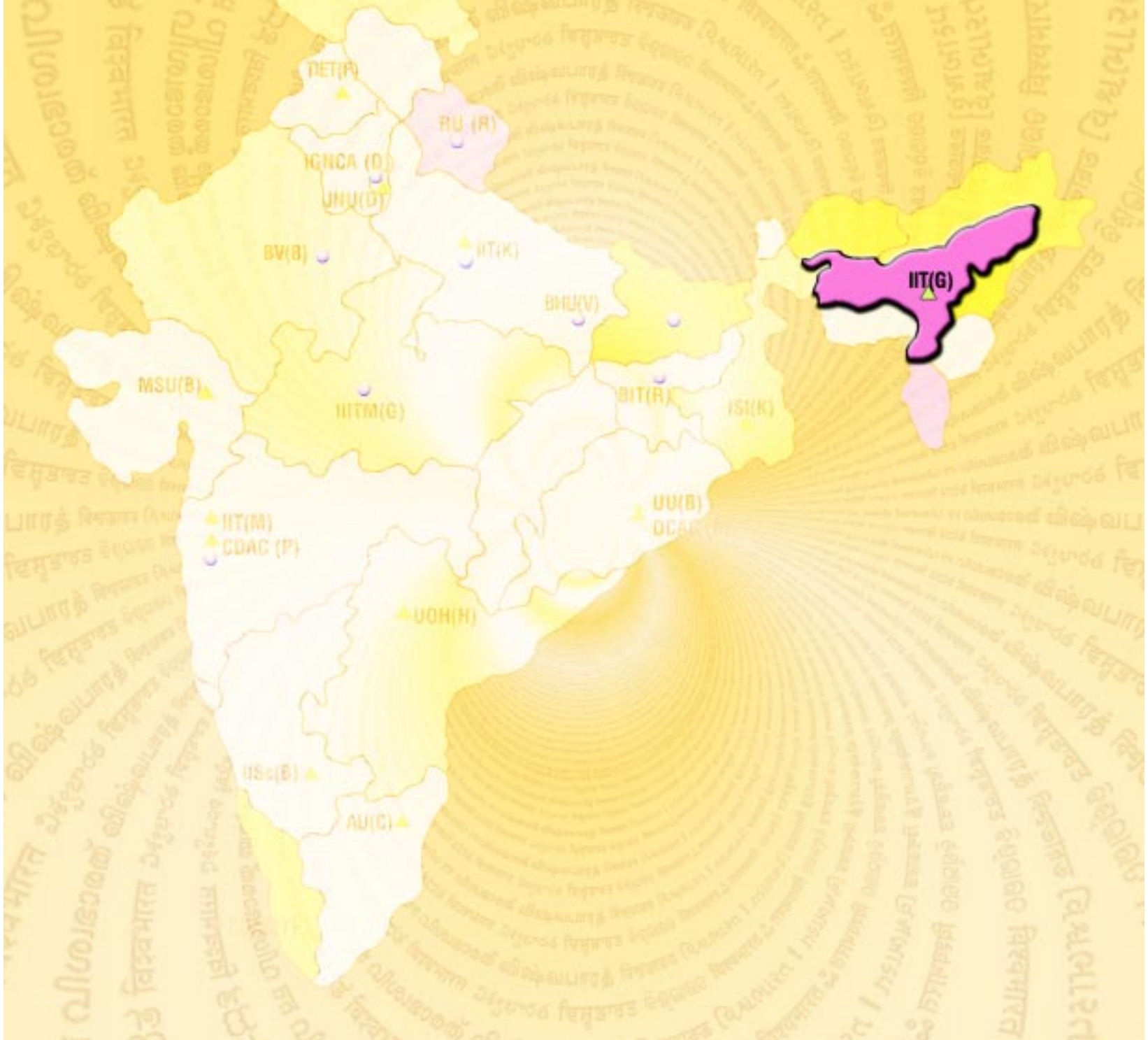


Resource Centre For
Indian Language Technology Solutions – Assamese & Manipuri
Indian Institute of Technology, Guwahati

