

ABSTRACTS

**ENDANGERED AND LESSER KNOWN LANGUAGES:
CHALLENGES AND RESPONSES- ELKL I**

DEPT OF LINGUISTICS, UNIVERSITY OF LUCKNOW

IN

COLLABORATION WITH

CIL, MYSURE

2012



How serious is language endangerment and how vital is language documentation?

Anvita Abbi
Jawaharlal Nehru University, New Delhi
anvitaabbi@gmail.com

Language endangerment is the biggest threat to linguistic diversity and human rights. Language loss is often the result of oppression of speakers of minority languages, violating their rights on their land, their forest, forcing them to dislocate themselves from their environment, and losing the social and spiritual identity. To arrest or decrease the rate of endangerment two courses of action are required, viz. shift in the language policy and documentation of the dying language.

Languages are witnesses to the diverse and varying ways in which human cognitive faculties deal with the world, so the study of endangered languages and their documentation in forms of text, grammar, dictionary, ethnolinguistic and sociolinguistic study, audio and video recordings of language in use open windows to a world long past. As a result of the richness of the information thus gleaned, we are allowed a unique insight into the world- views of the speakers of the language.

Actions to prevent language loss and to safeguard linguistic diversity will only be successful when meaningful contemporary roles for minority languages can be established by introducing them at school level through Four Language Formula in phases. Reversing the stream of negative attitude towards one's heritage language and associating its learning and teaching with economic gains will ensure the sustenance of dying languages.

The keynote lecture will share the exceptional experiences of compiling a multilingual and multiscriptal interactive dictionary of Great Andamanese. The ethno-semantic and ornithological account of the local birds and their names in the Great Andamanese language that feature in the dictionary and in the book 'Birds of Great Andamanese' in great part reveals the various ecological and archeological signatures of the original communities that maintained close ties with their environments. What is more, the grammar of Great Andamanese becomes a significant piece in the jigsaw of language evolution. Documentation of endangered languages is imperative in scientific understanding of human language and human evolution.

Tibeto-Burman Languages A Syntactic Typology

Kārumūri V. Subbārāo
kvs2811@gmail.com

The aim of the present paper is to provide a syntactic typological analysis of Tibeto-Burman languages. We shall also see how syntactic strategies are used in achieving politeness phenomena in some of these languages. Subbarao (2012) presents a detailed discussion of the syntactic typology of the four major language families of the South Asian subcontinent. Most of the Tibeto-Burman languages are under-analyzed and some of them are even, endangered (Ralte, for example) and some of them do not have even a phonological or any other description available. Tibeto-Burman languages present

a great opportunity to study the diversity in language. Studying the various construction types in a language or a group of languages provides us a chance to see the range of variation found in human languages which would enable us to get a better understanding of the nature of human language that will consequently help us to arrive at a better characterization of Universal Grammar (UG)

Out of the four major language families in the South Asian subcontinent it is the Tibeto-Burman and the Munda languages that exhibit an enormous diversity, as Subbarao (2012) demonstrates. For example, there is no subject-verb agreement in a subset of languages as in Manipuri (aka, Meithei), Kokborok, Sema, Liangmai and Tangkhul and in contrast, there are languages with robust subject, direct object, indirect object and even an oblique object agreement coupled with the phenomenon of rich incorporation of quantifiers, negative polarity items, adverbs and even postpositions. Adverb Incorporation in some of the Naga languages (Tenyidie and Liangmai, for example) is fascinating as manner and some other adverbs carry the conjunctive participial marker (cpm) when they occur independently. When incorporated in the verb, the cpm is obligatorily deleted.

This is the only group of languages that has Internally-Headed Relative Clauses (IHRCs) in addition to having Externally-Headed Relative Clauses (EHRCs). In a set of languages in IHRCs the overt post position that indicates the thematic relation of the argument with the verb is overtly present on the Head, and in the other set of languages, it is not. We shall show that such occurrence or non-occurrence of the postposition has implication in the relativization of a comitative PP as Head. Though the preferred strategy in relativization is to have the nominalizer occurring on the embedded verb, a few languages have a full-fledged relative clause too in which a dubitative marker occurs with the embedded verb and we shall show that this dubitative marker performs all the functions that such a marker performs in Dravidian.

With regard to lexical anaphors, Tibeto-Burman languages have a nominal and a verbal reflexive, just as Dravidian languages do. A verbal reflexive or reciprocal exhibits all the properties that are normally attributed to verbal anaphors universally. In addition Tibeto-Burman languages have some unique features found nowhere else to the best of our knowledge. These include the occurrence of a specific tone as a device to express reciprocity in Tenyidie or Swapping of the constituents of complex nominal anaphor in Liangmai.

We shall also show how Negative Polarity Items, quantifiers, negation, adverbs of manner etc. are incorporated in the verb. We shall also show how the incorporation of a postposition in the verb increases its valence and thus, an intransitive verb is transitivized which has implications to ergative case marking on the subject and verb agreement. We shall also show how verbs and adjectives behave in an identical fashion. The non-permissibility of the occurrence of the ergative marker in Non-nominative Subject construction also deserves a mention.

In this paper we shall thus discuss some major aspects of the syntax of Tibeto-Burman languages that include anaphora, agreement, forward and backward control, non-nominative subjects and Externally Headed Relative Clauses (EHRCs) and Internally Headed Relative Clauses (IHRCs).

