

A Manipuri-English Bilingual Electronic Dictionary -Design and Implementation

S. Poireiton Meitei, Shantikumar Ningombam, H.Mamata Devi, Bipul Syam Purkayastha

Abstract— Bilingual Electronic dictionaries are some kind of system that hopes to process natural languages as people have information about words, their meaning and relative information(from source language) to another(target) language. So, machine readable electronic dictionary becomes the central resources for Natural Language applications. Dictionaries and other lexical resources are not yet widely available in electronic form for Manipuri Language. This paper describes the design and implementation of a developing Electronic Manipuri-English Bilingual dictionary. This implementation will provide a more effective combination of traditional Manipuri-English bi-lingual lexicographic information and their conceptual information.

Index Terms: - Manipuri Language, Bilingual, Dictionaries, Tibeto Burman, Natural language Processing.

I. INTRODUCTION

Dictionary is a book in which words are listed alphabetically with their meanings, either in the same language or in another language, and other information about them. Modern dictionaries usually provide phonetic transcriptions, hyphenation, synonyms, derived forms, and etymology. However, a dictionary of a living language can never be complete; old words fall into disuse, new words are constantly created and those surviving frequently change their meanings. A dictionary can also be a book, which gives information about a particular subject, in which the entries are given in alphabetical order. Examples are biographical dictionary, pronunciation dictionary, dictionary of local dialect, dictionary of proverb, dictionary of terminology, dictionary of history. Dictionary of quotations, legal Dictionary and computer Dictionary etc. Electronic dictionaries are also programs that can be downloaded from the Internet or purchased on a CD-ROM or DVD and installed on a desktop computer or on a lap top. Other electronic dictionaries can be searched and consulted online on the Internet. The computer-installed dictionaries can often be consulted directly from within any application that uses editable text. Some electronic dictionaries contain only corpora for a single language (monolingual), but others are bilingual dictionaries and translation dictionaries and may also include medical or legal dictionaries, thesauri, travel Dictionaries, dictionaries of idioms and colloquialisms, a guide to pronunciation, a grammar reference, common phrases and collocations, and a dictionary of foreign loan words.

II. MANIPURI LANGUAGE

Manipuri, locally known as Meiteilon (the Meitei+lon ‘language’), is spoken basically in the state of Manipur which is in the north eastern India. It is also spoken in Assam, Tripura, Mizoram, Burma and Bangladesh. Manipuri is the only medium of communication among 29 different mother tongues. Hence it is regarded as lingua franca of Manipur state. Since 20 August 1992 Manipuri become the first TB (Tibeto Burman) language to receive recognition as a schedule VIII language of India. Before this every date, Manipuri has adopted as the medium of instruction and examination. More and above, Manipuri is taught up to Ph.D. level in the Indian Universities. It is also the medium of education in Manipuri Schools. This language has been a major adhesive in the integration of all ethnic groups of Manipur. In the classification of Sino-Tibetan family, Manipuri is placed in kuki-chin proper separated by Grierson-Konow (1903-1928), in Meitei branch of Kukish Section by Safer (1955, 1966, and 1974).

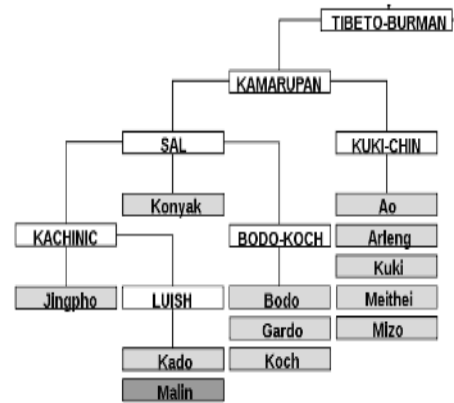


Fig 1: language tree for Manipuri (Meitei+lon)

(<http://www.translationdirectory.com/articles/article2300.php>)

Manipuri share genetic features of TB (Tibeto Burman) features including phonemic tones, SOV(Subject Object Verb) word order, agglutinative verb morphology and tendency to reduce disyllabic forms to monosyllabic ones (Delancy 1987). Grierson (1903) in his LSI, Manipuri is briefly described: vowels, consonants, word categories syllable, sentence type etc. however neither of this word is exhaustive in its description.

III. HISTORY OF MANIPURI- DICTIONARY

The development of Manipuri Dictionary started from 1960's. In 1964 the first Manipuri Dictionary **Manipuri to**

Manipuri and English Dictionary compiled by Shri Ningthoukhongjam Khelchandra was published. The entries in this dictionary are made in Manipuri and definitions are given in both Manipuri and English or in English only in some places. In 2004 **Manipuri to English Dictionary** compiled by Dr. Imoba was published. In 2005 the second edition of **Manipuri to Manipuri and English dictionary** compiled by Shri Ningthoukhongjam Khelchandra was published. In 2006 **Learners' Manipuri-English Dictionary** compiled by Dr. Surmangol Sharma was published. In this dictionary the entries are made in roman transliteration and it is again written in **Meitei Mayek** (Meitei script) and its meaning is written in English.