Achievements



Indian Institute of Technology, Guwahati
Department of Computer Science & Engineering, Panbazar
North Guwahati, Assam-781031 India
Tel. : 00-91-361-2691086 E-mail : tdil@iitg.ernet.in
Website : <http://www.iitg.ernet.in/rcilts>



**RCILTS-Assamese & Manipuri
Indian Institute of Technology, Guwahati**

Introduction

Conceived in the millennium year, the Resource Centre for Indian Language Technology Solutions Indian Institute of Technology Guwahati, is a project funded by the Ministry of Information Technology of the Government of India. The main objective of the Centre is to make electronic information available in native languages mainly Assamese and Manipuri thereby aiding the dissemination of information to the larger masses. This Centre is one among several in the nation and is equipped with all modern systems and language related software. Four investigators from the Indian Institute of Technology Guwahati and one collaborator from Guwahati University are presently involved in research in the areas of Natural Language Processing, Speech Recognition and Optical Character Recognition Systems. The Centre is equipped with all modern equipments, which include several Pentium-III and Pentium-IV machines, a Linux server and a Windows2000 server, scanner, web camera, printers, digital camera and audio recording systems. Eight project personnel, five technical and three linguists working in joint collaboration with Guwahati University, man the Centre.

The Centre has developed various language related tools and technologies over the last two and odd years. Details of the same follow:

1. Knowledge Resources

1.1 Corpora : A systematic collection of speech or writing in a language or variety of a languages forms a language corpus. The two corpora created by the Centre are-

- **Assamese Corpora:** Seven Assamese novels have been transformed into electronic form to form the base corpus. The salient features are
 1. Number of words : 6,00,000
 2. Fonts used : AS-TTDurga (C-DAC font) and Geetanjililight (A popular font used in DTP work).

3. Encoding Standard: ISCII.

- **Manipuri Corpora :** The Manipuri corpora were received from Ministry of Information Technology, Govt. of India. This corpora creation started at Manipur University under the aegis of Prof. M.S.Ningomba and Dr. N.Pramodini. The corpora were in LP1 format, which is compatible to Leap Office. The total word count comes out to be around 3,60,000. The fonts used in the creation of this corpus are BN-TTDurga, BN-TTBidisha and AS-TTBidisha. Further investigation to make it compatible to existing systems is in progress.

1.2 Dictionaries : Derived from the Latin word *dictio* (the act of speaking) and *dictionarius* (a collection of words), a dictionary is a reference book that provides lists of words in order along with their meanings. Dictionaries may also provide information about the grammatical forms, syntactic variations, pronunciation, variations in spellings, etymology, etc. of a word. The Centre has developed e-dictionaries, the details of which are shown in the Table 1.

Dictionaries	Root Words	Total words
English-Assamese	5000	49,627
Assamese-English	2000	9,015
English-Manipuri	2500	36,778
Manipuri-English	3000	7,228

Table 1. Dictionary details

Each entry in the Dictionary contains information on – i) the Grammatical form, ii) Meanings iii) Synonyms iv) Antonyms, v) Pronunciation of the word (in the form of sound files) vi) Transliteration in English, vii) Soundex code and viii) Semantic category. More words are being added on.

Features

- a) Provides information about a word.

- b) These dictionaries have been structured to support the spell checker and machine translation systems being developed at the Centre.

1.3 Design Guides

A design guide gives a brief overview of a particular language. The main topics covered by the design guide are the consonants, vowels, conjuncts and matras which form the character set of the language and numerals, punctuation marks, month names, weekday names, time zones, currency, weights and measurements used in that particular language. Additionally, some linguistic information such as Phonological features, grammatical features, the history of the language and the geographical description of the state where the particular language is spoken are also included.

The Resource Centre for Indian Language Technology Solutions, Indian Institute of Technology Guwahati has designed two design guides for Assamese and Manipuri.

1.4 Phonetic Guides

Pronunciation is also an important part in language learning for which we have a Phonetic guide in the website. This guide describes how a particular alphabet should be pronounced. The International Phonetic Alphabet (IPA) is used in the dictionaries as well as in the description of the languages (North –East India) available in the web site.

