1	I am not driving the car	
---	--------------------------	--

2	She is travelling from Kashmir	
3	We are attending the lecture series	
4	Mita is stitching a table-cover	
5	Mother is baking a cake	
6	Roson is not reading the newspaper	

Aspect (Past Progressive)

1	Uncle was ploughing the garden	
2	The baby was sleeping in the evening	
3	We were not attending the class	
4	Players were not playing cricket	
5	The man was reading a detective story when I called him	
6	The teacher was teaching Linguistics in the last semester	
7	You were not helping the delegates to find out the venue	
8	Milky was sleeping all the day yesterday	

Aspect (Perfective)

1	Ritu has gone to Mumbai	
2	Mithu had gone to Puri in July	
3	I have seen the missing dog	
4	Abhra had not finished the course on time	
5	She had called me around 7.30 am	
6	We have already booked our hotel	

Aspect (Habitual)

1	Itu used to love egg	
2	I used to jog in morning	
3	She used to smoke in younger age	
4	Dipak usually stays at Sukanta's place	
5	I don't watch TV usually	
6	Kids usually love ice-cream	

8.2.8.3. Mood

Imperative

Imperative mood expresses the speaker's order, comment, permission, prohibition, and request towards the hearer.

Sr. no.	Sentence (in English)	Sentence (in....)
1	You (non honorific) do	
2	You do	
3	You (honorific) do	
4	Please do it alone	
5	Remove your shoes	
6	You please bring your laptop	
7	Sit here	
8	Vidhya! Come here	
9	Don't pluck the flower	
10	Do finish your work quickly	
11	Maintain silence inside the school	
12	Don't go	

Subjunctive/ Conditional

Subjunctive mood expresses a condition which is doubtful, not factual or an undesired possibility.

Sr. no.	Sentence (in English)	Sentence (in....)
1	Hima entered into the room silently so that the baby won't wake up	
2	I hope that he would be able to cross this hurdle	
3	I am afraid that Rana would be elected as the director	
4	I am afraid India will lose this game	
5	Nupur wants that we go to her house tonight	
6	It is my desire that India will win the game today	
7	I have come so that she will go back home with me today	
8	If Mona does not come today her mother would be very angry	
9	If it will be sunny today then the day will be warmer	
10	Rina came with a musk so that the kid will be afraid	

Interrogative Sentences

These sentences express questioning mood of subject. These types of sentences either contain a question word or express the

interrogative mood by changing the word order and intonation.

Sr. no.	Sentence (in English)	Sentence (in....)
1	What is Sita's father's name?	
2	Who is her brother?	
3	What is this?	
4	Who is your brother?	
5	What did Tiku give to Miku?	
6	Where is your house?	
7	Where do you stay?	
8	How did you get the book?	
9	What is the price of onion?	
10	How much you pay for your tuition fee?	
11	How did you repair the tap?	
12	When will Sima come?	
13	Why are you so angry?	
14	What is your hobby?	
15	When did Rima leave for home?	

8.2.9. Yes/No Sentences

These types of sentences are asked to know the acceptance or denial/rejection of the addressee or of a fact.

Sr. no.	Sentence (in English)	Sentence (in....)
1	Does Abhi eat fruit?	
2	Would she come today?	
3	Has she written an essay on Indian Geography?	
4	Did Bithi meet Mona?	
5	Yes Abhi eats fruit	
6	No she has not written the essay on Indian Geography	
7	Yes Bithi met Mona day before yesterday	
8	Bithi met Mona day before yesterday evening near the river bank	
9	No she will not come today	
10	The flower is beautiful, is not it?	
11	Rita is not coming today, is she?	
12	Itu does not come everyday, does she?	
13	Does she have sister also?	
14	Are you Sita?	
15	Is she awake till now?	

16	Do you have any degree in Linguistics?	
17	Do you know anything about Rabindranath?	
18	Does she have sister also?	
19	She does not have any sister or brother	
20	Shall I come for the meeting?	

8.2.10. Exclamatory

An exclamatory sentence expresses the strong emotion of the speaker towards an event.