References

Subbarao, K. V.2012. *South Asian Languages: A Syntactic Typology*. Cambridge: Cambridge University Press; New Delhi: Cambridge University Press

Budhali: A Descriptive Note

M. K. Koul¹, Nazir Ahmad Dhar, ²Parvaiz A. Ganai*, Tanveer Ahmad^{*3}

neelofar.wani@gmail.com

Budhali is a link language in Budhal (Rajouri, J&K). There are three ethno-linguistic communities that live in Budhal viz. Kashmiri, Pahari and Gojri. There are native speakers of all the three languages. Budhali is the language, as per the people who speak it, of the descendants of Raja Budh Pal primarily, but at present it is used by all the communities as a link language. It has no literature and also lacks a script. The Census of India groups the Budhali speakers under Pahari language. However, Pahari and Budhali, though resembling with one another, are linguistically different speech varieties. The contemporary status of the language is that it is mother tongue to a significant number of people living in Budhal, Dandoot, Keval, Phalni, Triganyi etc. of the Rajouri district.

In the morphological study of this language, various resemblances were seen with other languages and dialects of the region e.g. [t^hi] used as auxiliary is also used in Kishtwari- a dialect of Kashmiri. It must be noted that some of the Budh Pal descendants had had settlements in Bhaderwah as well, and the natives claim that Budhali has links with Bhaderwahi. Since Bhaderwahi is, geographically, close to Kishtwari, the existence of [t^hi, t^his, etc.] in Budhali may be traced back to Kishtwari. Our survey of the language also revealed that in a local school the following breakup of the students on the basis of the language spoken by them was found.

Kashmiri	>	86	
Pahari	>	130	(Budhali speakers included)
Gojri	>	44	
Total	=	260	

40% population are Kashmiri speakers in Budhal whose mother tongue is Kashmiri. However, they also speak languages other than Kashmiri. The resemblance of Budhali with Bhaderwahi was stressed by our language consultant during our fieldwork in May, 2012 for the project “Survey of Kashmiri Dialects”. The present paper is a descriptive note on this so far unreported language.

Language documentation and preservation: CPDP an example from Nepal Vishnu S Rai

¹ Principal, Urdu Teaching and Research Centre, Lucknow and Coordinator, *Survey of Kashmiri Dialects* project- funded by CIIL, Mysore.

² Head, Department of Linguistics, University of Kashmir, Srinagar and Principal Investigator, *Survey of Kashmiri Dialects* project- funded by CIIL Mysore.

³ *PhD Research Scholars, Department of Linguistics, University of Kashmir, Srinagar and Research Assistants, *Survey of Kashmiri Dialects* project.

**Associate Profesor
Faculty of Education
Tribhuvan University
Kathmandu, Nepal
raivishnu1@gmail.com**

Language documentation is one of the most effective way to preserve an endangered language. Chintang Puma Documentation Project (CPDP) was conducted jointly by a team of linguists, anthropologists and psycholinguists from the Central Department of Linguistics, Tribhuvan University, Nepal and Institute for Linguistics, Leipzig University, Germany. Thus it was an inter-disciplinary project.

The aim of the project was to document the two endangered languages Puma and Chintang from Nepal so that they can be preserved. It also aimed to document the cultural aspects of the two language communities, and to conduct a longitudinal study of Chintang children's L1 acquisition.

The data was collected by audio-video recordings of different genres such as, myths, songs, folktales, biographies and autobiographies, stories, rituals and narratives about places, people and recipes. The recordings were transcribed, translated and glossed linguistically on the toolbox –a programme which enables to gloss the text as well as compile the dictionary of the language. Finally, they were put into ELAN, which enables a viewer to read the glossed text (song, story, ritual, etc.), hear them as well as see the person talking the text and/or performing the rituals.

The project collected a large corpus for archiving purpose, which can be used for further scientific research. The archive not only includes a large corpus for further research but also the linguistic analysis of the most of the texts, a sketch grammar and a dictionary of the two languages. It also contains anthropological descriptions of rituals and myths and an analysis of the Chintang child language from the language acquisition perspective. The project also produced dictionaries and grammars for both the language community as well.

The paper presents the processes and findings of Chintang Puma Documentation Project which aims to document Chintang and Puma the two endangered languages from Nepal for preservation and promotion purpose and how the community is going to benefit from the project. The paper presenter will also make a visual presentation of how the analyzed texts look like on screen.

Raji Orthography Development

**Dr. Kavita Rastogi
Department of Linguistics,
University of Lucknow
Lucknow**

krastogi@satyam.net.in

Raji is a little known tribal community whose descendants are the prehistoric Kiratas. They live in dense forests far away from the surrounding Kumauni villages of Pithoragarh district. In 2001 census their population was reported to be 680 in all the nine villages. Sir George Grierson, in his book 'Linguistic Survey of India' had named this language as 'janggali which has only spoken form.

Following the framework established by Wurm and the stages of threatenedness discussed in Fishman's GIDS Raji can be assessed as '*potentially endangered and at stage*

6,(language) which means the language is at risk.' While chalking out a revitalization programme for this oral language I realized the need of orthography development for this language. It is an established fact that Orthography gives stability to a language and not only conserves it but also helps in the standardization process.

So after preparing a small grammar book, with the help of collected phonological and grammatical material of Raji the next important task before me was to develop an orthography system. The present paper focuses on the early stages of orthography development for this previously undocumented indigenous language.

Language Attitude of the Chennai settled Malayalees

Dr. P.M. Girish

Assistant Professor

University of Madras, Marina Campus, Chennai

drpmgirish@gmail.com

The Present paper, entitled 'Language Attitude of the Chennai Settled Malayalees' attempts to analyze the language attitude of the Malayalees, settled permanently in the city of Chennai, the capital of Tamil Nadu.

An interesting fact is that the migrant Malayalees are distributed in almost all districts in Tamil Nadu. In Tamil Nadu migrant Malayalees occupy the fourth position, when compared with other linguistic minorities such as Telugues, Kannadigas and Urdu speaking people. Chennai is one of the major metropolitan cities in Indian having lot of migrated people from various parts of the country. This paper focuses on the migrant Malayees who settled in Chennai since 1950's.