2. Knowledge Tools

2.1 The RCILTS, IITGuwahati Website (<http://www.iitg.ernet.in/rcilts>)

As part of its core objectives the Centre hosts a website that offers a wide variety of information ranging from Assamese and Manipuri languages to Geographical and Cultural issues.

Information hosted may be categorized as:

- (i) **Linguistic Information:** The Northeast is known for its large diversity in languages.

Presently the website holds information on sixty-five of the existing languages.

- (a) **Northeastern Languages :** The website contains information on sixty-five North-Eastern languages like Assamese, Chokri, Chakesang, Zeliang, Pochuri, Lotha, Sangtam, Deuri, Dimasa, Kokborok, Tintekiya, Koch, Hurso, Miji, Chang, Khiamngamn, Konyak, Nocte, Phom, Tangsa, Wanchoo, Hmar, Karbi, Kuki, Lakher, Manipuri, Mizo, Riang, Khasim Monpa, Takpa, Tsangla, Sherdukpen, Sulung (Puirot), Adi, Apatani, Bori, Mishmi, Mishing, Nishi, Tagin, Liangmaialong with their classification.
- (b) **Linguistic Map :** Preliminary work commenced with the compilation and design of the Linguistic Map of the Northeast. After acquiring published and unpublished material on these languages of the Northeast and with inputs from the Census Department, Surveyors, Linguistic fieldwork and findings of Dr. Dipankar Moral, the Linguistic Map came into being. The website presently has a coloured updated version of this map. The size of the font connotes the density of speakers of a particular language in that area.



Figure :1 Resource Centre's Home page

- (ii) **On-line Dictionary:** A transliterated Assamese dictionary has been put up on the website. It carries information on the Meanings, (in Assamese and English), Grammatical categories and the Pronunciation.
- (iii) **Geographic and Demographic information:** Both geographic and demographic information are made available for each of the seven Northeastern states of Assam, Meghalaya, Manipur, Nagaland, Mizoram, Tripura and Arunachal Pradesh.
- (iv) **Guides:** Guides have been put up on the website to aid users to understand language related issues. Two such guides that are currently are:
 - (a) **Design Guides:** Design guides for both Asamiya and Manipuri that give a general idea of the respective language together with certain frequently used sentences has been hosted to aid a novice get a basic idea about the language.
 - (b) **Phonetic Guides:** Pronunciation is an integral part of the language learning mechanism. A Phonetic Guide describes how an alphabet should be pronounced. The International Phonetic Alphabet (IPA) has been used in the dictionaries as well as in the description of the languages (North–East India) to convey the manner in which the words are pronounced. The Phonetic Guide allows the user to interpret the same correctly. Figure 2 depicts the manner in which one can correlate the IPA symbols to the normal orthography especially for the pronunciation of vowels using the phonetic guide. Figure 2 depicts the IPA chart for vowels.

Vowels

IPA	English Word	Phonetic Representation
i	pin	pin
e	pen	pen
ε	Pat (in the production of	pæɪ

vowels the upper surface of the tongue is always convex, hence the result in tongue height. In this context it may be mentioned that Asamiya/ ε /is slightly higher than English / æ /as in / pæɪ / pæɪ

a	farm	fam
u	too	tu
ʊ	put	put
o	saw	so
ɔ	got	got

Figure 2. Vowel IPA chart

- (v) **Web based Dictionary:** A transliterated Assamese Dictionary has been put up on the website. This dictionary is possibly the first contemporary standard on-line Assamese dictionary. Moreover, this dictionary is designed in such a way that it will be helpful for the non-native speaker. It provides information on a) English Word b) Grammatical Category c) Meaning d) Pronunciation and e) Assamese Meaning.
- (vi) **Bookmarks:** Additional Links to various Newspapers of Assam have been provided. A Manipuri News letter titled “Manika” has also been put up. Further links to the various tourism related sites of Assam and Manipur, News Portals and other Resource Centres are also made available in the website.