1	What a cute child he is	
2	How many people have died in this accident	
3	Ohh, what a bad weather	
4	Alas, the principal died	
5	How beautiful the place is	
6	Ohh how did he manage the job	
7	Ouch I fell down from the bed today	
8	Oops I did a bad mistake	
9	Ohh god what an achievement this is	
10	I wonder how has he tolerated so much pain	

8.2.11. Negative and Prohibitive

These types of sentences contain a negative particle or a negative word which makes an affirmative sentence to a negative one.

Prohibitive sentences express the prohibition of an act.

Sr. no.	Sentence (in English)	Sentence (in....)
1	She does not eat rice	
2	Rice, she does not eat	
3	I don't know him	
4	Raka does not like watching movie	
5	Divya has not slept yet	
6	I do not have that book	
7	He is not a poet	
8	He left the place without eating	
9	Let us not eat now	
10	They are not here	
11	I never have gone to Agra	
12	I will eat neither chicken nor mutton	
13	Neither Babita will come nor Sabita	
14	Neither Raka will come nor Rita will come	

15	Not everyone came	
16	Not even my brother came	
17	Until Bobby reaches I will wait here	
18	Will you eat or not?	
19	It is not the case that Suren will fail the exam	
20	It is not that Rama will come tomorrow	
21	I don't think it will rain today	
22	I do think it will not rain today	
23	Smoking is prohibited here	
24	Parking car is strictly forbidden here	
25	She cannot eat meat	

8.2.12. Comparative and Superlative

Comparative constructions are sentences where comparison between two different entities or objects takes place. Intensifier or comparative adjectives are being used in this construction. Superlative constructions are constructions consisting of a superlative adjective which indicates the upper or lower limit of a quality or quantity among three or more entities or objects

1	Uncle walks slowly	
2	Train runs fast	
3	Car runs faster than train	
4	Soumya is a better cook than Roma	
5	Sona is best among all student	
6	Sheema is more clever than Hima	
7	Rekha is most beautiful in the class	
8	Mosquito is a worst creature	
9	Leela is an worse movie than Hate Story	
10	My paper was worst among all	
11	Ishani's presentation was most flawless one	
12	Ganga is the largest river in India	

8.2.13. Adjectival Clauses

Adjectival/adjective/relative clause is a dependent or subordinate clause introduced either by a relative pronoun or by a relative adverb. The entire clause functions as an adjective that qualifies or quantifies a noun.

1	The big house is Mr. Roy's	
2	The red book is mine	
3	The small child is my nephew	

4	The white laptop I bought 4years back	
5	He bought the new black mobile last week	
6	My small house was over crowded during puja	
7	Hima ate the big burger	
8	That cute baby is my neighbour's daughter	
9	I am lazy	
10	The green fruits are good for eyes	

8.2.14. Adverbial Clauses

An adverbial clause is a dependent or subordinate clause which functions as an adverb and the entire clause modifies either a verb or a noun or an adjective or another adverb.

1	Come on time	
2	The delegate will come from Delhi	
3	Father bought a toy for his son	
4	Come at 5 O'clock	
5	Come with your parents for the function	
6	The Principal would love to meet every student's parents	
7	I came to Kolkata from Bolpur	
8	I used to have a pet cat 10 years back	
9	Rima stays in this house since last 5 years	
10	Uma has fever from Wednesday	
11	Ira will stay here till 4 O'clock	
12	She has not reached yet	
13	As long as it will dark I will stay here	
14	As soon as the bell rang student rushed towards the class	
15	No sooner had the teacher entered the entire class became silent	
16	They finished eating before we ate	
17	The book will be inaugurated after the chief guest will arrive	
18	Our department is behind the jungle	
19	The aeroplane is above the cloud	
20	The house is in front of the church	

8.2.15. Infinitive

1	I want to come	
---	----------------	--

2	They went to play without eating	
3	He ought to come	
4	He had to go	
5	She need to finish it by today	
6	Pratibha loves to eat fruit	

