

Design challenges of Spoken Dialogue System in Bodo Language

Dr. Manoj Kr. Deka¹, Aniruddha Deka¹
¹Dept. of Com.Sc & Tech, Bodoland University

Abstract- In the last few years there is a significant development in natural language processing, speech recognition and text-to-speech conversion. In these significant development the research in spoken dialogue is Spoken dialogue can play a vital role between human-computer interactions. Spoken dialogue is the missing component which makes possible to develop conversational interfaces between human and computer. Again the designing a spoken dialogue system with a developing language like Bodo is more challenging. In this paper we give a brief survey on spoken dialogue systems, its challenges and discuss different issues for developing spoken dialogue systems in context of Bodo Language.

Keywords—Spoken dialouge system, importance and characteristics ,Bodo language, Bodo corpus.

I. INTRODUCTION

A. Spoken dialouge System

The interaction between human and computer try to minimize the gap for literate/illiterate and visually challenged users to access information. There are different kinds of ways like speech, text, gestures, symbols etc., or a combination of these for interaction between human and computer. An interaction or conversions consist of statements, expressions, questions and answers. A spoken dialogue system is computer agent that interacts with people by understanding spoken language [1]. The goal of a spoken dialogue system is to provide information by conversing with a human-being in a natural fashion. The main purpose of a spoken dialogue system is to provide an interface between a user and a computer-based application such as a database or expert system. The input in spoken dialogue systems consist of small set of spoken words like yes and no, digits such as 0-9 or combination of both. DTMF is also one kind of input for such systems. The output may be spoken or displayed as text on a screen, and may be accompanied by visual output in the form of tables or images [2].

B. Importance of Spoken dialouge System

IT and IT based applications are essential part of our day to day life. We use computer and computer based application like ATM, paying bills, online purchasing, flight booking etc. for our regular routine works. Most of these applications contain graphical user interface. These applications takes input through keyboard or mouse and finally get the response. But general people who are not habitual of the recent technology are not able to use these kinds of system. These inability put some question to the researcher that

- Could we somehow achieve human like performance and flexibility with computer interfaces?
- Wouldn't it make sense to have systems that could be used by having human-human like conversation, that is, the most natural way for a human to interact with somebody?

One of the solutions of these questions is spoken dialogue system. Spoken dialogue systems try to achieve this conversational gap between the user and the system. There is a substantial improvement in natural language processing, speech recognition and text-to-speech conversion in the past decade. Spoken dialogue is the missing component in between of these, making it possible to develop conversational interfaces to computers. [3]

II. ARCHITECTURE OF SPOKEN DIALOGUE SYSTEM

A. Component of Spoken dialogue system

A common spoken dialogue system is consisting of six components:

- Automatic Speech Recognition component (ASR): This ASR module takes speech as input and returns a string of words. But due to the different pronunciation, background noise, emphasis, pauses common systems have high error-rate.
- Natural Language Understanding component (NLU): It produces a semantic representation of words from the strings using syntactic and semantic analysis. There are several methods which can be used in this analysis namely semantic grammar, probabilistic semantic grammar, semantic hmm, template based semantic etc.

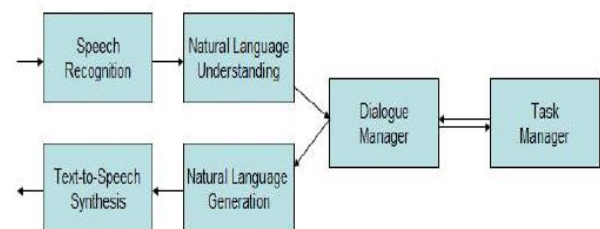


Fig.1: Basic architecture of spoken dialogue system

- Natural Language Generation component (NLG): It generates in each dialogue state an appropriate expression. The generation of task consist of two parts:
 - a) What to say, addressed by the content planner module
 - b) How to say, addressed by the language generation module
- Text-To-Speech synthesis component (TTS): This module converts the generated words into audio.

- Dialogue manager: It takes the output from the NLU module and passes it to the task manager and vice-versa. It also controls the structure of the dialogue.
- Task manager: The Task Manager (TM) assists other components in recognition of user's intention and in execution of problem solving steps with respect to the task at hand. It answers queries about objects and their role. Intention recognition services are usually used by the Interpretation Manager (IM). The TM also provides a generic interface to task-specific agents that do the executing.