2.2 Fonts : The Centre has developed a True Type Font “Asamiya” using Fontographer software.

This font can be used for typing Assamese text in Microsoft Word. Fine Tuning of the existing Font

Set has been done where the spacing and Matras have been adjusted. A Manipuri True Type font has also been created.

ক খ গ ঘ ঙ
চ ছ জ ব ঞ
ঠ ঠ ড ঢ ণ

Figure : 3 Snapshot of Assamese font “Asamiya”

2.3 Spell Checker

A Spell Checker forms a vital ingredient of a word processing environment. The basic tasks performed by a spell checker include comparison of words in a document with those in a lexicon of correct entries and suggesting the correct ones when required. The two commonly used methods for detection of *non-words* are the Dictionary look-up and the N-gram analysis. Isolated-word error correction is achieved by either of the Minimum edit distance technique, the Similarity key technique, Rule-based methods, N-gram, Probabilistic and Neural Net techniques. Context-dependent error correction methods usually employ Natural Language Processing (NLP) and Statistical Language Processing for correcting *real-word* errors.

Spell checking strategy for Assamese

The development of a Spell Checker for Assamese has been undertaken at the Resource Centre, and code has already been developed in Perl. The Spell Checker exists in the form of separate modules for error detection and correction, as well as a stand-alone system in which all the spell checking routines have been integrated. The strategy used is described below.

Non-Word Detection

The *non-words* are detected by looking up document words in a dictionary of valid words. The dictionary used is actually a word list in which 5000 distinct Assamese words have been extracted from the English-to-Assamese online dictionary developed by the Centre, and the rest of the words

were taken from a corpus of about 67,000 words. This means that the look up dictionary contains around 72,000 words. A hash table has been used as a lexical lookup data structure. The performance of a dictionary using a hash table is quite adequate, even for complex access patterns. Perl has an efficient implementation of a hash table data structure, which has been used in the non-word detection module.

Isolated-Word Error Correction

A three-pronged strategy has been used for generating suggestions comprising of the Soundex, Edit-distance and Morphological processing methods.

Soundex Method

This method maps every word into a key (code) so that similarly spelled words have similar keys. A Soundex encoding scheme for Assamese has been designed based on the encoding scheme for English which comprises of a set of rules for encoding words and 14 numerical codes. The Soundex code of the misspelt word is computed, and the dictionary is searched for words, which have similar codes. An example is given below (Table 2).

Word	Soundex code
মৰম meaning affection	ম৬৫
স্বাধীন meaning independent	স্বা৩৫
ফঁহিয়ালেহেঁতেন meaning would have analyzed	ফ৭৯৪৭৩৫

Table : 2 Assamese words and soundex codes

Edit-Distance Method

Candidate suggestions are obtained by ‘Damerau’s error reversal’ method. The four well known error-inducing edit actions are the insertion of a

superfluous letter, deletion of a letter, transposition of two adjacent letters and substitution of one letter for another. In 'Damerau's error reversal' each edit action is applied to the misspelt string and a set of strings is first generated. These are checked in the dictionary to see which of them are valid words to finally produce the suggestions.

Morphological Processing

Morphological analysis is performed to extract the root word from the misspelt word as also a list of valid affixes that can be attached to that root word. By attaching the affixes that closely match those of the misspelt word to the root word, a list of suggestions is generated.

A module for ranking of suggestions, based on Minimum edit-distance methods has also been developed.

A run-time snapshot of the Assamese Spell Checker is shown in Figure 4.

Features of The Stand-Alone Spell Checker

- GUI developed using Perl/Tk, with simple text editing facilities.
- Currently supports C-DAC's AS-TTDurga font.
- Assamese text files can be loaded onto the GUI, edited, and saved.
- Misspellings can be marked by clicking the 'Det' ('Detect') and then the 'Show' button.
- Selecting the misspelt word and clicking 'Sug' ('Suggest') button can generate suggestions.
- A facility to add new words to the dictionary exists.
- 'Select All/Unselect All' option available in the 'Edit' menu on right clicking on the text, for selecting/deselecting entire text.
- 'Copy'/'Cut' and 'Paste operations possible for text within the GUI window.