Transitive

1	I love flower	
2	Ram sees Sita	
3	Rabindranath Tagore wrote Gitanjali	
4	She eats egg	
5	Gargee does not like rice	

Intransitive

1	I sleep late	
2	The child fell down	
3	I was waiting at the main gate	
4	The kid stands there	
5	My father snores	

Ditransitive

1	I gave a book to Rima	
2	Nadiya will write a letter to Chomsky	
3	He took the letter from the post office	
4	Raina will give Raka her stamp box	
5	Mother did not give the kid the chocolate	

8.2.16. Complex Predicate Construction

A complex predicate construction consists of two or more predicates, which give the meaning of the main predicate only. The term predicate here refers to Noun and Adjective also along with verb.

1	Nita fell down from the tree	
2	Sikha broke off her engagement	
3	Labours call off their strike	
4	I have eaten up the whole rice	
5	They have lighten up the house	
6	Go and see	

7	Have a look and take	
8	The bird flew away	
9	Mira cleaned up the room	
10	The yogi greets the Sun every morning	

8.2.17. Causative

Causative constructions are constructions where one action or event is not being done by the subject but being done by someone else. These sentences contain one main verb and the causative verb or the main verb is inflected by a causative particle.

1	Mother fed the baby	
2	Father made the mother feed the baby	
3	I made the man to clean the bathroom	
4	I made the man to ask his wife to cook for us	
5	Mother makes the kid to sleep on time	
6	Tomba made Kiro make Rambha sleep	
7	Dipak asked Dipika not to eat chicken	
8	Dipak made the Dipika not to eat chicken	
9	Dipak forced Dipika not to eat chicken	
10	Lipika will make the teacher to demonstrate the theory to students	

8.2.18. Gerundial Construction/Verbal Noun

Gerund is a verbal noun which plays the role of noun within a noun phrase construction. In English, gerund is formed by adding -ing to the verb root, e.g., smoke+ing=smoking.

1	Smoking is injurious to health	
2	Gardening is my hobby	
3	Obedience to traffic rule is a compulsion for every citizen.	
4	Swimming is good for health	
5	Reading is a good habit	

9. Specifications of Tools

Audio Recorder

This should be a .WAV file recorder with following capacity:

24-bit/48kHz (or 44.1kHz) uncompressed recording quality.

Up to 320 KBPS MP3 playback and recording capacity.

SD or SDHC memory card compatibility with up to 16 GB storage.

High-grade stereo condenser microphone built in Mic and Line audio inputs.

Easy operation, user-friendly graphic display quality.

Ultra portability.

Long battery life.

Video Recorder

This should be a full HD recorder with following capacity:

3 inch LCD Monitor.

4.14 Megapixel Camera.

MOS Sensor 49 mm Filter Diameter HDMI Support f/1.8 - f/5.6

Battery Chargers

Battery charger is a basic backup tool, which should be necessary for field work with compatible with various tools like:

Audio recorder

Video recorder

Laptop/Tablet

Battery backups

Storage Devices

This secondary storage device will be needed for backup storage support for audio and video recorders. An external portable hard disk having following features will be required:

Up to USB 3.0 connectivity.

1 TB storage capacity.

Formatibility.

Other

These tools and stationary may be suit at the time of field work as per requirement of weather and location etc.

10. Storage and Reporting of Output

The issue of storage during field work has always been important step towards drawing conclusion. A linguistic field work will have various types of ingredients for result point of view. The main items of product and its storage feature are given bellow:

Image

The images captured during field work will be focused on related speech communities and its social and cultural events. The file format of images should be .PNG and .JPEG with resolution 19201080.

Audio

The audio recorded based on dataset or questionnaire of surveyor will be main asset of field work and simultaneously it is most useful resource for linguistic analysis. Thus sound recording environment should be noise free and natural. The format of audio should be .WAV or .MP3.

Video

As output of field work, video will be focused on ethno-linguistic, cultural and ecological features of targeted speech community. The specification of video should be full HD with video mode 1440p, frame size in pixels (WH) 2,560×1,440 and pixels per image 3,686,400.