This paper may bring out the socio-political attitude of the migrant Malayalees towards the dominant language of the region.

Linguistics Situation of Karnataka

BasavarajaKodagunti

Central University of Karnataka, Gulbarga

bkodagunti@gmail.com

This paper intends to provide a comprehensive picture of linguistics situation of Karnataka. The paper provides details of number of languages recorded from Karnataka, and details of scheduled and non scheduled languages, which are listed in Indian senses, 2001, and also an attempt is made to list out the languages which are not listed in the Indian senses and are used in Karnataka. Major intention of the paper is to provide the statistical facts about languages of Karnataka, which would help us to work on language policy for Karnataka. The paper majorly depended on Indian senses 2001 for the data of scheduled languages and non scheduled languages and listed in the senses. And for the languages which are not listed in the Indian senses, varied sources are made use of.

There are 22 scheduled languages, all of which are used in Karnataka and there are 100 non scheduled languages listed in Indian senses, out of these, quite interestingly, 73 languages are reported from Karnataka. Out of these 73 languages, only 6 languages have about 1000 reported speakers. 8 scheduled languages have more than 1,00,000

speakers reported, and excluding those languages there 6 languages for which more than 10,000 speakers are reported.

The number of speakers of scheduled languages is 50,865,780 and of the non scheduled languages listed in senses are 1,984,782. And quite interestingly the number of speakers of languages which are not listed in senses and given in the same as other languages is 271,689, which is really considerable number. There are about 20 languages having less numbers of speakers which are not studied well. And they are not considered by the state government at any level for any kind of activity. The endangered languages listed in the UNESCO report, e.g. Tulu, and other lesser known languages e.g. Koraca, and languages for which there are considerable number of speakers e.g. Lambani/Banjara, are to be considered linguistics map of Karnataka. And there are languages for which number of speakers is about 1,000 and fewer languages have lesser than 1,000 number of speakers.

Nyishi world through words: A presentation of Nyishi-English bilingual dictionary.

Sudhanshu Shekhar' Manisha Singh
Jawaharlal University New Delhi, shekhar921@gmail.com
Banaras Hindu University, manisha.singh39@gmail.com

This proposed work concerns itself with demonstration and presentation of English-Nyishi bilingual dictionary using the software called toolbox. Nyishi, earlier known also as Nissi, Nishing, belongs to Tibeto-Burman stock of Sino-Tibetan language family. It is the name of both the tribe and the language. Speakers of Nyishi inhabit the districts of Kurung Kumey, Papum Pare, Lower Subansiri, Upper Suibansiri and West Kameng of Arunachal Pradesh. Very few works can be found on the linguistic study or sketch of this language. This work is mainly based on data collected during the year 2009.

With the globe metamorphosing into a global village, the shifting from a local, traditional language to rising global languages is a phenomenon common to many parts of the world and Nyishi is no exception to it. Intergenerational gaps in local language transmission are increasing. Extinction is predicted for thousands of the world's languages. Nyishi speakers are shifting to major languages like Hindi and English. Nyishi children and young adults who have been growing up even in small towns speak fluent English and Hindi, but they cannot speak Nyishi fluently. Hinton (2001) observes that as a language loses speakers, knowledge accumulated by its associated traditional culture is also diminished. Stories, histories, knowledge and traditional practices are in danger of dying with the language death. Therefore, children who do not know Nyishi language will not have awareness of the Nyishi culture, and thus the culture will disappear in time to come. Language not only helps us to perceive the environment, but it informs the unique culture of the group, rooted in various aspects of its past. Language is a part of

culture and identity. And at the same time a whole culture is enshrined in or built into language.

This would be an encyclopedic dictionary as it gives lexical and encyclopedic information. We will begin with the inclusion of lexical (i.e. linguistics) and non-lexical (i.e. encyclopedic) information in the dictionary. The lexical or linguistics information pertains to linguistic characteristics of the lexical unit viz. gloss, pronunciation, definition, sounds, grammatical category, etc. the work is in progress and we will try to include as many entries possible.

Motion verbs in Bodo

Aleendra Brahma & Bidisha Som

aleendra.iitg@gmail.com & bidishasom@gmail.com

Indian Institute of Technology Guwahati

Language based hypothesis of cultural relativism argues that linguistic patterns might influence people's conceptualizations during shaping the schemata used for representing the environment, resulting into a strong link between the non-linguistic cognitive processes and the form and content of the language spoken. In the backdrop, motion events form a core area of investigation within the cognitive semantics framework of linguistic studies, helping to understand the underlying cognitive processes of perceiving 'motion' through an analysis of the relationship of meaning and surface expression in languages. The way a language lexicalizes the motion events reflect the way the event is conceptualized.

The present paper investigates the phenomenon of "motion events" (Talmy 2000b; 25) in Bodo language, spoken in Assam. Specifically, we look into the issue whether the verb encoding information about the cause/manner (e.g. float) is able to combine with a path (into the cave) to convey directed manner/cause of motion within a single clause as in *The bottle floated into the cave* (1991: 488) as in English. Other languages (e.g. Spanish, French) cannot combine a manner of motion verb with a path/locational expression to conflate two events in a minimal clause, deeming ungrammatical the equivalent of *The bottle floated into the cave* (1985). Rather, only path of motion verbs can occur as the main verb of directed motion constructions; the causal/manner event must be expressed separately as an adjunct: *La botella entro' flotando a la cueva* 'the bottle entered floating to the cave' (1991: 488). In this study, we have examined Bodo constructions of the type discussed by Talmy (1985) in the domain of spontaneous motion. Following the typology proposed by Talmy, the present research presents an analysis of the Bodo motion verbs with a focus on the preferred construction type of motion verbs in this language.