- Text can be copied/cut from an I-leap document and pasted onto the GUI window.

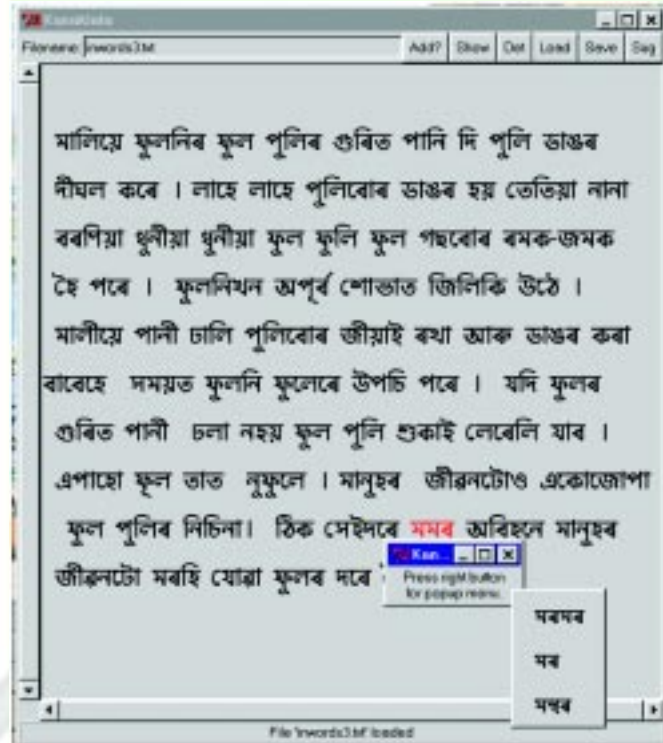


Figure 4. The Assamese spell checker.

The Soundex encoding scheme for Assamese has been refined into a more fine-grained one, and now comprises of 21 numerical codes. Added functionality has been incorporated into the Soundex code generator for handling matras attached to consonants and conjuncts in the first letter position, as also for khandata/chandrabinu/ anuswar/bisarga attached to consonants, vowels and conjuncts.

The Edit-distance module now accounts for 2234 distinct letter combinations (matras attached to consonants/conjuncts, khandata/chandrabinu/ anuswar/bisarga attached to consonants/ vowels / conjuncts.)

Tests conducted against a corpus of about 67,000 words reveal that the Edit-distance method gives the best results, followed closely by the Soundex method. A non-word detection module for Manipuri has also been developed. Work has also been undertaken to integrate a spell checking facility for Assamese in the Microsoft Word environment.

2.4 Assamese language support for Microsoft Word : Word processors like Leap Office provide the facility of document typing in Indian languages. Content creation can be done with the help of these editors by using the inscript keyboard layout (provided by editors). The Centre has developed a macro that allows the user to type Assamese text in Microsoft Word without the need of available Indian Language Editors.

Technology Description

This Microsoft Word macro maps the inputs to appropriate glyphs. The macro supports the Inscript keyboard layout. Typing can be done using the font “*Asamiya*” developed by the Centre. The use of the Inscript keyboard layout facilitates smooth migration from C-DAC to Microsoft technologies.

Features

- Inscript keyboard layout
- Template macro
- Supports “*Asamiya*” font and C-DAC Assamese fonts (As-ttDurga, As-ttBidisha, As-ttDevashish, As-ttkali, As-ttAbhijit).
- All the features of MSWord (such as justification, different font styles and different font sizes etc.) can be used.
- Documents are stored in glyph form.