Text

The investigator should type the every Indian script in Unicode and English in New Times Roman font with font size 14.

11. Ethical Issues

Each survey should begin with an ethical statement. Each and every informant should be informed about the future use of the data he/she provides. Their approval/permission—either audio-visually or in black and white—should be kept on record for publishing the data. Responsibility of errors should be borne by the transcribers.

To work in remote and politically sensitive areas—like Nagaland, Manipur, etc.—work permit has to be procured from the competent authorities. In some of these areas use of GPS may not be allowed—digital/visual recording would be the only option.

References

- Abbi A. 2001. *A Manual of Linguistic Field Work and Structures of Indian Languages*. Lincom Europa.
- Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Structures of Indian Languages*. Muenchen: Lincom Europa.
- Bowern, Claire. 2008. *Linguistic Fieldwork: A Practical Guide*. New York: Palgrave Macmillan.
- Census of India. 2011. "Census of India 2011: Towards a bright future". <http://www.languageinindia.com/feb2014/censusofindia2011data.pdf> (accessed 18 February 2016).
- Comrie B. and Smith N. 1977. "Lingua Descriptive Studies: Questionnaire". *Lingua* 42:1-72.
- Dorian, Nancy. 1998. 'Western language ideologies small-language prospects' in Grenoble, L.A & L.J. Whaley, eds. *Endangered Languages: Current Issues and Future Prospects*. Cambridge: Cambridge University Press.
- Fishman, J. A., 1991, Reversing language shift, Clevedon, UK, Multilingual Matters Ltd.
- Gordon, Raymond G., Jr. (ed.), 2005. *Ethnologue: Languages of the World*, Fifteenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com/15>.
- Grimes, Barbara F. (ed.), 2000. *Ethnologue: Languages of the World*, Fourteenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com/14>.
- Kibrik E. 1977. *The Methodology of Field Investigations in Linguistics (Setting up the Problem)*. Mouton, The Hague-Paris.
- Lewis, M. P. and G. F. Simons. 2010. "Assessing endangerment: Expanding Fishman's GIDS". *Revue Roumaine de Linguistique* 55(2):103–120. http://www.lingv.ro/RRL_2_2010_art01Lewis.pdf. Accessed 11 January 2011.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the World*, Eighteenth edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*. 3rd Edition. Paris: UNESCO Publishing. ISBN 978-92—3-104095-5.
- Singh, Udaya Narayana. 1990. "On Language development : the Indian perspective". In Bahner, Joachim Schildt & Dieter Viehweger, eds. *Proceedings of the 14th International Congress of Linguists*. Berlin: Akademie Verlag. 1460-71.
- , 2002. "Linguistic landscaping: An Overview". In Udaya Narayana Singh, N.H. Itagi & S.K. Singh, eds. *Linguistic Landscaping in India*. Mysore: CIIL 7-19.
- , 2003. "Life after death: Patterns of Language loss in South Asia". In Jennifer Lindsay & Tan Ying Ying, eds. *Babel or Behemoth: Language Trends in Asia*. Singapore: Asia Research Institute. 57-76.
- , 2007. "Bhaarat kaa nayaa bhaaSaa-sarvekSaN". *GaveSaNaa* 87: 124-45. (in Hindi)
- , 2009a. "Status of lesser-known languages of India". In Anju Saxena & Iars Borin, eds. *Lesser-Known Languages in South Asia : Status and Policies, Case Studies and Applications of Information Technology*. Mouton de Gruyter.
- , 2009b. "The Sense of Danger: An Overview of Endangered Languages of India". In Kamalini Sengupta, eds. *Endangered Languages of India*. 39-56. New Delhi: INTACH.
- , 2009c. "Devaluation, Danger and Death". Keynote address, International Conference on World Languages and North-East Languages: Convergence, Enrichment or Death?? NEHU, Shillong; March 18.
- , 2010. "Erosion of Cultural and Linguistic Bases in South Asia". Keynote address, 21st ECMSAS, 2010; Universität Bonn; Institut für Orient und Sienwissenschaften Abteilung für Indologie, Germany.
- UNESCO. 2003. Language Vitality and Endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Paris, 10-12 March 2003. <http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf> (accessed 11 February 2016).
- Wilson, E.O. 1992. *The Diversity of Life*. W.W. Norton & Co., New York.