B. Characteristics of Spoken dialogue System

Some of the main characteristics for spoken dialogue system are:

- To achieve the goal system must recognize user's intention first. What the user is really trying to do?
- When discussing and exploring the solution usually there is mixed-initiative interaction, i.e., both parties can control the dialogue enabling everybody's needs to be met.
- Interactions are contextually interpreted, depending on the discourse history and the current context.
- Ability to understand user's goal and to reach an appropriate and a satisfied solution.
- Ability to carry out sub-dialogues to achieve sub-goals.
- Ability to pass control from one sub-dialogue to another.
- Ability to vary the dialogue initiative modes from system initiative to user initiative.
- Use of a user model to expect user's utterances and act aptly.
- Directing the user towards task completion.

III. BODO LANGUAGE

Bodo (pronounced as Boron) is a language of North East India spoken by the Bodo tribe of people which is a prominent ethnic group inhabiting Assam, Bhutan, Nepal and Bangla Dash. It falls under the Assam-Burmese group of languages. It branched off from the Tibetan and Burmese family of languages. It is one of the twenty two recognized languages by the Eighth Schedule of the Indian Constitution and is among the official languages of the Indian State of Assam.

Bodo is derived from the term 'Bod & rsquo'; which means Tibet. This signifies that majority of the Bodes arrived from the Bhutan passes. Bodo is a tone language like many of the African and South East Asian languages. The tonal languages use pitch to signify a difference in meaning between words. [4]

Before 1953, the Bodo language had no standard form of writing. It had a history of using Deodhai, Roman and Assamese scripts. At present, Bodos adopted the Devanagari script. But, there is a huge difference in the usage of the letters in Bodo language from the Devanagari script. Bodo language shares some common salient features with other languages belonging to the Bodo group. These features are similar in terms of phonology, morphology, syntax, and vocabulary. Bodo language is closely associated with the Dimasa language

of the state of Assam and with the Garo language of the state of Meghalaya, and also with Kokborok language of Tripura. It important to note that, among the four districts of present Bodo land, namely, Kokrajhar, Chirang, Baksa and Udalguri, the language is heard in pure form only in the district of Udalguri [5]. Bodo has at least four varieties:

- North Goalpara variety, spoken in the northern regions of Goalpara and Kamrup districts. However the North Goalpara and the North Kamrup variety have salient differences.
- A variety spoken in South Goalpara, Garo Hills and South Kamrup.
- A variety spoken in Darrang, Lakhimpur and a few places of Arunachal Pradesh, Called the north central Assam dialect.
- Another variety spoken in the Nowgong, North Cachar and Karbi Hills.

The Bodo language has its written record from the last part of the 19th century. This language was introduced in the primary level of education in Assam from the year 1963 and presently is the medium of instruction up to 10th standard in the state of Assam. It was recognized by the government of Assam as official language in the Kokrajhar district and Udalguri sub-division from the year 1984. The language also got Indian govt. recognition as scheduled language from 2003. According to the census of 1991 it has a total of 11, 84,569 speakers. The Bodo population has basic concentration in the northern part of the Brahmaputra valley of Assam [6]. They have also thin concentration in the southern part of the valley. Besides that they have also concentration in small number in the border areas of Meghalaya, Nagaland, North Bengal, Nepal, and Bhutan adjoining Assam. This language has a total of 22 phonemes: 6 vowels and 16 consonants.

a. Vowels: अ, आ, इ, उ, ए, औ

b. Consonants: ख, ग, ङ, ज, थ, द, न, फ, ब, म,
र, ल, स, ह

c. Semi Vowels: य, व

The Bodo language has different special characteristics such as [7]:

- It has intonation pattern, juncture and two types of tones.
- The words in Bodo are highly monosyllabic.
- It has agglutinative features.

IV. TONAL ISSUES IN BODO LANGUAGE

Bodo is a developing language which has a very less linguistic and literature resource. The vocabulary list of this language is very less and new and new words are being discovered, coined and added. As a result, to collect a Bodo corpus has typical and frequent problems.

Sometimes to represent a particular concept of Hindi, Bodo language requires a complete sentence or long phrase, rather than one single word or combination of two words.

Bodo is a tonal language. Tones refer to the distinctive pitch level of a syllable. In many languages the tone carried by

the word is very essential for the meaning of the word. Such languages are called tonal languages. If the meaning of a word is changed by the pitch of word, that language is called a tone language. The pitch can change nuances as well as core meaning of the language. The different tones are produced by the different pitch level of the language. Eg.

Anjalua _____ bungdungmon
Anjalu _____ said

The acoustic result of the speed of the vibration of the vocal cord in the utterance of the voiced part of the sound produced *pitch*. The rapid vibration of vocal cord produces high-pitched sound and slow vibration produces low-pitch sound. The tones may fluctuate due to pitch contour movement and thus raising and falling tones are produced. All languages have pitch variation. The function of this pitch is different from language to language. In Sino-Tibetan family of languages, the pitch difference distinguishing the meaning of one word from the other though they have the same phonetic structure.

