Vocabulary Enhancer

Rishabh Chanana , Sanya Taneja, Ayushi Dang, Anand Prem Rajan

Abstract- This study deals with creating a smart bot for improving and enhancing the English of an individual. Keeping in mind the importance of the English Language as a medium of expression across the globe and the lack of Indians to speak English proficiently such technologies have to be made and brought into the Indian Market. The model that we have proposed will basically help people of all ages to learn and improve their English-speaking Skills. The smart bot will further be adaptive with the person who is using it to improve his/her ability to speak English.

Keywords- SmartBot, enhancement of English, Vocabulary Enhancer, Vocabulary builder

I. INTRODUCTION

In the 21st Century the importance of English as a language and a medium of expression has increased rapidly. English is the most-used language online, with nearly 1 billion users typing and chatting in the language. Yet the inability of a majority of Indians to speak English fluently is one of the biggest hurdles that India faces today.[1]

But let's consider the fact that India is a country of 1.34 Billion people, second only to China. This gives our country a population density of about ~400 people per square km. The statistics on English speaking ability tends to be unreliable for a various of political reasons, , however it's usually accepted that somewhere in the range of 30% are able, to varying degrees, speak English—though only a third have some semblance of reading and writing aptitude.[2]

Now, in terms of the motivation for this paper we wanted to offer a solution to the inability of the people of our country to be able to speak and write English confidently. Basically, a solution that can act as an enzyme to make learning English more interesting, easily available, affordable and for all the age groups of people in our country.

English is one the most internationally recognized language in the world. English is the language of business, of international communication, gives more access to the internet and so knowing English empowers an individual. The idea is to create a smart bot for improving and enhancing the English of an individual and being able to present this application to all the sections of people in our country so that they can benefit from it.

Revised Manuscript Received on October 05, 2019

***** Correspondence Author

Rishabh Chanana, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India. Email: rishabhchanana8@gmail.com

Sanya Taneja, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India. Email: sanyataneja32@gmail.com

Ayushi Dang, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India. Email: ayushidang3414@gmail.com

Anand Prem Rajan, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, India. Email: janandpremrajan@vit.ac.in

II. LITERATURE SURVEY

One of the current systems which was proposed in the year 2018 highlighted that the need for enhancing vocabulary development, mobile application systems (apps) was considered as the one of the best ways in the field of vocabulary learning. The purpose of the study is to design and implement a mobile-based application aiming to enhance English vocabulary teaching and learning. Using the Design-Based Research methodology, this study suggests the steps taken to develop a vocabulary learning based mobile app named VocUp. [3]

Another work in the same category which was published in the year 2015 focuses on how mobile devices can be used to facilitate vocabulary learning for English learners. A theoretical approach coupled with studies in vocabulary acquisition was highlighted to point to related practices for all teachers and students. Based on the urged theories AN analysis was drawn on however numerous mobile apps to illustrate vocabulary acquisition with four research-based vocabulary learning strategies: dictionary use, phonological analysis, morphological analysis, and contextual analysis.[4]

This paper studies intermediate-level English learners' performance before and after introducing mobile applications to the study group as the part of the study. It examines if multimedia system courseware affects the vocabulary learning within the second language acquisition.[5]

The study examined the impact of a mobile vocabulary learning app on secondary school students' vocabulary performance, also giving some indications about the students' development of foreign language learner autonomy. [6]

In this paper we describe Indian English and discuss how English is and has been used in the Indian subcontinent. This paper discusses typical examples of pronunciation, vocabulary and grammar of Indian English. This paper also states that including the number of imperfect speakers of English who can manage to communicate through English, the total number of the speakers of English may reach 300-400 million. [7] This article provides a survey of the developments that have taken place within the description of Indian English within the past 2 centuries, with specific attention to the phenomena of language (e.g. phonology, lexicogrammar, and pragmatics) that are examined from a descriptive perspective. The criticism during this article, however, demonstrates that the linguistic descriptions except those in regard to society area unit scant and also the often-made intuitive observation that Indian English is extensively studied doesn't apply to the outline of linguistic phenomena.

[8] In this paper, calibration was introduced to improve English vocabulary learning for learners to reduce the number of repetitions and to improve vocabulary memorization. Thus, an App for the iPhone 4 called English Vocabulary Learning was designed for learning English vocabulary. The findings from this study implied that a practice scheme of calibration can be implemented in learning English vocabulary or in learning any other languages.[9]

Feedback is often taken in word learning paradigms, in both research studies and commercial language learning apps. providing critical (right/wrong) feedback on a spoken response influences two different parts of vocabulary learning, the training of a new phonological kind, and also the learning of a semantic property of the phonological kind. We discover that receiving critical feedback improves retention of descriptive linguistics forms, however not of linguistics facts.[10]

This is significantly true in language learning, wherever new fascinating word games become a mean to enlarge vocabulary and improve the lexical accuracy of the user. Marco Rocetti and seven other authors in 2016 created a game to boost English of a personal. This work presents Magic Word, a game based on wide identified game mechanisms, however specifically designed and developed so as to support beginners learning Italian through an A1 level e-learning course. The paper presents the look and implementation of a demo version of Magic Word, through the most challenges encountered throughout its creation.[11]

III. PROPOSED WORK

This project aims to enhance vocabulary by giving its meaning for the word specified, giving the word of the day, giving a history of past searched words. We are providing a Dictionary interface, a Count which increases on the basis of the number of searches for every word.