Appendix

SURVEY TEMPLATE For Individual Languages

1. Introduction

Language / Mother tongue name (s)

The notion mother tongue whether it is language of the Mother or language of the Household, each assumes the primacy of a single language, but does not account for the simultaneous acquisition of two / more languages, for either childhood or adult language loss or replacement.

2. Alternative names, if any/if known

3. A brief general introduction (generalities)

4. Review of earlier literature, if any

5. Demography : Statistical and Geographical Data (Language statistics both comparative and diachronic with decadal growth rates by district / state / country)

[Tracing the trends of language maintenance and shift in India from 1901 – 2001 (fairly systemic data is available from 1901 – 2001, except some deviations in 1941, in 1947 and in 1971 and thereafter)- 1941 (language data was not tabulated), 1947 (because of partition of the country there was linguistic imbalance), 1971 (non-availability of data on minor languages fewer than 10,000 speakers since 1971). However, these Census reports are a natural starting place, as Fishman has claimed for the American Census reports, for tracing the trends of language maintenance and shift in India . Therefore the whole idea was to prepare a Census history for about 200 odd languages from 1901 till date.

6. Geopolitical concentration of languages (which became the basis for linguistic state formation): *Distribution of every 10,000 persons by language(s) (for Scheduled Languages), distribution of every 1,000 persons by language (non-scheduled languages) in the states. There is a lot of difference between a hundred thousand speakers scattered around, as opposed to a hundred thousand in a block.*

7. Segmental language figure : Sex, Rural / Urban, Religion(by cross-tabulation)

8. Amount of bilingualism / trilingualism (multilingualism) among the speakers of the language. (with age, sex, education).

Bilingualism / Trilingualism are part of the mother tongue strength.

9. Second language (s) speakers

The question on bilinguality can be viewed in two ways, 1) Bilinguals who are part of mother tongue strength, 2) Second language speakers who are added to the strength of a mother tongue; that can result in either stable or replacive

bilingualism. In the case Indian languages, second language speaker strength is marginal, vast majority of the speakers are mother tongue speakers, barring a few exceptions Indian languages do not attract non-native, second language speaker. In India English sets the highest limit of the second kind i.e., more than 90% English speakers are second language speakers – a fact which adds to the tremendous prestige of English in India. English and Hindi are two most preferred second languages in India.

10. Language and Ethnic Data

Mother tongue and bilingualism play significant roles in planning ethno-linguistic identity. Since language and identity are interrelated.

11. Language by country (Diaspora)

Although, in this survey geographical distribution of languages will be primarily based on speaker strength in India, but since many Indian languages are spoken across the boundaries and beyond – Indigenous natives of contiguous territories have a common past. Moreover, linguistic boundaries between languages in the international borders are marked by fuzzy boundaries., our survey should take note of them.

12. Language Map, Charts and Graphs

13. Language Corpus (Relation to other languages)

13.1 Family, Sub-family and group affiliation (Historical developments).

13.2 Earlier description (s), if any

13.3 Neighboring languages

14. Dialect Variation (extent of dialect diversity within a language)

14.1 Major dialects already identified (if any, even if notionally)

14.2 Different types of Dialect Variation: Regional, Social Class, Ethnic, Gender and Stylistic

Variations: In different levels, Lexical, Phonological, Morpho-Syntactic and Discourse Variation

Style variation : formal style, casual style, reading style, etc.

Social Class : variation of social stratification

14.3 Dialect Atlases : Distribution of features and Isoglosses

How much dialect variation is there in the language ? If the language in question has no standard form but shows several major dialect areas with strong feelings of dialect identification by the speakers (Roofless dialects / dialectalised languages).