IV. WRITING STYLE AND GRAMMAR

The writing style of Manipuri Language is from left to right. SOV (Subject Object Verb) word order is obeyed in Manipuri Language. There are two types of roots-the free roots and bound roots. All nouns (except the kinship terms), in Manipuri, are free roots. There is also no grammatical gender in Manipuri, the human and animal. Meitei had its original script named Meitei-Mayek which was in use up to the 18th century. However, later the Eastern Nagari or Bengali script adopted for scripting Manipuri. It is also the script used till now. However, efforts to revive the Meitei-Mayek are still on. The Grammar of the Manipuri Language is interesting yet simple; the phonological system of the Manipuri Language can be basically divided into three levels- Vowels, Consonants and Tones. In Manipuri nouns are addressed according to the masculine or feminine on the basis of the natural sex.

V. ELECTRONIC DICTIONARY

Now, the computer has made the creation and manipulation of information the central activity of the economy. This era has also been known as the information age. If information explosion is the most pronounced phenomenon of the present era, the Internet is and will remain its greatest enabler. This explosion is triggering a global knowledge revolution. In the early 1990s computer technology made possible the release of dictionaries on floppy disks or CD-ROM, e.g. the electronic edition of The Random House Unabridged Dictionary (1993). Electronic dictionaries also became available as part of multi-volume reference book packages such as Microsoft's Bookshelf CD-ROM and as a feature of on line services Computer technology provided new ways to search for and link words and new ways to illustrate them, e.g., prerecorded pronunciations that users can play back. By the end of the 1990s many dictionaries were available in various print and electronic editions; the new Encarta World English Dictionary (1999) was released both as a printed book and a CD-ROM.

VI. COLLECTION OF DATA

The data are collected from many sources. Some of them are as given below.

A. Daily Newspapers

Words collected from the daily local newspapers become one of the sources of data collection for the work. Some of local newspapers published in Manipur are as.

- a. The Sangai Express.
- b. The Poknapham.
- c. The Naharolgi Thoudang.
- d. The Yakairol.
- e. The Huiyen Lanpao etc

With the development of internet and Information technology, we can download and read some of the newspapers daily directly through the internet from anywhere throughout the world.

B. Weekly or monthly journals

We also collect the words from the journals which published weekly or monthly. Some of the journals areas.

- a. Sanathong
- b. Kangla Pao.
- c. Eastern Frontier. Etc

C. Manipuri Text Corpus

Manipur corpus is the backbone for researchers of Manipuri language. Some of the words are also picked up from the Manipuri Corpus which is created in Manipuri University under the aegis of Prof. Ningomba and Dr. N.Pramodini.

D. Manipuri lexicons (paper dictionaries)

- a. Manipuri to Manipuri and English dictionary
Compiled by Sri Khelchandra Singh.
- b. Manipuri to English Dictionary
Compiled by Dr. Imoba.
- c. Learner's Manipuri-English Dictionary
Compiled by H. Surmangol Sharma. Etc

VII. METHOD USED IN WORD SEARCH

We are using **Binary search** algorithm for our dictionary searching mechanism. Algorithm is quite simple. It can be done either recursively or iteratively.

1. get the middle element;
2. if the middle element equals to the searched value, the algorithm stops;
3. otherwise, two cases are possible:
 - ✓ Searched value is less, than the middle element. In this case, go to the step 1, before middle element.
 - ✓ Searched value is greater, than the middle element. In this case, go to the step 1, after middle element.

Now we should define when iterations should stop. First case is when searched element is found. Second one is when no elements. In this case, we can conclude, that searched value doesn't present in the dictionary database.

VIII. IMPLEMENTATION OF SOFTWARE AND DICTIONARY DATABASE

The software's used for designing the Electronic – Dictionary is as:-

Frontend: - **PHP** (Hypertext Preprocessor)

Backend: - **MYSQL SERVER**

WEB SERVER: - **UNISERVER (Orion)**

IX. DICTIONARY DATABASE AND SEARCH TECHNIQUES

The software which is proposed for the development of this Bilingual dictionary is: PHP as front-end and MYSQL database as backend database. Here (Fig.2) shows the context diagram of the data flow diagram (DFD) of the Manipuri-English Electronic bilingual dictionary. The USER gives the word to search in the dictionary. The DICTIONARY SEARCH program searches the word from the DICTIONARY DATABASE. If the word is in the database then the program will return the word with its corresponding information through the dictionary interface. If the word is not in the database then we can enter the word with its corresponding information through the PERSONAL-DICTIONARY-ENTRY.