We propose that Bodo is an S-framed language in that the verb root encodes 'manner' of motion. The path/locational information is reflected through a 'satellite'. These constructions also show a very strong bias towards encoding location rather than path information. This phenomena in the present study will be analysed from the point of view of whether atelic path phrases are (relatively more) felicitous with manner of motion verbs than those which are telic (Aske 1989) in case of the Bodo language.

References:

- Aske, J. (1989). Path predicates in English and Spanish: a closer look. Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society (pp. 1–14). Berkeley, CA: Berkeley Linguistics Society.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), Language typology and lexical description: Vol. 3. Grammatical categories and the lexicon (pp. 36-149). Cambridge: Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. Proceedings of the Berkeley Linguistics Society, 17, 480-519.
- Talmy, L. (2000). Toward a cognitive semantics: Vol. II: Typology and process in concept structuring. Cambridge, MA: MIT Press.

Case-Markers in HO

Richa Ojha
Research Scholar
University of Lucknow

This paper aims to describe case-marker of Ho language. Ho, one of the major languages spoken in the state of Jharkhand in eastern India belongs to the Austroasiatic (Austic) language family.

The aim of this paper is to establish the case relations in Ho language. The present study emphasized on the semantic aspect of the case, as a result postposition has also been taken into account. In researcher's opinion the way of expressing case in Ho requires postpositions marked with case ending.

In order to exemplify it, we can say that in the sentence '*table cetan-re plait mena?*' (The plate is *on* the table) *cetan-re* expresses the postposition with case ending, wherein *cetan* means above which acts as a postposition here and *-re* case marker representing location. Thus in this sentence '*-re*' acts as a locative case-marker.

A standard elicitation technique, 'Topological Relation Picture Series' was used in the study which is a booklet of seventy-one pictures. The study covers a large number of relations that would be coded in English using such prepositions as on, in, under as well as complex prepositions like inside, on top of.

Ethnolinguistic Study of Gulgulia

Sneha Mishra

Research Scholar
sneha.daffodils@gmail.com
Dept of Humanities and Social Sciences
IIT-Dhanbad

This paper briefly discusses the ethnolinguistic vitality of the Gulgulia community of the Patherdih area in Dhanbad, Jharkhand keeping in view the socio-economic aspects of Gulgulia life and its influence on the Gulgulia language and culture across time in the multilingual setting.

The Gulgulias are a nomadic community, often found on the out-skirts of villages, near railway stations or forests. They are widely distributed in the states of Jharkhand, Bihar

and Bengal who live by arranging shows of animals, begging and even petty thieving. The gradual extinction of the Gulgulia community from our society and most importantly the threat that the Gulgulia language is undergoing triggers our focus towards a closer look at the ethnolinguistic vitality of the Gulgulia speech community.

This paper chiefly examines the ethnolinguistic vitality of the Gulgulia language on the basis of indicators as proposed by M.Lynn Landweer and the factors as supported by the UNESCO to evaluate a language's vitality and state of endangerment. Taken together, these indicators and factors are useful for characterizing the language's overall sociolinguistic situation, thereby identifying its imperative needs. Moreover, the endangered situation of the Gulgulia language and the plausible direction towards which the speech community can proceed has been discussed in the context of a social-psychological framework of language maintenance and shift.

Magahi Complex Predicate and Agreement

Nilu, Rajesh Kumar

nilu@iitp.ac.in, kumar@iitp.ac.in

Department of HSS, IIT-Patna

Magahi varies from Indian languages in agreement pattern among many others. This has implications for many syntactic aspects too. One such aspect is complex predicate (CP) construction. Butt (1994) and Verma ed. (1993), define complex predicate as a multi-headed construction which behaves as a mono-clausal unit in various syntactic mechanisms such as passivization and agreement. The phenomena of complex predicates (V-V, N-V or A-V) is common in almost all the South Asian linguistic area, occurring in all of the major language groups of Dravidian, Tibeto-Burmese, Indo Aryan. In case of compound verb, main verb (V1) is either in participial form (as in Bangla, Oriya and Assamese) or a root (as in Hindi) and light verb (V2) is semantically bleached and carries inflection (Paul 2004). The very first verbal member of the compound verb construction has variously been called 'main verb', 'polar verb', and 'principal verb'. The second delexicalized verb (form) is known as 'operator', light verbs, explicators, and vectors, intensive auxiliary.

'Agreement' is the grammatical phenomenon in which the form of one item forces a second item in the sentence to appear in a particular form. Agreement is a language specific feature found in almost all languages of the world. Magahi should be categorized under as one of the endangered language on the basis of regular decline in the number of speakers of the language. On the basis of the census of 1881, Grierson mentioned the total number of Magahi speakers 6, 504, 817. The census of 1951 shows the reduced number of Magahi speakers, i.e. only 3, 728. According to the census of 1961 the number of Magahi speakers is 3, 792, 447. According the census of 2001 the number of Magahi speakers is 1,05, 66, 842 (Abbi 2000: 15). Magahi (Indo-Aryan language) is a language with its own uniqueness in context of agreement system. For Hindi, structural case and verbal agreement are intimately connected as structural case can be assigned via Agr positions (Mahajan 1989, 1992). But this is not the case with Magahi. The agreement in Magahi is shown through person and honorificity.

The verb in Magahi may agree with both the subject (1) and object (2) in person and honorificity.

(1) rahiim khaanaa khaa le-l-aii/le-l-kaii
Rahim food eat take-Pt.Perf-3.NH
'Rahim has taken the meal.'

(2) ham toraa jaae de-l-iau
I you.NH go give-Pt.2.NH
'I let you go.'

In Magahi, the argument carrying case markers has no effect on the agreement. The verb agrees in the same way with the subject with case marker in person and honorificity. One of the reasons behind it is the absence of inflections based on gender in Magahi. In this paper, we try to present a full description of agreement system of Magahi especially in the context of complex predicates.