15. Grammatical Description

15.1 Phonology

Phonemic Inventory

Segmental Phonemes

Vowels (*Apart from classification of vowels according to tongue and lip positions it could be classified as nasalized vowels, retroflex vowels, tense/lax vowels, voiceless vowels.*)

Semivowels

Diphthongs

Consonants

Major allophonic distributions

Phonemic distributions

Clusters

Phonotactics (*conditions on the order and type of sequences*)

Supra-segmental phonemes (or prosodic features of quantity (length), stress(loudness),

Tone (pitch).

Tone (for Tonal languages – tonemic distributios)

Stress (prominence relation between syllables – fixed stress, lexical stress, paradigmatic stress)

Intonation

Patterns of intonation (yes – no, tag, reconfirmation and echo questions, statements, information question, imperatives and blessings, curses and doubt).

Length (length of vowels and consonants is an important feature of many languages)

Co-articulation

Secondary articulation or other modifications like nasalization, retroflexion, pharyngealisation, tenseness, etc.

Syllables : Syllable structure

15.2 Script – orthography

15.3 Morphophonemics (*Alternations of phonemes in the same morpheme*)

The process involve nasal assimilation, retroflexion, high vowel shortening, palatalization, vowel raising, vowel lowering.

15.4 Morphology

Word formation

Structure of words: Derivation and Inflection, root and stem.

Word formation processes

Affixation – reduplication – suppletion – compounding – coining new words – backformation – internal change

echo-word formation

Categories : Each word belongs to a category: Nouns, Verbs, Adjectives, Adverbs, Prepositions (*Although grammatical properties of a given part-of-speech class are quite specific to a given language*).

Noun Morphology

Person

Number

Gender

Cases

Pre and Post positions

Compound Nouns

Pronouns (Personal, Demonstrative, Relative, Interrogative, Indefinite and Reflexive) : Morphological anaphora

Mary sees herself (case of pronoun), admirer – self admirer, denial – self denial, employed – self employed.

Adjectives (Variable or Invariable)

Numerals : decimal / vigesimal (Cardinals, Ordinals and Fractional)

Classifiers

Particles (A small group of clitic elements)

Nominalization Processes

Verb Morphology

Verb (Inflection, Derivation) verb system : Temporal, Aspectual and Modal as well as person and number)

Finite

Non-finite

Transitivity

Intransitivity

Types of Verb: Main Verbs, Auxiliary Verbs, Subsidiary Verbs, Additive Verbs

Types of Verb bases: Primary bases, Conjunct bases, Compound bases

Tense

Aspect

Mode

Verbal Noun

Infinitive

Gerund

Verbalization or (New) Verb formation processes

Adverbs (Time/Space/Reason/Manner/Others - esp clausal)

Clitics : affixes vs. clitics (*a small group of clitic elements is traditionally classed as particles*)

15.5 Syntax

- a) Order of Words and Phrases in simple sentences
(Including permitted change of order)

b) Subject and Objects (incl. Types of Subjects)

c) Description of Phrases (Noun Phrase followed by Verb Phrase, with Final Verb)

Noun phrase

Verb phrase

Post positional phrase

Adjectival phrase (Attributive, Predicative)

Adverbial phrase

Clauses

d) Description on types of sentences

Interrogative

Imperative

Negative

Causative

Serial Verbs

Coordination

Copular

Passive

Relative Clauses

Correlative Clauses

Complementation

Conditionals

Elliptical constructions (e.g. Headless constructs)

Semi-sentences (of various kinds)

16. Languages in contact : *Intense language contact for centuries resulting linguistic convergence : Patterns of convergence, linguistic borrowing- lexical and structural (linguist typology or linguistic areas)*

17. Basic Socio-linguistic Information (overall societal picture of a language)

Linguistic (On the basis of Abstand and Ausbau)

Heinz Kloss conceptualize as Abstand that is language by distance, a linguistic one and calls Ausbau, language by development, a sociological one.