The pitch difference used in this way is called tones. Tone may be on a single level of pitch, called level tone or may fluctuate and thus produce contour type of tones. As a result of the fluctuation, the level of tone may change and produce different categories of tones. If the pitch level rises during the articulation of the sound it is called rising tone. If the pitch level falls, the tone is called falling tone. There may be fluctuation in the middle to produce the tones rising-falling and falling-rising. Based on the pitch movement from the starting position, the tones may also be classified as mid-level, high-level and low-level due to their level-wise movement or they may be mid-rising, mid-falling, high-rising, high-falling, low-rising and low-falling due to their fluctuation from the starting position.

Bodo language has two contrastive tones of contour type – rising, which rises still higher than its original pitch registered at the beginning of the syllable and falling, which falls still lower than its original pitch registered at the beginning of the syllable. Any of the two tones must co-occur with every syllable in the language. [8]

V. CHALLENGES IN BUILDING A BODO CORPUS

Some of the challenges for building a corpus in Bodo corpus are given below.

1. No standard or uniform spelling system is followed by the authors or writers in Bodo language for their writings though standardized language is followed. Most of the authors and writers go their own wishes. Due to this difficulty spelling variation is occurs in their scripts which create problem in corpus design.
2. Splitting of words occurs frequently in Bodo Language. It creates problem while entering the text into desired format.
3. Joined sentence is one of the issues in the Bodo texts.
4. As the Bodo vocabulary is very less and the authors write by their own style hence a large number of incorrect punctuation may occur in the texts materials.

5. Due to the dialect variation sometime many dialect words are found in the texts.
6. Grammatically incorrect sentences can damage the smoothness of the system. So it has to be rectified before deploying the system.
7. Bodo also have hyphenated words, those are in case of multiword expression words. But surprisingly, there are a few hyphenated words in Bodo within a word which are found in the texts.

VI. CONCLUSION

This survey paper mainly comprise of two parts. First Part describes the spoken dialogue system, its importance and characteristics. Second part describes a survey on Bodo languages, tonal issue associated with Bodo language and probable challenges faced during the building of a Bodo corpus. As spoken dialogue system is a recent development issue and there are various researches as well as various developments is going on, so through this paper we try to give a small description to Bodo researchers who are working on designing Bodo language based Spoken dialogue system.

VII. REFERENCES

- [1] Jason D. Williams, Spoken dialogue systems: challenges, and opportunities for research, ASRU – December 2009.
- [2] Michael F. Mctear, Spoken Dialogue Technology: Enabling the Conversational User Interface, University of Ulster.
- [3] Kimmo Karhu, Spoken dialogue systems, Helsinki Unive of Technology.
- [4] <http://www.veethi.com/articles/bodo-article-225.htm>
- [5] Laba Kr. Thakuria, Purnendu Acharjee, Akalpita Das, Prof. P.H. Talukdar, "BODO Speech Recognition based on Hidden Markov Model Toolkit", International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013 2309.
- [6] Dr. Shikhar Kr. Sarma, Moromi Gogoi Biswajit Brahma, Mane Bala Ramchiary, "A Wordnet for Bodo Language: Structure and Development".
- [7] Chandan Sarman, Prof. P.H. Talukdar, "Dialect variation in Boro Language and Grapheme-to-phoneme conversion rules to hand;e exical lookup fails in Boro TTS System," International Journal of Scientific and Research Publications, Volume 2, Issue 9, September 2012
- [8] Utpal Bhattacharjee, "Recognition of the tonal words of BODO language".



Dr. Manoj Kumar Deka, received the M.Sc. degree in Instrumentation in the year 2007 and Ph.D. degree in Signal Processing in the year 2011 from Gauhati University, Assam. He achieved INSPIRE Fellowship sponsored by Department of Science & Technology (DST), Govt. of India, New Delhi during his Ph.D. work. Presently he is working as Assistant Professor in the Department of Computer Science & Technology, Bodoland University, Assam. His area of interest includes Signal Processing, Embedded system, Artificial Intelligence.



Aniruddha Deka, received the B.E degree in Computer Science from North Eastern Hill University from Shillong in the year 2006 and M.Tech degree in IT from Gauhati University, Guwahati in 2013. He worked as an Assistant Project Engineer in IITGuwahati from the year 2007 to 2012, mainly focusing on development of speech based application in Assamese Language. He is currently pursuing the Ph.D. Degree in Computer Science and Technology at Bodoland University, Assam. His research interest includes speech processing, spoken web and spoken dialogue system.