References

- Abbi, A. 2000. Forgotten Indian Heritage Languages of Minority Communities of India. In O. Koul and L. Devaki (eds) *Linguistic Heritage of India and Asia*. Mysore: CIL Printing Press
- Butt, Miriam. 1994. *The Structure of Complex Predicates in Urdu*. Ph. D dissertation. Stanford University.
- Mahajan, A. 1989. Agreement and agreement phrases. M.I.T. Working Papers in Linguistics. Vol. 10, 217-252.
- Mahajan, A. 1992. The Specificity Condition and the CED. *Linguistic Inquiry*. Vol. 23 (3). 510–516.
- Paul, Soma. 2004. *An HPSG Account of Bangla Compound Verbs with LKB Implementation*. Ph. D dissertation. University of Hyderabad, Hyderabad.
- Paul, S. 2004. *An HPSG Account of Bangla Compound Verbs with LKB Implementation*. Ph. D Dissertation, University of Hyderabad, Hyderabad.
- Verma, M. K. 1993. Complex predicates and light verbs in Hindi. In M. K. Verma, ed., *Complex Predicates in South Asian Languages*. New Delhi: Manohar Publishers and Distributors.

Honorifics in Maithili

nupurc1@gmail.com

The English and Foreign Languages University

This paper intends to discuss the honorifics used in Maithili language. Maithili is an Indo-Aryan language spoken in states of Bihar and Jharkhand. An honorific is a word or expression with connotations conveying esteem or respect when used in addressing or referring to a person. Maithili uses a very complex system of honorifics where the form of addressal varies with varying relationships. For instance, the second person singular pronoun can take the following forms:

(i) tu kaaj kar jaa kau.

'You go and work.' (tu is non-honorific)

(ii) ahaN bhojan kau liya

'You (HON) eat.' (ahaN is used for addressing those whose social status is higher)

(iii) i jethin?

'Will you go?' (*i* is used for in-laws relationships)

The present work will study two forms of honorifics: honorific pronouns and terms of address.

Sanskrit in Varanasi: A Struggle for Survival

Ms. Shrishti Singh
singhsriss@gmail.com
Dept of Linguistics

Banaras Hindu University, Varanasi

Sanskrit is the language of the Vedas, Upanishads, Puranas, ethics and several other ancient scriptures. It was a widely spoken common man language in India for a long time. There is a great affinity shared between *The Capital of all Knowledge*, Varanasi and the language regarded as the *earliest and purest* of all, Sanskrit.

Varanasi, an amazing spot of social and cultural mix, fantastic combination of physical, metaphysical and supernatural elements, and face of India's tradition, art, philosophy, mysticism; therefore is very much close to the repository of India's cultural heritage, Sanskrit, in many ways. Being a pilgrimage it is rich in *Dharmik Anushthans* (religious customs). People here perform different ceremonies at every twist of life from birth till death. Inhabitants here start their day by offering prayers to the rising Sun and no new engagement, ceremony or foundation is laid without offering prayer to lord Ganesha. Each day of week is named after God and Goddesses worshiped there. It is engraved in the hearts of people in such a way that an uneducated citizen is able to recite hundreds of '*Shlokas*' and '*Mantras*' effortlessly. In spite of having a tremendous history, it is pathetic to state that this language has lost its wide speech community some thousand years back. The language which scientists are now taking as model in current developments in fields of *computational and machine translation* for its explicit grammar is suffering great loss of its actual speakers. In a survey in 1991 it was found that the estimated number of Sanskrit speakers all over India is reduced to 0.5 lakh only. Language once placed among top 18 Indian Languages is now limited to be a sticker to the punch lines of administrative councils and some other organizations only. For past two decades the initiatives taken by youths in preserving the culture by actively participating in Sanskrit related workshops, as researchers in Universities and finding their career in this field cannot be overlooked. Some other administrative measures are there that if taken will surely help in promoting and taking it to the same height where it ought to be. In this paper, I would like to highlight the linguistic changes that have made language of the elite, Sanskrit, an endangered language.

References:

Gary A. and Emery A. Bose 2006. Scholastic Sanskrit: A Handbook for students, American Institute of Buddhist Studies and Columbia University Press

Grierson, George Abraham 1906. Journal of the Royal Asiatic Society of Great Britain & Ireland, Cambridge University Press

Hardy, F.E. 1994. The religious culture of India: Power, Love and Wisdom. Cambridge University Press

The Impersonal Pronoun *jaNe* 'one': A Classifier in Odia?

KalyanamaliniSahoo

This paper deals with the morpho-syntactic properties of the impersonal human pronoun *jaNe* 'one' in Odia language. *jaNe* 'one' carrying 3rdsg. features refers only to humans and allows a generic reading of the construction. *jaNe* can be morphologically realized, or can be a null category as well:

- (1) a. (**jaNe**) adhikaaraapekhyaa kartabya-kubesimahatwadebaadarakaara
(one) right - than duty-ACC more importance give-INF necessary
'One should give more importance to duty than rights.'
'I/we/you/he/she/they should give more importance to duty than rights.'

Interestingly, this pronoun gives a generic reading as a subject (e.g.(1)), but is compatible for both generic and arbitrary/specific interpretation in the object position (e.g. (2)):

- (2) **jaNaku** Daaka, dekhibaa
one-acc call, see-fut-1st pl.
'Call one (of them), we will see.'

In the generic reading, although it can include the speaker as well as the hearer, morphologically it does not show any special connection to 1st or 2nd person. Roughly, it involves 'generic simulation', that is, keeping anyone in a position meeting the relevant conditions. Although, *jaNe* 'one' (as in (2)) can carry case features, it cannot take passive morphology. Thus, it carries a special phi-specification (Chomsky, 1995) as an argument.

