

SPELLChecker

CRBLP Bangla Speller is a Java library to provide Bangla spelling checking services. One of its key strengths is in generating “good” suggestions for misspelled words, for which it uses similarity of pronunciation in addition to other techniques. It can be used as a standalone application to check the spelling of Bangla Unicode documents, or it can be embedded in other applications (e.g. email client, word processor, etc), or it can be used as Web Service.

Project Page:
<http://www.bracu.ac.bd/research/crblp/speller.php>

JKimmo

JKimmo is a GUI for PC-KIMMO – a two-level processor for morphological analysis. PC-KIMMO was developed in the C language and does not support Unicode. JKimmo adds the Unicode support to PC-KIMMO, and creates a multilingual interface for morphological analysis.

Project Page:
<http://www.bracu.ac.bd/research/crblp/jKimmo.php>

Localized URL, Corpus Analysis & Corpus Collection, Lexicon, Parallel Corpus

Project Page:
<http://www.bracu.ac.bd/research/crblp/currentresearch.php>

Join us in making Bangla language computing a reality

CRBLP, BRAC University
66, Mohakhali, Dhaka-1212
Phone: +88 (02) 8824051-4 Ext: 4023
Fax: +88 (02) 8810383
crblp-info@bracu.ac.bd

<http://www.bracu.ac.bd/research/crblp/index.php>

Center for Research on
CRBLP
Bangla Language Processing

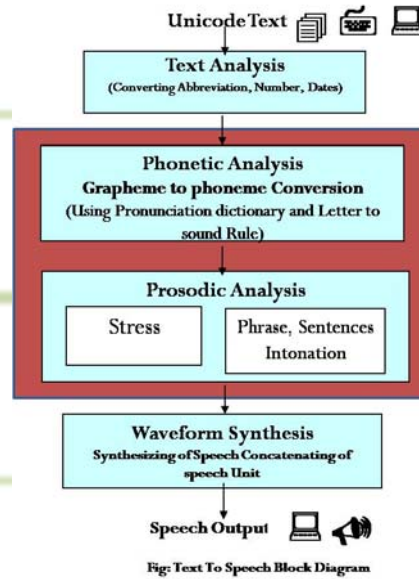
Center for Research on Bangla Language Processing, CRBLP of BRAC University has been working on Bangla Language Processing since 2004. This center is supported in part by grants from the PAN Localization Project (PanL10n) of the International Development Research Corporation (IDRC) of Canada, and Microsoft Corporation of USA. Our research team has been working on Bangla Document Authoring, Information Retrieval (Spelling Checker, Search Engine), Optical Character Recognition, Speech Processing (Speech Synthesis, Speech Recognition), Pronunciation Generator, Morphological Analysis, Parts of Speech Tagging, Syntax, Grammar Checker, Text Categorization, Language Modeling and related research areas. All our software are released under an open source license.

Our Major Research Projects

TTS

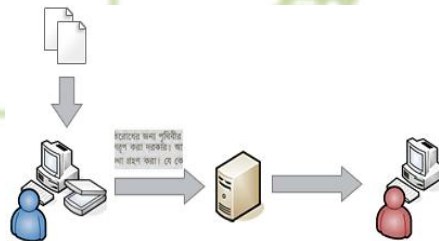
We are working on generating speech from Bangla Text. Text To Speech (TTS) system for the Bangla language can help tackle the illiteracy problem, empower the visually impaired and increase the possibilities of improved human-machine interaction. This project aims to develop a TTS system for Bangla using diphone and unit selection concatenation techniques based on the Festival speech synthesis technology.

Project page:
<http://www.bracu.ac.bd/research/crblp/tts.php>



OCR

Optical Character Recognition (OCR) is the process of converting printed text images to editable text. We are developing a Bangla OCR that takes the scanned image of a printed page or document as input and converts it into editable Unicode text.



Project Page:
<http://www.bracu.ac.bd/research/crblp/ocr.php>

WORDNET

One measure of the “digital maturity” of a natural language is the richness and diversity of linguistic resources available for that language – from the simple digital dictionaries to the complex annotated corpora. Bangla, despite being very widely spoken, is only just beginning to see the development of these linguistic resources. One such important resource is WordNet, a lexical semantic database for a language. It is a linguistic resource that can be used in, and for, a variety of applications ranging from a digital dictionary to an automatic machine translator. We are developing a BanglaWordNet (BWN) for Bangla language.



Project Page:
<http://www.bracu.ac.bd/research/crblp/wordnet.php>

CRBLPConverter

CRBLP converter was designed to address a very common problem with Bangla documents – how to convert documents encoded using non-standard ASCII fonts to Unicode. CRBLPConverter has a set of built-in converters for commonly used ASCII fonts, including the ones used by the Prothom Alo newspaper. The converter can be extended to incorporate other ASCII fonts very easily.

Project Page:
<http://www.bracu.ac.bd/research/crblp/converter.php>