



CENTRE FOR ENDANGERED LANGUAGES SIKKIM UNIVERSITY

(A central university established by an Act of Parliament of India, 2007 and accredited by NAAC in 2015)

Report

5 Day Workshop on Scripts and Font Development for Endangered Languages of Sikkim and the Himalayas

23rd-27th September 2019

The Centre for Endangered Languages (CEL), Sikkim University (SU) organised a 5 day workshop on “*Scripts and Font Development for Endangered Languages of Sikkim and the Himalayas*” from 23rd - 27th September 2019 at Barad Sadan Hall, Sikkim University. The workshop was led by Prof. Pramod Pandey, Jawaharlal Nehru University along with Prof. Bhim Narayan Regmi, Tribhuvan University; Dr. Ritesh Kumar, Dr. B.R. Ambedkar University, and Mr. Rupesh Rai, Central Institute of Indian Languages, Mysuru, Karnataka.

Following the call for participation, a total of eighty six participants including eleven (11) from the Centre for Endangered Languages, Sikkim University; eleven (11) from Magar community; six (6) from Bhujel community; six (6) from Gurung community; five (5) from Sherpa community; one (1) from Bhutia community; four (4) from Newar community; two (2) from Lepcha community; seven (7) from Rai community; four (4) from Tamang community; two (2) from Sunuwar community; one (1) from Limboo community; ten (10) from Department of Limbu, Sikkim University; eight (8) from Department of Nepali, Sikkim University; two (2) from Department of Anthropology, Sikkim University; four (4) from Department of Hindi, Sikkim University and two (2) from Department of Lepcha, Sikkim University participated in the said workshop.

On 23rd September 2019, the inaugural session was chaired by the Vice-Chancellor Prof. Avinash Khare. Dr. Samar Sinha, the Coordinator of Centre for Endangered Languages, Sikkim University, in his inaugural remarks, briefed about the establishment of the Centre and its mandate to document endangered languages of Sikkim and North Bengal. He also mentioned that the CEL has established good contacts with the various communities of Sikkim and North Bengal, their ethnic, literary and youth organisations and various other agencies over the last three years. On the theme of the workshop, the Coordinator mentioned about the various indigenous endangered scripts like *Magar Akkha*, *Kripasalyan* (now known as *Rai lipi*), *Khema Pri*, *Kharpa*, *Sambhota* and various other orthographies based on the Devanagari — as the derivatives of Brahmi script. He expounded that a script is a collection of signs used to represent spoken signs. Unlike speech, script is non transient, and is essential for posterity. Writing systems can be placed into broad categories such as alphabetic, syllabic and logographic and of various other systems and designs.

The Vice-Chancellor in his remarks stated that there are more than 370 million indigenous

people spread across 70 countries. Indigenous peoples represent the greater part of the world's cultural diversity, practice unique traditions, and are the holders of unique knowledge systems and beliefs. He reminded that Sikkim University is committed to the development of the indigenous peoples and the endangered languages of the region. It was followed by Vote of Thanks by Dr. Meiraba Takhellambam, an organiser of the workshop.

The first talk in the academic session was given by Prof. Pramod Pandey on “*Going back to the community: Scripts for Under-Resourced Languages of India.*” He mentioned scripts as an instrument of representing speech, and defined *Script* as a set of symbols and diacritics used for representing speech sounds of one or more languages, and a *writing system* as a system governing the use of a script for writing and reading a specific language. He discussed that scripts differ in the shapes and sizes of symbols and diacritics and the convention of relating them to sounds. An example was taken as reference from Pandey (2007) that Arabic depends more on lexical awareness than English and Hindi. English depends more on morphological awareness than Hindi, and Hindi depends more on phonological awareness than Arabic and English.

Prof. Pandey also discussed about the properties of writing systems where he mentioned that as a representational system, orthography needs to adhere to some general usability requirements that are not always compatible. He mentioned that three of these are especially important: Perceptual distinctness, Learnability and Alphabetic Paradox. He also stated that a planar regular language can be generated by a regular grammar — Regularity and Consistency. He greatly emphasised on reading and spelling development mentioning the three major challenges in reading acquisition: Availability - knowledge of the rules of the language, Consistency - spelling to sound relation, phonemic or phonetic, and Granularity of spelling to sound mappings (mapping between sound and size of symbol).

Dr. Ritesh Kumar gave a lecture on “*Character encoding.*” He discussed about internationalisation of character sets, encoding declaration, ISO character sets and Unicode. He briefed about the ten Indic scripts that have been Unicodised, viz. Devanagari, Gujarati, Punjabi, Bengali, Assamese, Oriya, Telugu, Tamil, Malayalam and Kannada. He said that there is one code, different encodings for these ten different Indic scripts. On designing fonts, he discussed the software Font Forge, and demonstrated the installation processes in a computer system on different platforms. He showed how the Indic scripts can be represented as glyphs of the underlying Roman characters using the latter's coding system.

In continuation of the technical session, Mr. Rupesh Rai accompanied Dr. Ritesh Kumar with the installation of the software Font Forge in the computers of the participants and gave a short explanation on it.