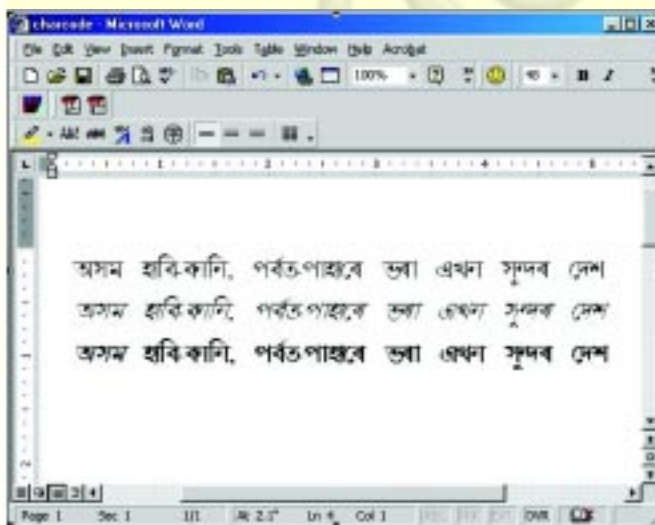


Figure 5. Snapshot of Assamese macro with different font styles.

2.5 Morphological Analyzers : Given any word or a group of words, a morphological analyzer determines the root and all other inflectional forms. Morphological Processing plays a vital role in the development of Spellcheckers and Machine Translation Systems. The Centre has developed morphological analyzers for both Assamese and Manipuri.

a) **Assamese Morphological Analyzer:** The analyzer has been developed for use with the Spell checker and the Machine translation systems.

Technology Description Stemming technique forms the base of the Assamese morphological analyzer. In the above technique, affixes are added/deleted according to the linguistic rules. The derived words are verified with the existing Corpus/Dictionary to treat as valid words.

Features

- Currently works with Eleven linguistic rules
 - More rules can be added without alteration in code.
 - Modules are available in the form of API's for customization.
- b) **Manipuri Morphological Analyzer :** A design for developing a Morphological Analyzer for Manipuri is being investigated. Development of a root dictionary, a morpheme dictionary and an affix table has commenced. Identification of linguistic rules for nouns and pronouns has also commenced. Further linguistic rules for other grammatical categories are being studied.

Technology Description

The Manipuri Morphological Analyzer will be realized using the same techniques used in the development of Assamese Morphological Analyzer.

Features

- Five rules are being used currently

- Can be easily upgraded for more complex patterns
- The modules, which have been developed so far are available in the form of API's and can be used according to the need by other applications like, spell checker, etc.
- A Graphical User Interface has been developed (see Figure 6).



Figure 6. Graphical User Interface of Manipuri Morphological Analyzer

3. Translation Support Systems

3.1 Machine Translation System The term Machine Translation (MT) refers to the process of performing or aiding translation tasks involving more than one human language. It is a system that translates natural language from one source language (SL) to a target language (TL).

Technology Description

This MT system developed is basically a rule based one and relies on a bilingual dictionary. It can currently handle translation of simple sentences from English to Assamese. The dictionary contains around 5000 root words. The system simply translates source language texts to the corresponding target language texts word-for-word by means of the bilingual dictionary lookup. The resulting target language words are re-organized according to the target language sentence format. In order to improve the output

quality, the system performs morphological analysis before proceeding to the bilingual dictionary lookup.

Currently it can handle general-purpose simple translation. It is being upgraded to handle complex sentences. Efficiency of the system can be improved by selecting a specific domain. Currently the rule based Machine Translation system contains 22 rules. The system has been tested for more than 250 frequently used simple sentences. An example of Machine translation from English to Assamese is depicted in Figure 7.



Figure 7. English-Assamese Machine Translation System

4. Human Machine Interface Systems

4.1. Optical Character Recognition for Assamese and Manipuri: Optical Character Recognition (OCR) is the process of converting scanned images containing texts into a computer processable format (such as ASCII, ISCII, and UNICODE etc.). An MoU to transfer the OCR technology from the Resource Centre at the Indian Statistical Institute, Kolkata was signed in August 2002 and accordingly the same was effected in September 2002.

Technology Description

The system takes a gray level (8bit) TIFF (Tagged Image file format) image as input. The images scanned from either Assamese (or Manipuri) books or from paper documents can be processed by the OCR software. The current version of the OCR system produces output in the ISCII format,

which can be viewed or edited using any editor supporting ISCII.

Features: Table 3 shows the salient features of the OCR system while Figure 8 shows the associated GUI.

