# Research Proposal Paper on Sanskrit Voice Engine: Convert Text-to-Audio in Sanskrit/Hindi

Piyush Mishra[1]
Lovely Professional University
Lovely Professional University, Punjab, India

Jainendra Shukla[2]
Lovely Professional University
Lovely Professional University, Punjab, India

## ABSTRACT

This paper presents, Methodology,Application area& some Results obtained in association to the proposal of an "Automated System". Our proposed system is capable of teaching "Sanskrit Language" with the help of "Hindi Language". System would have two major modules under its consideration i.e. "Teaching" & "Evaluation". The system's main motivation is to utilize the similarities of the two languages &add on to the Sanskrit learning environments in accordance to the "Modern Education Scenario".

## General Terms

Natural Language Processing, Computational Linguistics, Neural Network, Text-to-Speech (TTS), Speech Recognition.

## Keywords

Sanskrit Language, Hindi Language, Teaching, Evaluation, Modern Education Scenario.

## 1. INTRODUCTION

With advancement in the field of Computational Linguistics &the Natural Language Processing there has been a great uplift in the extraction & understanding of the ancient languages, their scripts &knowledge.Modern days are of Technology & Performance, both of which relates to the ease of work & effective use of Time. Enhancement & penetration of computer in different fields made it necessary to be knowledge available & information ready.The proposed system justifies the above statement in "Indian" context,which will be beneficial for providing a bridge type connection to the modern technology & Ancient Indian Sciencealong with the complete relevancy to the Indian origin language "Sanskrit".

The system comprises two distinct modules for two inter-related tasks of any education system i.e.

1. Teaching &
2. Evaluation

## 1.1 Teaching

Implementation of a machine for teaching Sanskrit with support of Hindi, will impart the Text-to-Speech (TTS) for delivering the speech output from the Sanskrit and/or Hindi text of the books. System would be accepting various Sanskrit texts, their Hindi interpretations & read it out using the module comprising speech synthesis. The system needs Text-to-Speech synthesis for Hindi as well as for Sanskrit, both languages are different but requirements for these languages are not. Since both the languages has their base script as "Devnagri", also they are phonetic in nature they do not require separate TTS, which is one of the major beneficial aspect of the system, since the capabilities of the system would not be limited to single language, instead, it might be possible it would support all the languages having Devnagri as base script, for example Hindi, Sanskrit, Marathi,though it is tough to state it with a good level of confidence as the "Teaching" module is completely a proposed concept &yet not many efforts has been made to this part in present research.

The current research is completely focussing on the "Evaluation Module", as teaching is incomplete without any procedural evaluation methodology.

## 1.2 Evaluation

The job performance of the Evaluation module in the system has been categorised in following sub tasks

A. Asking Questions

B. Manipulate learner's Answers

The Evaluation process will be conducted as a speech oriented question/answer session between the tutor (Machine) & the learner (Human), where the tutor will ask the questions& the learner need to answer them. The interesting part of the evaluation process is that it will be in the verbal manner. On the basis of answer the tutor will evaluate the learner about their learning. For implementing the evaluation part to the teaching machine there is a necessity of indulging a speech recognition feature in thesystem so as to get the verbal, answer speech, as input, process it & check the correctness in accordance to the question asked.

In this paper we are presenting the "Asking questions" named sub task of the evaluation system. The system is sufficiently capable of forming questions on its own from the input Sanskrit text;this feature makes the system self-sustained & independent of supervision. The details regarding Methodology & Results in association with "System Formed Questions" are discussed in later sections.

## 2.    LITERATURE REVIEW

In this research there has been an extensive study of various fields of computers; in accordance to the system proposed, there has been none of the work done on the part of facilitating with an intelligent Sanskrit Techno system, we discuss the recent developments in these following areas comprising in relation to the proposed system.

- ➔ Speech Recognition.

- ➔ Text-to-Speech (TTS).

- ➔ Natural Language Processing (NLP).

### 2.1 Speech Recognition

### [5] [8] [9]

This is the field with at mostimportance in concern to any Human-Machine Interaction System (HMIS), such as that of the system proposed. Mainly the HMIS are developed of two major components

1.    Signal Processing Component.
2.    Language Processing Component.

Signal Processing Component for making the system aware about the nature of the input signal, & Language Processing Component to associate the input signal in accordance to the words in the Language so as to generate proper & accurate responses.