18. Legal status: Official (by country, by state, by region) : *The official recognition of a language raises its status and gives its speakers a new dignity and sense of identity and that the census data are a favourite instrument of those who call on government to begin or to change their language status. There may be different levels of officialization, the language regarded as official may be used for various purposes of government. But there may be several gaps between the ideal and the reality.*

Scheduled / Non-scheduled : The languages of India have been divided along the line of the Indian Constitution as scheduled and non- scheduled languages. The languages belonging to non schedule category appeared or disappear as the case may be, depending upon the strength of their returns i.e., 10,000 and above.

19. Whether written, if written, in what manner (Script and spelling)
If not written – Oral Literature
A primary index of language unfolding
20. Disseminated in what form (pamphlets, periodicals, news papers, Books, Films, CDs / Cassettes, Akashvani, Doordarshan)
Thus provides baseline data in the realm of written as well as spoken traditions.
21. Used for what purposes (Education, Religion, Government, Industry, Legislature, Courts of Justice)
Key domain to ascertain the ranges and intensity of language usage or functions. This will offer the possibility of quantifying functional usage both in terms of extensiveness and frequency of use, and hence allows for a measurement or weighting of this usage.
22. Lexicon
Basic words of the language
It is the repository of basic items on which grammar rules operate(words).
23. Texts
Prescribed
Free
Discourse
24. Reference Frame Work : Reference Books, Dictionaries, Grammars, Bibliographies, Basic Socio-linguistic Studies ; Specialists and Consultants, Agencies / Institutes concerned with promoting the language.
25. Bibliography

IPA Transcription and transliteration scheme for Indian Languages

Devnagari	IPA Symbol for Hindi	Bangla	IPA Symbol for Bangla	Roman Transliteration
अ	ə	অ	ɔ	a
आ	a:	আ	a	ā
इ	i	ই	i	i
ई	i:	ঐ	i	ī
उ	u	উ	u	u
ऊ	u:	ঊ	u	ū

ऋ	ri	ঋ	ri	ṛ
ए	e:	এ	e	e
ऐ	ɛ:	ঐ	oi	ai
ओ	o:	ও	o	o
औ	ɔ:	ঔ	ou	au
क	k	ক্	k	k
ख	k ^h	খ্	k ^h	kh
ग	g	গ	g	g
घ	g ^h	ঘ	g ^h	gh
ङ	ŋ	ঙ	ŋ	ṅ
च	tʃ	চ	tʃ	c
छ	tʃ ^h	ছ	tʃ ^h	ch
ज	dʒ	জ	dʒ	j
झ	dʒ ^h	ঝ	dʒ ^h	jh
ञ	n	ঞ	n	ñ
ट	t	ট	t	ṭ
ठ	t ^h	ঠ	t ^h	ṭh
ड	d	ড	d	ḍ
ढ	d ^h	ঢ	d ^h	ḍh
ड़	ɽ	ড়	ɽ	ṛ
ढ़	ɽ ^h	ঢ়	ɽ ^h	ṛh
न	ɳ	ণ	n	ṇ
त	t̪	ত	t	t
थ	t̪ ^h	থ	t ^h	th
द	d̪	দ	d	d
ध	d̪ ^h	ধ	d ^h	dh
न	n	ন	n	n

প	p	প	p	p
ফ	p ^h	ফ	p ^h	ph
ব	b	ব	b	b
ভ	b ^h	ভ	b ^h	bh
ম	m	ম	m	m
য	j	য	d̪	y
র	r	র	r	r
ল	l	ল	l	l
ৱ	ʋ	ৱ	b	v
শ	ʃ	শ	ʃ s	ś
ষ	ʂ	ষ	ʃ	ṣ
স	s	স	s	s
হ	ɦ	হ	h/ɦ	h
ঙ	ŋ	ং	ŋ	m̐
:	h	ঃ	h/ɦ	ḥ
ঁ	~	ঁ	~	◌̣
ক	q	ক	q	q
খ	x	খ	x	kh
গ	ɣ	গ	ɣ	g
জ	z	জ	z	z
-	ʒ	ঝ	ʒ	zh
ফ	f	ফ	f	f
-	v	ভ	v	v