The structure of this dictionary (lexical) database is divided as:

1. *Root word*:- the main word in the dictionary
2. *Lexical item*:- the part of speech category of the word
3. *Transliteration*:-transliteration of the word in (target language) English
4. *Meaning in English*:- meaning of the word in English
5. *Example in Manipuri*
6. *Example in English*

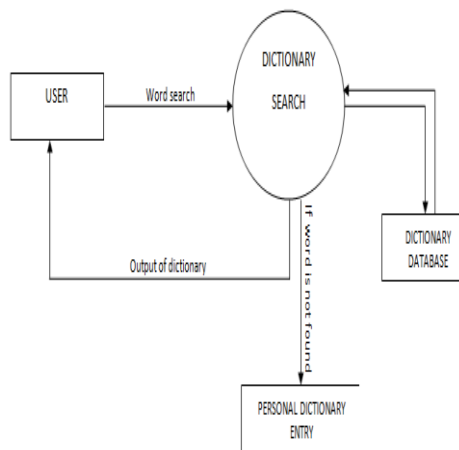


Fig.2: Dictionary Search DFD (Context Diagram)

Table I. Table structure in the MYSQL dictionary Database

Field	Data Type	Null	comment
root	varchar	no	root word
p.o.s	varchar	no	part of speech of the Dictionary word
trans	varchar	no	transliteration of the Dictionary(Manipuri) word in English
e.meaning	varchar	no	meaning in English
man.ex	varchar	no	example in Manipuri
eng.ex	varchar	no	example in English

The algorithm of word search in the dictionary is given as:

1. START.
2. Search of the word in the Dictionary (BINARY SEARCH).
3. If word is found. Then Print the word and its corresponding information: STOP.
4. If word is not found: and do you want to Entry it in PERSONAL DICTIONARY ENTRY.
 - 4 (a). If (Option="yes")
Personal Dictionary entry is done: STOP.
 - 4 (b). If (Option="no")
Display a message: Word is not found: STOP.

X. CONCLUSION AND FUTURE WORK

This paper is to show that the Electronic-Dictionary which is going to develop has gained positive remark from computer experts and experienced Manipuri language users. The paper has also shown that Electronic-Dictionary can help the teaching and learning process of Manipuri language with savvy, smooth and time-saving application. It also has educational values on inspiring the interest of students to learn and helping the instructors to achieve the objectives of lesson. This paper is to be a contribution to the quality of education in Manipuri language corresponding with the development of education technology. Further research should be carried out either directly or indirectly by those who involved in the study of Manipuri language. It is to ensure that Manipuri language is highly valued and continues to be developed in future. And finally we need young researchers and linguistics to develop CLIR(cross lingual information retrieval systems) and MT(machine translation) for Manipuri Language as future work.

REFERENCES

- [1] Bhat, D. N. S. and Ningomba "Manipuri grammar" Munich: Lincom Europe, 1997.
- [2] Chelliah, Shobhana L. "A grammar of Meithei" Berlin: Mouton de Gruyter. ISBN 0-19-564331-3, 1986
- [3] Delaney, Scott "The Sino-Tibetan languages" in B. Comrie (ed.) The World's Major Languages, New York: Oxford University Press, 1987.
- [4] Ch. Yashwanta Singh "Manipuri Grammar", Rajesh Publications, New Delhi, 2000.
- [5] G.A.Grierson LSI, Vol III. Part III .1st Edition Motilal Banarasidas, Delhi, 1903.
- [6] H.Mamata Devi, S.Poireiton Meitei, B.B.Chaudhuri "Development of a Manipuri-English Electronic Dictionary":



ISSN: 2277-3754

ISO 9001:2008 Certified

International Journal of Engineering and Innovative Technology (IJET)

Volume 2, Issue 1, July 2012

In the XI Manipur Science congress, Manipur University. pp
(32-33), March, 2008

- [7] S. Poiraiton Meitei, Shantikumar Ningombam, Prof. Bipul Syam Purkayastha, "An Analysis towards the Development of Electronic Bilingual Dictionary (Manipuri-English) -a report", International Journal of Computer Science and Information Technology (IJCSIT) ISSN: 0975-9646, Vol. 3 (2), 2012, pp 3423-3426
- [8] Singh Imoba "Manipuri to English Dictionary", 2004.
- [9] Sharma H. Surmangol "Learners Manipuri Dictionary", Sangam Book Store, Paona Bazar, Imphal, 2006.
- [10] T. C. Hudson "The meiteis", Neeraj Publication, House, New Delhi, 1975.
- [11] www.dictionary.com.
- [12] www.alphadictionary.com
- [14] www.translationdictionary.com
- [15] www.virtualbangladesh.com