According to Holmberg (2005), consistent null-subject languages have a D-feature in I(nflection), which means that when a null Φ P (Phi-phrase) enters into an Agree relation (Chomsky, 2001) with I, it can be interpreted as definite, referring to an individual or a group. That is, a null subject cannot have a generic interpretation. Thus, to express generic meaning, null subject languages have to use a different type of pronoun. However, Odia does not confirm to this generalization fully, as being a consistent null-subject language also, the null subjects can provide a generic reading.

I propose that although the 2ndsg. polite and 3rd pl. personal pronouns can be used as impersonals, the real impersonal pronoun in Odia is *jaNe* 'one', which is the combination of the classifier *jaNa*_[+human] and the bound morph *-e* (*-e* is a historical derivative of the numeral *eka* 'one', but has been reanalysed as a determiner: a singular number marker or indefinite article like English 'a/'an', and thus, is generated in the D slot in a DP (Sahoo, 1996)). The generic reading of this pronoun can be attributed to the bound indefinite affix *-e*, which is realized in a (NOM) subject position, but is substituted by the case marker *-ku*_{Acc/Dat} in the object position, thus allowing a specific reading to the object NP as Acc case-marker carries [+def] feature in this language.

Compiling Kashmiri Corpus: A Case Study in Empirical Method of Language Documentation

Shahid Mushtaq Bhat

shahid.bhat3@gmail.com

Linguistic Data Consortium for Indian Languages

Central Institute of Indian Languages, Mysore

Empirical method corpus linguistics in linguistics (corpus linguistics) has gained momentum since the recent past due to various achievements in natural language processing & speech processing. It has led to development of huge corpora. However, corpus development should not be seen from the perspective of language technology only, rather, it should be seen as an empirical or quantitative method of language documentation & field linguistics but it is possible only when corpora are created as reference works rather than domain specific works & application oriented works. Further, corpora are store houses of linguistic & non-linguistic information which can be analysed & annotated to carry out different descriptive & cognitive studies besides bringing technology to the concerned language.

This paper tries to blur the boundaries between corpus linguistics, field linguistics & language documentation and reports a research programme of developing 1 million words Kashmiri text and approx 45 hrs speech corpora, at LDC-IL. It also reports a separate research attempt of developing small scale text corpus (50,000 words) for Dependency Treebank of Kashmiri. Finally, the paper argues that there is urgency in using quantitative methods & techniques in language documentation so that the efforts will fetch more values & more could be done to cater the technological needs of lesser known languages besides the traditional analysis.

Reflexives in Santali: A Typological View

J. Mayuri Dilip, Martin Everaert,

Kārumūri V. Subbārāo

& G. Uma MaheshwarRao

Nominal expressions are classified into three classes.

- i. Nouns/ Referential expressions (R-expressions)
- ii. Pronouns
- iii. Anaphors
 - a. Reflexives
 - b. Reciprocals

These forms obey three binding principles:

Principle A- An anaphor must be bound in a local domain.

Principle B- A pronoun must be free in a local domain.

Principle C- An R-expression is free.

(Chomsky 1986:166)

This paper discusses reflexives in Santhali. The principal device used in most of the Munda languages is the verbal anaphor both in reflexives as well as reciprocals (Subbarao 2012), when the anaphor occurs in a subcategorized position. The optionality of the nominal reflexive when co-occurring with a verbal reflexive as in (1), shows that the nominal reflexive is a secondary device for reflexivity. We hypothesize that Munda languages indigenously did not have of a nominal reflexive. However, due to persistent language contact with adjacent Indo-Aryan languages, Munda languages might have borrowed the nominal reflexive device. In some languages, they have calqued it and in some, they have used the corresponding native forms. Nevertheless, there are also various other functions of the verbal reflexive marker such as an intransitive marker, passive marker and a self-benefactive marker (Subbarao 2012).

Santhali

1. unī- y_i (ac'- lāgit')_i joma?- e khɔj- kan- n- a
 He- agr self- for food SAM ask- pres- vr- [+fin]
 'He is asking food for himself.'

Structural differences with sentences having nominal reflexive:

Most of the Munda languages use a verbal reflexive marker as their first preference of structure to convey reflexivity, than the nominal reflexive in subcategorized positions. The presence of a nominal reflexives is obligatory, if the verbal reflexive is absent as in (2a). However, it is optional when the verbal reflexive is present as in (2b).

- 2b. hɔn-kɔ_i (ac' te)_i arsi- re nel- ke- n- a- kɔ_i
 child- pl self.acc mirror- in see- pst- vr- [+fin]- SAM
 'The children saw themselves in the mirror.'

Reflexivity in Long distance binding

In this paper, we shall discuss long distance binding too. Long distance binding depends upon the position of occurrence of the anaphor and the morphological complexity of the anaphor. It is blocked under the following two conditions.

- a. When the nominal anaphor is either located in a sub-categorized position. or
- b. It is polymorphemic.

It is permitted when the anaphor occurs in a non-subcategorized position (3a).

Anaphor in a non-subcategorized position

- 3a. baha_i arel_j [PRO_j ac'_{i/j}- lagid' cā benaw e -]
 Baha Arel self- for tea make-
 meta- de- (y)a
 tell- [+tr]- [+fin]
 'Baha asked Arel to make some tea for self.'

Anaphor in a subcategorized position

- 3b. məntri_j gidrə_i [PRO_j ac'_i/*_j- tɛ sarhawe-] metak'-
minister child self- acc praise- tell-
'The minister asked the child to praise self.'