Acknowledgements

Abbi, Anvita; Former Professor of Linguistics, Centre for Linguistics, School of Languages, Literature & Culture, Jawaharlal Nehru University, New Delhi – 110 067; E-Mail anvita@mail.jnu.ac.in

Banerjee, Purnendu Kishore; DRG, Office of the RG&CCI; E-mail: purnendukb@yahoo.com

Bhattacharya, S.S.; Formerly Senior Research Officer, Office of the Registrar General of India, Kolkata; Email: shyamsundarbhattacharya@gmail.com

Bokova, Irina; Director-General of UNESCO; from Sofia, Bulgaria; Email: i.bokova@unesco.org

Devy, Ganesh; Director, Bhasha, 62, Shrinathdam Society, Near Dinesh Mill, Baroda – 390 007(Gujarat); E-Mail: ganesh_devy@yahoo.com

Gunasekaran, D.; Former Registrar, Visva-Bharati, Santiniketan; Now Registrar, IIT, Bhubaneswar, Odisha; Email: dgshekaran@gmail.com

Minasyan, Anahit; Overseeing editor at the UNESCO Atlas of the World's Languages in Danger of Disappearing; Email: A.Minasyan@unesco.org

Mishra, Awadesh Kumar; Director, CIIL, Mysore; Email: akmishraciil@yahoo.com

Mishra, Rameshwar Prasad; Former Adhyaksha, Bhasha Bhavana and Professor of Hindi, Visva-Bharati, Santiniketan; Email: rpmishra@gmail.com

Moseley, Christopher; BBC, General editor, UNESCO Atlas of the World's Languages in Danger; Now University College, London; Email: chrismoseley50@yahoo.com

Mukhopadhyay, Tapati; Director, Rabindra Bhavana, Visva-Bharati, Santiniketan; Email: tapatimukherjee1952@gmail.com

Munda, Late Ram Dayal; Former M.P., & Former Vice-Chancellor, Ranchi University)

Nicolas, Alexandre; the Cartographer of the UNESCO Atlas.

Ojha, S. N.; Former In-Charge, Computer Centre, Visva-Bharati, Santiniketan; Email: snojha2@gmail.com

Opgenort, Jean Robert; trained in Leiden, a software development engineer, linguist and e-learning consultant associated with UNESCO Atlas project; formerly an Assistant professor of the University of Bern and now in Apple; ISM eCompany; E-mail: jeanrobert@opgenort.nl

Patnaik, B.N.; Formerly Prof. of English, Dept of HSS, IIT/Kanpur; Residence: 8061, Sobha Daffodil Apartments, Somasundara Palya, HSR Layout, Sector II, Bangalore – 560 102; E-mail: bn.patnaik@gmail.com, patnaik@iitk.ac.in

Prakash, Ved, currently Chairman, UGC; formerly also Vice-Chairman and Secretary, UGC and Vice-Chancellor, National University of Educational Planning and Administration (NUEPA), New Delhi – known as a specialist of Education and Measurement at the NCERT; E-mail: cm.ugc@nic.in

Prasanna Sree, Andhra University, Waltair, Visakhapatnam – 530 003 (A.P); E.Mail: sathupati@gmail.cpm

Reddy, B. Ramakrishna; Formerly Professor of Linguistics, P.S. Telugu University, Hyderabad; Email: brkrin@yahoo.com

Sen, Sabujkoli; Director, Sriniketan and Professor, Department of Philosophy & Religion, Visva-Bharati, Santiniketan; Email: sabujkolisen@gmail.com

Singh, K.P. Joint Secretary, University Grants Commission, New Delhi-110002; E.Mail: kpsingh@ugc.ac.in

Thorat, Sukhdeo, formerly Chairman, UGC, and also Chair-person, ICSSR, Government of India – an economist who worked on the Dalit cause; taught at the Jawarlal Nehru University; E-mail: skthorat@hotmail.com

Vatsyayan, Kapila; Former M.P.,Chairperson,IIC – Asia Project India International Centre, 40, Max Muller Marg, New Delhi – 110 003; E.Mail: asiaprojectiic@yahoo.com