Features	Specifications of Assamese & Manipuri OCR
Scanning resolution	300-600 dpi
Input image	TIFF (8 bit gray level)
Skew detection	+5 to -5 degrees
Skew correction	+5 to -5 degrees
Font name	Assamese: Gitanjalilight, Luit & AS-TTDurga Manipuri: Font used by the publisher, Manipuri Sahitya Parishad, Imphal
Font size	Assamese: 12-28 points Manipuri: 12-18 points
Test data size	Assamese: 600 pages from books & printed documents Manipuri: 100 pages from books.
Template size	Assamese: 2600 Manipuri: 1800
Post processing	Morphological analysis
Output file format	ISCII format
Accuracy	Assamese: 95% (without post processing) Manipuri: 90% (without post processing)

Portability/ Expandability	Windows 98/2000/XP & LINUX
-------------------------------	-------------------------------

Table 3. Specifications of the Assamese/Manipuri



OCR System

Figure 8. GUI of the OCR System

4.2 Speech Recognition System

Automatic Speech Recognition (ASR) for Assamese, which is concerned with the problem of recognition of human speech by a machine, is the core of a natural man-machine interface. A speaker-dependent continuous speech recognition system with a list of vocabulary has been developed for the Assamese language. The objectives of the system are:

- Assamese and English Spoken Digit Recognition
- Study of Noise Effects on Assamese Recognition
- Phonetic Alignment of Assamese Digits

A set of acoustic rules has been formulated from the acoustic-phonetic features of Assamese and English that will be able to classify a given sound. The rules based on these features show an overall success rate of around 76% for randomly collected test utterances in Assamese. The classification rate is least among the sounds falling under the nasal class (65% success rate). The sounds falling under the vowel, stops and the fricatives show better results (80%, 85% and 85% success rate respectively) while the classification results for diphthongs (70% success rate) were not so good.

The unique Assamese sound of /x/ is found to be very similar to /s/ acoustically. A comparative study of acoustic properties of Indian spoken English and Assamese vowels has been performed

4.3 Interface for e-Dictionary : An interface for the English-Assamese and Assamese-English dictionaries has been developed that allows users to choose between one of the two languages viz. English and Assamese, enter a word and find its equivalent in the other language. Based on *Client-Server* Architecture, the system allows users to access the Dictionary Server using a Java applet opened via a browser.

Technology Description: The basic components that make up the on-line dictionary are:

1. **Client:** The client is provided with a Graphical User Interface (GUI) in the form of a Java applet shown in Figure 9 and initiated from the web browser. The user can invoke the URL of the dictionary and choose the source language; type the word in the search window provided in the applet and click on the *Search* button. The query from the user is sent to the server and the results are displayed at the client-end.



Figure 9 Graphical User Interface of the dictionary

2. **Web Server:** Apart from the main html page, this server hosts the Java jar files and serves the applet classes to the client.
3. **Dictionary Server:** Coded in Visual Prolog,

the Dictionary Server has been developed as a menu-based application that assists the administrator to maintain the dictionaries, to stop and restart the server and monitor incoming requests. The server provides information on the requests and the actions taken during run time. The current version runs on a Windows platform but the can be ported to a Linux system with minor changes, making it virtually platform independent. Whenever a connection to the server is established, it checks the incoming request from the client, consults the database, searches for information on that particular word after which it acknowledges the request by serving the information.

4. **Database:** The e-dictionaries that comprise the database were initially developed using MS Access. These are then converted to Prolog facts for faster access by the search engine. The server provides the administrator an option to open an ODBC (Open DataBase Connectivity) connection to the MS-Access database and perform the conversion thereby uploading the latest versions of the dictionaries. The current version of the system supports two (English-Assamese and vice versa) of the four electronic dictionaries (viz. English-Assamese, Assamese-English, English-Manipuri and Manipuri-English) being created.

Features

- a) User Friendly GUI available.
- b) Dictionaries can be used to aid machine translation and/or spell checking applications.
- c) Pronunciations of words are available in the form of wave files.
- d) APIs provided for prospective programmers to allow them to mould the dictionary information to suit their custom applications.

- e) A stand-alone and a web enabled version of e-dictionaries available.