There is required a rigorous work to be done in developing the system for the analysis of any speech signals of any particular language. Correct interpretations of the signals are necessarily major area of consideration in the case of development of such sophisticated systems, wrong manipulations would result something meaning-less & hence useless.

In such systems as soon as the input signals are arrived there must be the capabilities of inter-relating the speech signal with the trained data or pre-defined rules & perform correct language processing.

At the application level of such systems, these are complicated & difficult to embed with other applications since they work at the back-end & their outputs are critical for the overall output of the system, because if there has been gone anything wrong at the speech recognition level, the rest of the manipulations are never going to be correctly manipulated.

## 2.2    Text-to-Speech [1] [7] [12]

TTS are mainly the Speech Synthesis systems comprising the abilities to produce natural alike speech outputs from the text documentation. These systems can be used so as to make read the text, importantly such systems are unique of their kind & are very beneficial to the visually challenged people to connect themselves with the community & go parallel with the mainstream, even more not only these systems are good to a single particular population of our society but can be beneficial in other fields such as teaching & training purposes, may be to the children or also to the illiterate, after all education is for all. Similar to Speech Recognition systems TTS systems consists of two modules

1.    Textual Analysis" &

2.    Speech output Generation".

## 2.3 Natural Language Processing [2] [3] [6] [11]

NLP is the field of computers which deals to enhance the capabilities of system with the knowledge of Human's Languages so that to make compatibility with the user in various distinct modes of interactions. Its knowledge also comprises the analysis of different aspects related to the languages, other than the Grammar & Word meanings, such as Ontologies, Morphological analysis, Word formations, Question Formations, detailed study of Lexicon, development of Computational Corpora and many more.

NLP deals with Linguistics or in other words Computational Linguistics as its part, & both the fields run parallel to each other complementing the other's development.

## 3.    METHODOLOGY

The text provided as input for the system is developed in UTF-8 format using the "Aparajita" font, and then the text is tagged using a Part-of-Speech (POS) tagger. The POS developed by Jawaharlal Nehru University, New Delhi as a Language Processing Tool is used for our research purpose which is named as "Sanskrit Consortium POS". The tagger is developed in coordination with Microsoft Research Labs India. Also both the institution has developed their individual "Tagsets" & their "Annotations" along with their guidelines,available at their website. Then before using the text as input for the system there are somepre-processing required to be done on the text so as to make it appropriate in accordance for the system.

The formations of the question are done on the basis of analysis & manipulations of various defined rules & patterns. These patterns are extracted from distinct texts & then after the rules are defined. Up till date in our research we have formulated58 patterns of question formations from 33 tagged sentences of different Sanskrit documentations. Sanskrit has many Question forming units, but the research is confined to "8" of them. The elementary building blocks for question formations are

कः *, * कः *, किं *, * किम् *, कुत्र *, * कुत्र *, * किमर्थं *, * कानि*.

काशीनगरे\N एकः\SMK पण्डितः\NP अस्ति\NP ।\PUNC

Out of 58 patterns, 47 patterns are identified to be unique& 11 are redundant (i.e. approx. 20%), the redundant patterns are not being in the consideration for the research purpose & were omitted.

The methodology can be understood more easily through a simple example as follows

Example:

⇨ Sanskrit sentence in UTF-8 format using "Aparajita" font

काशीनगरेएकःपण्डितःअस्ति।

⇨ Same sentence tagged using the POS-tagger, we get

⇨ Sentence after pre-processing

काशीनगरे N एकः SMK पण्डितः NP अस्ति NP

The pre-processing of the sentence has been done, so the text is now appropriate to be considered as the input for the machine & now can be passed to the system for producing the self-formulated questions on the basis of the knowledge of patterns & rules that are governing the overall computation intelligently. After passing the input to the system, it will produce the output as single "Notepad" file naming "possible_questions.txt", displaying all the possible question formations from the sentence, & will also display the relevantly complete tagged forms of the questions with the help of an interface; this is explained & represented in Figure 1.
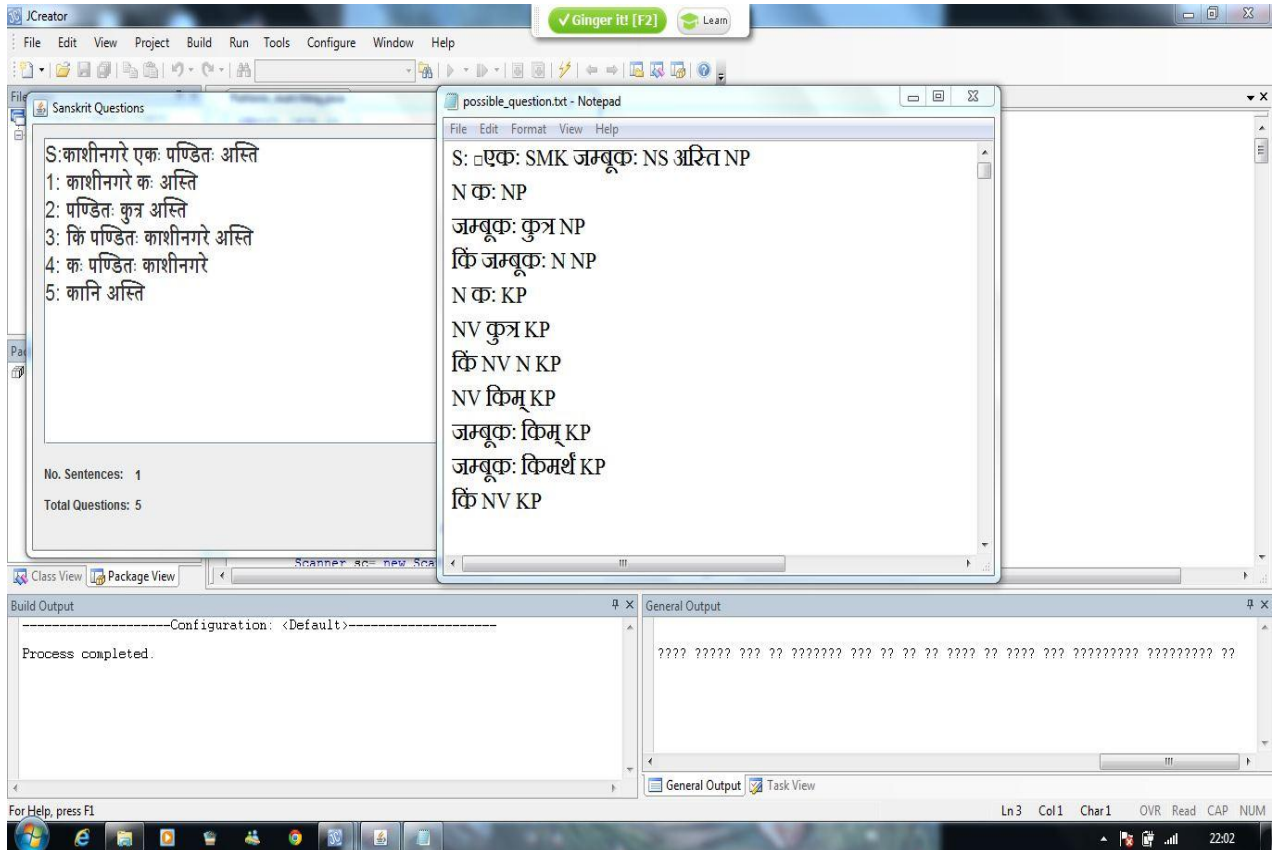


**Figure 1**

From the above example & pictorial representation the working & the functionality is clearly explained, we may now discuss the actual experimental results obtained by providing the system with an unseen Sanskrit text input. These are mentioned in further sections.

## 4.     APPLICATION AREA

As we have already discussed about the diversity of the fields involved in the system, accordingly its application areas are very broad. Our proposed system would provide a leap to the research areas including Sanskrit language, its grammar & its

computational relations. Scientific & research activities in regards to the precious ancient Sanskrit Knowledge will be boosted. System is being proposed with keeping in mind the requirements of an "Intelligent Tutor", which contributes to the teaching & learning scenario for Sanskrit language. Since the system would be a speech oriented machine it makes the teaching-learning process very easy & simple, though the system would be the complex one. It comprises the various fields of "Artificial Intelligence" for education, which changes & modernizes the traditional methods of Sanskrit learning or in other words Language learning.

One of the most significant benefits of the "Sanskrit Tutor System" would be that, there is no requirement of Human intervention or supervision for the purpose of "Teaching" & "Evaluation", to make it easy to understand we just state that "To make a well Qualified Sanskrit Faculty there is a requirement of at least 30-35 years of grooming the being, which is a very slow process w.r.t. the evolution of "Intelligent Systems". Once the system gets evolved, they can be propagated & penetrated to the society & education system very easily & effectively without much capital involvement, which further adds to the benefits.