The Language called Rana

Ashok Kumar
Research Scholar
Department of Linguistics
Banaras Hindu University
danushlng@gmail.com

Rana is a language spoken in the district of Khiri in the tehasil Paliakalan. This area is known as Tharu schedule tribe area. The people of this region speak a language known as Rana language. There are two other varieties of this language: Kothari and Dangarua. In this paper, I try to trace out the roots of this language and the reasons for its endangerment. There are so many reasons for language loss. One and the most important reasons is the minority status of the language. The speaker of this language is very few in number; this language is neither considered as standard language nor is it used for official and educational purposes. The language does not have written record and is not associated with prestige either. This causes the speakers to shift to the dominant language of the region and subsequently slowly forget their mother tongue. Traditionally, Hindi scholars club this language under the dialects of Hindi and Bhojpuri scholars under the dialects of Bhojpuri, a claim that does not stand linguistic analysis (Kumar 2011). Rana deserves attention of the linguists to be analyzed, categorized and preserved, as the loss of a language is loss of many significant socio-cultural facets of the people and the society at large.

Reference

Kumar, Ashok (2011). "Rana: A language". National seminar on Endangered Language. Anna Malai University, Tamil Naidu.
Mother Tongue Survey of India. 2008

Roots and its technical applications of Manipuri (Meetei) language

Thokchom Kamala Devi
Ph.d. Informatics and Language Engineering
Centre For Technology Studies, School of Languages

Like a human being language is also mortal. It has to be fight for it's survival also. To make a language survive, it is to be in use. The report of the Endangered Language in UNESCO made shock all the people of Manipur. Among the Tibeto Burman Languages in North-East India Manipuri is the only language which has its own Script known as Meetei (Kanglei) script. According to Cheitharol Kumbaba , a chronicle maintained by the state and suppose to be the daily record of all important events, the Manipuri script is said to introduce in 1616 A.D. by Khagemba the then reigning king of the land. Manipuri is the official language and Lingua franca of Manipur. Because inadequate maintainance of Manipur and India Government the status of Manipuri language is becoming poorer day by day. According to UNESCO report it has included in the list of 97 endangered languages in the world. Several Literary words of this language are missing day by day. Structurally Manipuri is agglutinating in nature. Word is formed through three process- prefixation, suffixation and compounding . Words are formed from two types of roots. Roots in Manipuri can be classified into two as:- (1). Free Roots and (2). Bound Roots.

The former can be used separately without a suffix in sentence `whereas the later cannot be used without an affix. Free roots take number and case suffixes expecting the ones that occur with free roots.

1. Free Roots :- Free roots in Manipuri may take siŋ or k^hoy "Plural" to distinguish singular from plural. Such roots may also take case suffixes namely nə "nominative" , bu "accusative" , də "locative", gə "associative" and gi "genitive". Free roots in Manipuri can be divided into six different structures as given below (V for Vowel and C for Consonant).

(I). V	<u>u</u>	"tree"	(IV). CVC	<u>kok</u>	"head"
(II). VC	<u>ok</u>	"pig"	(V). CCV	<u>kwa</u>	"betelnut"
(III). CV	<u>ŋa</u>	"fish"	(VI). CCVC	<u>kwak</u>	"crow"

2. Bound Roots :- Bound roots in Manipuri can be divided into two types as

(1). Nominal and (2). Verbal

(1). Nominal Root :- It consists of three sub types .They are the roots found in

a. Kinship terms - pu "grandfather".

b. Basic Numeral – ni "two"

c. Non-human objects – na "leave"

Such types of bound roots are invariably used with prefix ma "its" i.e. məna "its leave".

(2). Verbal Roots :- The verbal roots may be broadly divided into two as a. Dynamic and b. Stative under the basis of meaning difference and different grammatical functions.

Dynamic Verbs

Phu “beat”

Ca “eat”

Sat “bloom”

Stative Verbs

ten “short”

saŋ “long”

waŋ “high”

All the stative verbs are intransitive (e.g. sa “hot” , pik “small”) but the dynamic words may be transitive or intransitive.

E.g. cen “run” and tum “sleep” are dynamic verbs. They are intransitive verbs. Pi “give” , ca “eat” are transitive. Stative verbs are frequently used as adjective and they can be used in triplication.

E.g. p^hidu ŋaŋsəŋ ŋaŋsəŋ ŋaŋi.

There are number of roots used in both human and non-human entities, such as mit “eye” , kok “head” , na “ear” etc. In such cases roots are treated as free when referred to human being otherwise they are treated as bound roots. The rationale off such treatment is that mə “pronominal prefix” is invariably used with such roots when referred to non-human entities.

The grammatical structure of Manipuri language is fully depended on these roots. By the analysis of this can carry into many different theoretical researches as well as technical development. Word is formed in Manipuri through two processes- Affixation, Compounding. New words can be derived through two process of affixation. How these roots works in the word formation are to be discuss technically which can further developed into various auxiliary language tools in a productive and communicative way.To bring back Manipuri into full secured language, it should be globalized through Natural Language Processing.

Constructing Kannada Text Corpus

Vijayalaxmi .F. Patil, Shahid .M. Bhat

jeeveshwarip@gmail.com, shahid.bhat3@gmail.com,

Linguistic data Consortium for Indian Languages (LDC-IL)

Central Institute of Indian Languages, Mysore

Corpus linguistics has become one of the major areas of Linguistics in which the study of any natural language can takes place as expressed in samples (corpora). Corpus contains a large collection of representative samples of texts, covering different varieties of language used in various domains. It has become an essential media through which modern technology and linguistics can go together. In LDC-IL we are developing corpus for all the Indian languages. In this paper we are focusing on corpus development in Kannada, a south Dravidian language. We have developed 32, 35,370 words of Kannada corpus of various domains like Aesthetic, fine arts, natural, physical and professional sciences, social sciences etc. Out of which approximately 9, 00,000 words are cleaned and ready to be used for various NLP tasks and linguistic research.