5. Language Technology Human Resource Development

5.1 Workshops Conducted

Two workshops were conducted to disseminate and make people aware of how to use Indian Languages in IT.

- (i) A one-day workshop on “Natural language Processing” was conducted at the Indian Institute of Technology Guwahati on the 31st of March, 2001. The workshop was intended to act as a forum for promoting interaction among interested graduate students and researchers working on natural language processing and allied areas. The topics that were addressed in the workshop were :

- Introduction to Computing
- Natural Language Processing
- Linguistics
- Demonstration of language technology solutions

- (ii) A Training Programme on “Web page Design and Office Automation Using Assamese” was held at the Indian Institute of Technology Guwahati on the 15th & 16th of March, 2002. The Training Programme was designed for the benefit of Web Designers and people who work with Assamese Word Processors. With State Government officials using Assamese as a medium of communication, the need for Office Automation in this language is picking up fast. Likewise, the demand for putting up information in the native language on the web is also increasing. This training programme was aimed at disseminating techniques to generate e-content in Assamese and use this language for routine office tasks.

The training programme was specifically designed to encourage state government employees to use Assamese for office automation. Participants of the programme ranged from employees of Central, State and Public sector undertakings to Private entrepreneurs. The course contents of the workshop were :

- Office Automation
- Web development in Assamese.
- Creation and manipulation of Database and Spreadsheets in Assamese.
- Hands-On session.

6. Standardization

A draft of the Assamese Unicode Code-Set was prepared after due consultation with linguists and the State Govt. and submitted to the Ministry of Communications & Information technology, Govt. of India for further evaluation. The Assamese Code –Set is similar to that of the Bangali, exceptions being.

- The letter Ra ৰ has Code 09F0 instead of 09B1 and the letter Wa ৱ has code 09F1 in lieu of 09B5.
- An additional letter Khya ৰ্খ with code 09BB is introduced.

7. Publications

1. Monisha Das, S.Borgohain, Juli Gogoi, S.B.Nair, Design and Implementation of a Spell Checker for Assamese, Proceedings of the Language Engineering Conference, LEC2002, December 2002, Hyderabad, Published by the IEEE Press.
2. Monisha Das, S.Borgohain, S.B.Nair, Spell checking in MSWord for Assamese, Proceedings of the ITPC-2003:Information

Technology: Prospects and Challenges in the 21st Century Kathmandu, Nepal, May 23-26, 2003.

8. Manika Newsletter

Snapshot of the Manika: e-newsletter in Manipuri



Description

Manika, an e-newsletter in Manipuri, was launched on the Republic day this year (26th January 2003) to mark the beginning of a new electronic information era in Manipur - the Land of Jewels. The newsletter hosted by the RCILTS at Indian Institute of Technology Guwahati will bring news and share knowledge of the heritage and local innovations in Manipur and help enable the Manipuris to be well informed of the electronic age and forge ahead.

9. The Team Members

Prof. Gautam Barua	gb@iitg.ernet.in
Dr. S.B.Nair	sbnair@iitg.ernet.in
Dr. S.V.Rao	svrao@iitg.ernet.in
Dr. P.K.Das	pkdas@iitg.ernet.in
Dr. Dipankar Moral	dipankarmoral@postmark.net
Samir Kr. Borgohain	samir@iitg.ernet.in
Sushil Kr. Deka	sdeka@iitg.ernet.in
Monisha Das	moni@iitg.ernet.in

Nilima Deka	nilima_deka@iitg.ernet.in
Juli Gogoi	juli@iitg.ernet.in
Dr. L.Sarbajit Singh	sarbajit@iitg.ernet.in
Sirajul Chawdhury	si_chow@iitg.ernet.in

*Courtesy : Prof.Gautam Barua
Indian Institute of Technology
Department of Computer Science & Engineering
Panbazar, North Guwahati,
Guwahati -781 031, Assam
(RCILTS for Assamese & Manipuri)
Tel: 00-91-361-2690401, 2690325-28
Extn 2001, 2452088
E-mail : gb@iitg.ernet.in, g_barua@yahoo.com)*