# 5.   EXPERIMENTAL RESULTS

Following are the results obtained in regards to the tagged document having description as given below:

## Document Features

**Total No. of Words:** 40

**Total No. of Sentences:** 15

**Possible No. of Questions:**

47 (No. of Patterns) * 15 =702

**Completely Tagged Questions** [Actual Retrievals]**:** 101

**Correctly formed Questions:** 74

**Percentage of Correctness:**

(74 * 100) / 101 = 73.26 %

Chart-1, shows the No. of Questions formed by the system from the sentences, with following parameters

X- Co-ordinate = Sentence No.

Y- Co-ordinate = No. of formed Questions

### Chart-1
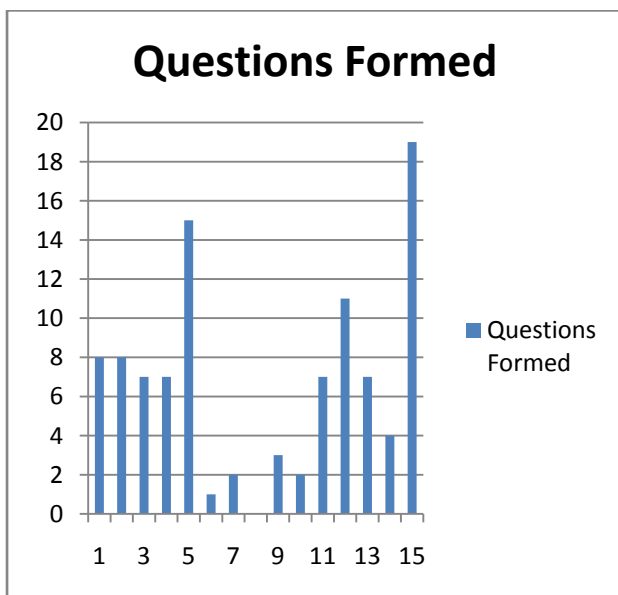


Chart-2, shows the No. of Correct Questions formed by the system from the sentences, with following parameters

X- Co-ordinate = Sentence No.

Y- Co-ordinate = No. of Correctly formed Questions
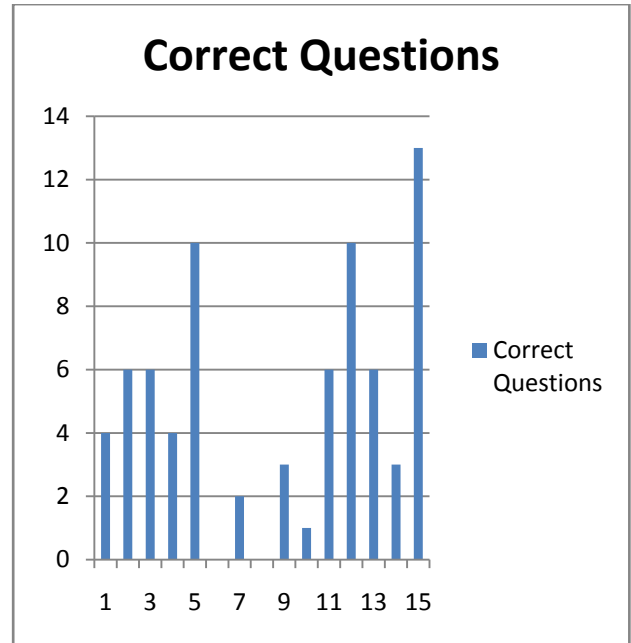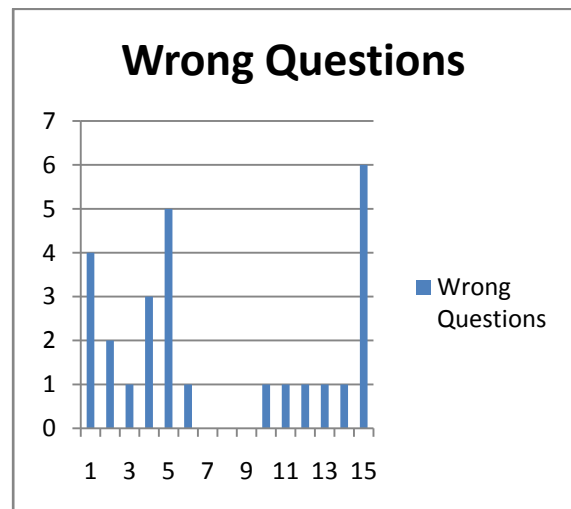
### Chart-2



Chart-3, shows the No. of Wrong Questions formed by the system from the sentences, with following parameters

X- Co-ordinate = Sentence No.