This paper is organized in 7 sections. First section will be on Kannada language, second section will be on sampling and data collection, third section will be issues in Data

collection, fourth will be on meta data and data inputting, fifth section will be on corpus cleaning and issues, sixth will be on preprocessing and finally will be on conclusion.

Bibliography

Biber, D., Conrad, S., Reppen R. *Corpus Linguistics, Investigating Language Structure and Use*, Cambridge: Cambridge UP, 1998. [ISBN 0-521-49957-7](#)

Ed. Paul Baker (2009), 'Contemporary Corpus Linguistics'- Continuum International Publishing group, New York.

McCarthy, D., and Sampson G. *Corpus Linguistics: Readings in a Widening Discipline*, Continuum, 2005. [ISBN 0-8264-8803-X](#)

Documentation of Lepcha Literature *Namtho Namthar*:

Interpreting the Indigenous Knowledge and Culture

Satarupa Dattamajumdar (Saha)

dattamajumdards@gmail.com

Documentation and archiving of the available indigenous knowledge (oral or written) of a speech community can be considered as an attempt to 'keep the existence operative within the trace', to understand the heterogeneity of the 'signified' by means of analyzing the 'sign', a means to safeguard the linguistic identity of the members of a speech community. Documentation is seen as a preventive measure to check linguistic erosion especially for a minority speech community like Lepcha. Lepcha, a Tibeto-Burman language spoken in Darjeeling, Sikkim, Bhutan and Nepal has been referred many a times by the scholars (and also by UNESCO) as endangered.

Namtho Namthar are a collection of 'record of literary tradition of the Lepchas' which are either translation works from Tibetan Buddhist text to Lepcha or are concerned with the worship of nature god and myths of the land. The exact time and the authorship of these literary contributions cannot be ascertained. It is believed to have been in existence before the Tibetan migration in the land, i.e., 14th c approximately.

With a view to record the indigenous knowledge of an endangered speech community like Lepcha, and the need for interpreting the indigenous knowledge and culture of this community, the present paper will concentrate upon the documentation and interpretation of the indigenous knowledge and the world view of the speech community by analysing one such *Namtho Namthar* text. The study will particularly meditate upon a *Namtho Namthar* text entitled *Nyolik Nyosung* /nyoli?k nyosun/. The text centers round a mythological story of the land – the divine understanding, universal characteristics of existential understanding and realization of the ultimate truth of life. The paper will characterize the linguistic features of the text with special reference to lexical and morpho-syntactic structures. The co-variance of language and society will be viewed in respect to time and space. Language being the direct reflection of conscious human existence, the present study in course of its analysis will be reflective on the identity of the Lepcha speech community.

The data of the present paper has been collected by means of phonetic transcription (as read out by a Lepcha speaker) using International Phonetic Alphabet, from field investigation at Kalimpong subdivision of Darjeeling district in West Bengal.

The Sociolinguistic Situation in Kinnaur with reference to the language use pattern

Harvinder Kumar Negi
negi.harvinder@gmail.com

Department of Linguistics, University of Delhi

'KINNAURI' is the language spoken by the *kinnaura* tribe that inhabits the Kinnaur region of Himachal Pradesh in India. The language has nine known varieties that varying intelligibility. The KINNAURI varieties can be classified as two generic KINNAURI; one belonging to the Tibeto-Burman family of language and the other belonging to the Indo-Aryan family. Varieties of the generic Kinnauri are; *hamskad, shumcho, hangrami, pooh boti, chitkul-rakcham, kunnu-charang boli, nesang boli*. There is a class variety, spoken by the scheduled caste community in the region - *chamakad* and it has been places in the Indo- Aryan family of languages (Sharma 88).

KINNAURI is one of the least documented languages of the West Himalayish group of the Tibeto Burman family of languages (DeLancey 89, Saxena 95). Baring one variety, which is also the variety of majority and has been documented to some extent, no literature is available on any of the varieties of Kinnauri. As a consequence, much of the rich linguistic and cultural heritage is gradually being lost, as is the case with other tribal languages of the world.

Aim of the present paper will be to describe the linguistic area in the light of previous literature available on KINNAURI and researchers own views. Then the discussion will illustrate the sociolinguistic situation in the region and discuss issues like language use pattern to bring out factors associated with language endangerment so as to propose measures to develop resources for language maintenance.

References :

- Grierson, G.A. 1906 (1967 reprint). *Linguistic survey of India. Tibeto-Burman family*. Volume III, Part I. Moti Lal Banarsi Dass. Delhi.
- Negi, H. 2008. *A Grammar of Kinnauri*. Unpublished M.A. dissertation. University of Delhi
- Saxena, A. 1992. *Finite verb morphology in Kinnauri*. In Cahiers de linguistique - Asie orientale, vol. 24 no. 2, 1995. pp. 257-282
- Sharma, D.D. 1988. *A Descriptive Grammar of Kinnauri*. Sharma, Mittal Publications. Delhi.

Reduplication in Angika

Lucy

singh.lucy@gmail.com

The English and Foreign Languages University

This paper discusses reduplication in Angika language. Angika is an Indo-Aryan language spoken in parts of Bihar, Jharkhand and West Bengal. Angika is not listed in the 8th schedule of the constitution of India. According to SIL Ethnologue reports (1997), the number speakers of this language was 725,000.

Reduplication is a morphological word formation process in which the word is reduplicated completely or partially. It can be of two types: (i) Morphological (ii) Lexical.

Examples of morphological reduplication can be onomatopoeic expressions such as '*jham-jham*' which stands for the sound of rainfall.

(i) *pani jham-jhamaib karai chai*

'It is raining.'

Examples of lexical reduplication is '*jhaat-jhaat*' which represents the idea of loneliness as in:

(ii) *baal-butru ke nai rehla se ghara jhaat-jhaat karai chai*

'The house seems lonely when children are away'

The paper will further discuss the reduplication process in Angika in detail.