On 24th September 2019, Prof. Bhim Narayan Regmi gave a talk on “*Scripts in Nepal: Commonalities, issues and development.*” He focussed on the scripts being used in different languages of Nepal. He added, despite the fact that the shapes are different in Tibetan or Sambhota, Srijanga or Limbu, Newar and Devanagari, the basic system on which the script is built is the same. All the scripts have some features in common (and referred to Yonjan-Tamang (2073BS), Kainla (2073BS), Tamot (2073BS) and Pokharel (2073BS)).

The technical session was led by Dr. Ritesh Kumar and Mr. Rupesh Rai. They helped the participants in developing and designing new fonts for their languages with the help of the software Font Forge. The participants from in and around Gangtok and participants from the

University were keen to learn to design new fonts.

On 25th September 2019, Prof. Bhim Narayan Regmi continued his talk from the previous day dealing with different issues associated with the scripts of Nepal: such as linguistic issue, cognitive issues and socio-cultural issues. He discussed about the sufficiency and redundancy of symbols to represent contrastive sounds of the language, whether the shape/use of a diacritic and arrangement of the symbols follow any system or not. He stated that the symbols/signs are easy to recognise and are distinctive.

Dr. Samar Sinha, Coordinator of the Centre for Endangered Languages, gave a lecture on “*Nepali Varnamala: Contemporary cleavages, claims and contentions*.” He highlighted that a script is a set of visible or tactile signs which are used to represent units of language in a systematic way, and orthography is a set of rules which defines the set of characters and about how to write characters - punctuation and spelling. He gave an introduction to the Devanagari script and its fundamentals. He raised the issues in the Nepali orthography beginning from the *varnamala*, its cleavages like linguistic, traditional, graphical, cultural, technological including “ignorance and arrogance” witnessed in the contemporary debate on the Nepali orthography.

The technical session was continued with the designing and developing of fonts by the participants which was guided by Mr. Rupesh Rai. He also defined that fonts are for defining glyphs, not script. He mentioned that scripts need to be encoded as character and designing font in the Font Forge software is the easiest way to represent some characters as a glyph which resembles character in the script.

On 26th September 2019, Prof. Bhim Narayan Regmi emphasised various socio-cultural issues such as identity and prestige, literacy, education, power, administration, Information and Communication Technology (ICT), linguistic human rights, language preservation and promotion focussing on language documentation and script development. He mentioned that earlier efforts for writing various languages in Nepal begun after 1990. Prior to 1990, the Nepalese people used diacritics to the existing Devanagari characters to represent the sounds of the individual languages which are not found in the Nepali language.

One of the research scholars of the Department of Nepali, Sikkim University, Ms. Reena Rai’s paper on “*A script of Kulung: thoughts and practices*” which was earlier presented in the University of Sydney, Australia on 28-29 June 2019 was also presented. Owing to the personal loss in her family, Ms. Reena Rai’s paper was presented by her supervisor Dr. Samar Sinha, Coordinator of CEL, on her behalf. The paper introduced the Kulung people, their language, linguistic reality of Sikkim - waves of language endangerment in Sikkim. Regarding the writing system for the Kulung language, the paper reviewed the Kripasalyan script (now known as Rai lipi) and Devanagari characters mixed with the IPA symbols. Building upon the advantages of the orthography based on the Devanagari system by Ms. Reena Rai to write Kulung language, the paper highlighted the phonological, orthographic, and Unicode related encoding proving that the devised system of the Kulung orthography based on the Devanagari satisfies the criteria laid down for phonological representation, literacy development, technical requirements, etc.

Mr. Bishnu Lal Bhujel, Resource Person, CEL, SU presented a paper on “*A Devanagari based Orthography for Bhujel Language: Problems and Prospects*.” He introduced Bhujel as a Trans-Himalayan language, which is one of the endangered languages of India and Nepal and it is mostly spoken in India and Nepal. The Bhujel community of Sikkim and Darjeeling aspires to revive their language,

he added. His objective was to develop writing system based on the Devanagari and the Kharpa script. The latter is the indigenous script developed by the Bhujel community of Sikkim. He showed the phonological system of the language, and its representations in the Devanagari and the Kharpa scripts. In his talk, he highlighted how various Bhujel phonemes which have no equivalent representation in the Devanagari script are devised orthographically to represent the same.

On the final day of the workshop, Mr. Rupesh Rai and Dr. Ritesh Kumar did a hands-on session with the participants related to the Font Forge software and discussed how to customise the fonts, its shapes and sizes. They also demonstrated the installation process of the newly designed font in individual computer systems.

After the technical sessions, the workshop concluded with a discussion and feedback session led by the community members and students regarding the issues in their respective languages and orthographies. The session was moderated by the Coordinator, Centre for Endangered Languages, Sikkim University. The following requests emerged from the various indigenous communities regarding their endangered languages in the discussion and feedback session:

1. Training in International Phonetic Alphabet (IPA).
2. Development of technologically compatible font and keyboard layouts.
3. Workshop on writing system to address the issues.
4. Development of writing manual.

The Coordinator assured the community members that their requests will be timely addressed by the Centre for Endangered Languages, Sikkim University being true to the objective of the Centre and of Sikkim University.