Y- Co-ordinate = No. of Wrongly formed Questions

### Chart-3



Data plotted in all the three Charts are collectively provided in the Table-1 for the purpose of reference, calculations and data analysis.

**Table 1**

| Sentence No. | Question Formations (Actual Retrievals) | Correct Question Formation | Wrong Formations |
|---|---|---|---|
| 1 | 8 | 4 | 4 |
| 2 | 8 | 6 | 2 |
| 3 | 7 | 6 | 3 |
| 4 | 7 | 4 | 1 |
| 5 | 15 | 10 | 5 |
| 6 | 1 | 0 | 1 |
| 7 | 2 | 2 | 0 |
| 8 | 0 | 0 | 0 |
| 9 | 3 | 3 | 0 |
| 10 | 2 | 1 | 1 |
| 11 | 7 | 6 | 1 |
| 12 | 11 | 10 | 1 |
| 13 | 7 | 6 | 1 |
| 14 | 4 | 3 | 1 |
| 15 | 19 | 13 | 6 |
| **Total =** | **101** | **74** | **27** |

## 6. FUTURE SCOPE

In this paper we presented the Methodology, Application Area, & Results related to the self-formed questions, for the "Evaluation" module, with the sub-task named "Asking Questions"

In future as the extension of this research we will work on extraction of more generalised & broad pattern formations to enhance the Question forming system's efficiency from the present work. Also we will perform research on the other components, modules, tasks, & sub-tasks of the proposed system to validate the complete "Intelligent Sanskrit Tutor".

## 7. CONCLUSION

The Research Methodology & Results discussed through the charts, snapshots & tables hereby conclude the total efficiency of our system which accepts Sanskrit Text as input & forms Questions as output with a "Precision" of 73.26 %

**Precision =**

**[(Correct Questions) / (Actual Retrievals)] * 100**

= [(74) / (101)] * 100

= 73.26.

## 8. REFERENCES

[1] Ong Singh goh, 2008, "A Framework & Evaluation of Conversation Agents", PhD thesis, School of Information Technology, Murdoch University.

[2] Lei Li,YanquanZhou,LinZang,Yxin,Zhong, 2006, "Enhancing Speech Based Information Service with Natural Language Processing",0-7803-9701-0/06, IEEE.

[3] F.Siasardjahantighi, M.Norouzifard,S.H.Davarpanah, M.H.Shenassa, 2008,"Using Natural Language Processing in Order to Create SQL Queries",978-1-4244-1692-9/08, Kuala Lumpur, Malaysia

[4] MallikarjunHangarge,B.V.Dhandra, 2008,"Shape and Morphological Transformation based Features for Language Identification in Indian Document Images",978-0-7695-3267-7/08,2008.

[5] Kiruthiga S &Krishnamoorthy K,2012,"Design Issues in Developing Speech Corpus for Indian Languages – A survey",978-1-4577-1583-9/12,IEEE, Coimbatore,India

[6] Lou Guohuan, Zhang Hao, Wang Honghui,2009, "Research on Methods of Semantic Disambiguation about Natural LanguageProcessing",978-0-7695-3901-0/09,IEEE.

[7] T.Kalaiyarasi,RParthasarathi,T.V.Geetha,2003, "Poongkuzhali–An Intelligent Tamil Chatterpot",Tamil Internet,Chennai,Tamilnadu,India.

[8] K.K. Paliwal,1990,"Lexicon-Building Methods For An Acoustic Sub-Word Based Speech Recognizer", CH2847-2/90/0000-0729,IEEE.

[9] TrymHolter and Torbjgrn Svendsen,1996,"A Comparison of Lexicon-Building Methods for Subword-Based SpeechRecognisers",0-7803-3679-8/96,IEEE.

[10] Rita Singh,Bhiksha Raj and Richard M. Stern,2002,"Automatic Generation of Subword Units forSpeech Recognition Systems", 1063-6676/02,IEEE.

[11] VipulArora, PawanGupta, Laxmidhar Behera,2007,"Analysis Of Sanskrit Text:Parsing and SemanticRelations",France.

[12] Harsh Jain,VarunKanade,Kartik Desikan,2004,

"Vani-An Indian Language Text To Speech Synthesizer",Final Stage Report,Vani-Project.

[13] HyancithS.Nwana, 1990, "Intelligent Tutoring Systems: An Overview", Artificial Intelligence Review.