We are implementing Speech to Text(S-T) and Text to Speech(T-S), using Speech recognition in Python and Google Text to Speech respectively. We can take input of sentences which then removes stop words and gives meaning.

All the data is stored in an excel file and further used for interpolation and usage.

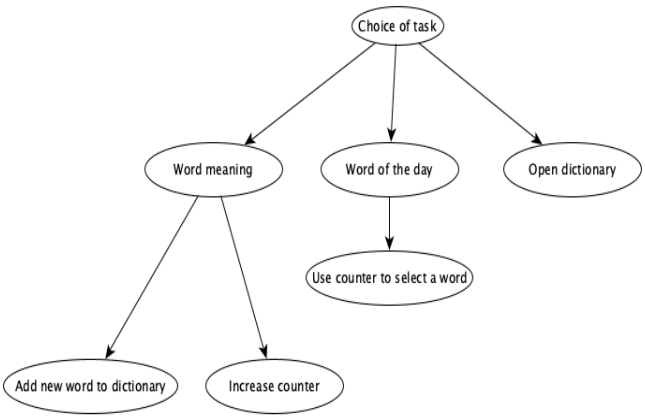

Fig. 1.Flowchart of components

IV. PREPROCESSING

Hardware Requirements-Any device which has python installed on it.

Software Requirements- We need various libraries that help us build our program such as:

- Anaconda
- Jupyter Notebook
- GTTS
- Speech recognition
- Nltk
- tkinder

To start with we firstly install Speech recognition in Python, Google Text to speech service and Play Sound (Library) incorporated. To convert Text to Speech we use the Google Text to Speech Service which creates an mp3 file.

For implementing the task of Word meaning – We use NLTK, Wordnet interface which gives a database for all the major words used in English language and those are being used to fetch the definition of the required words.

The deciding function plays a vital role in our program as it checks if the word is already fetched by the program or not. Two possibilities:

Word existing – Counter increases for that word

Word not existing – A new counter created

Word of the Day – This uses two features to predict the specific Word of the Day which are, the Count, the number of times the word has been searched, the level of the word, the difficulty of a specific word which can be used defined or fetched from the internet, provided a database is given.

Open Dictionary – We are using Tkinter to give an overview/glossary of all the searched words. It uses the csv cells that are further sent to a Tkinter grid which are then displayed.

V. IMPLEMENTATION

Whole process is divided into various functions for specific tasks-

- Initialization – First we install Speech recognition in Python, Google Text to speech service and Play Sound (Library) incorporated
- Text to Speech – We use Google text to Speech service, Call the API, select any desired language which sends back an MP3 file that is stored in the Local library, which is played by the PlaySound library in Python
- Speech to Text – With microphone as source, the audio is recognized using Speech recognition Python library
- Word meaning – We use NLTK, Wordnet interface which gives a database for all the major words used in English language and those are being used to fetch the definition of the required words.
- Stop Words – When a sentence is sent, it is tokenized, broken down into words, the program further removes stop words of English language such as like, is, the, etc and a new array is returned to the calling function.
- Deciding function – This function plays a vital role in our program as it checks if the word is already fetched by the program or not. Two possibilities:

  - Word existing – Counter increases for that word
  - Word not existing – A new counter created
- Count increment/Add to Dictionary – To maintain the read/write concurrency of the data, we have two csv files which are constantly being synced to each other after every read/write operation. Counter plays an important role as it specifies the number of times the word has been searched. This improves our data as it tells us that the specific word will be used for Word of The Day.
- Word of the Day – This uses two features to predict the specific Word of the Day which are, the Count, the number of times the word has been searched, the level of the word, the difficulty of a specific word which can be used defined or fetched from the internet, provided a database is given.
- Open Dictionary – We are using Tkinter to give an overview/glossary of all the searched words. It uses the csv cells that are further sent to a Tkinter grid which are then displayed.
- These all functions are then used to make a flow of tasks. The program will ask us for what we want to do, which can be either looking for a word, a word of the day recommendation or opening up a list of previously used words. When we look for a word, it further checks if the input is a sentence, which is broken down to words. It also checks if the word is a previously searched word or a new word. Word of the day and open dictionary works as per the function usage.

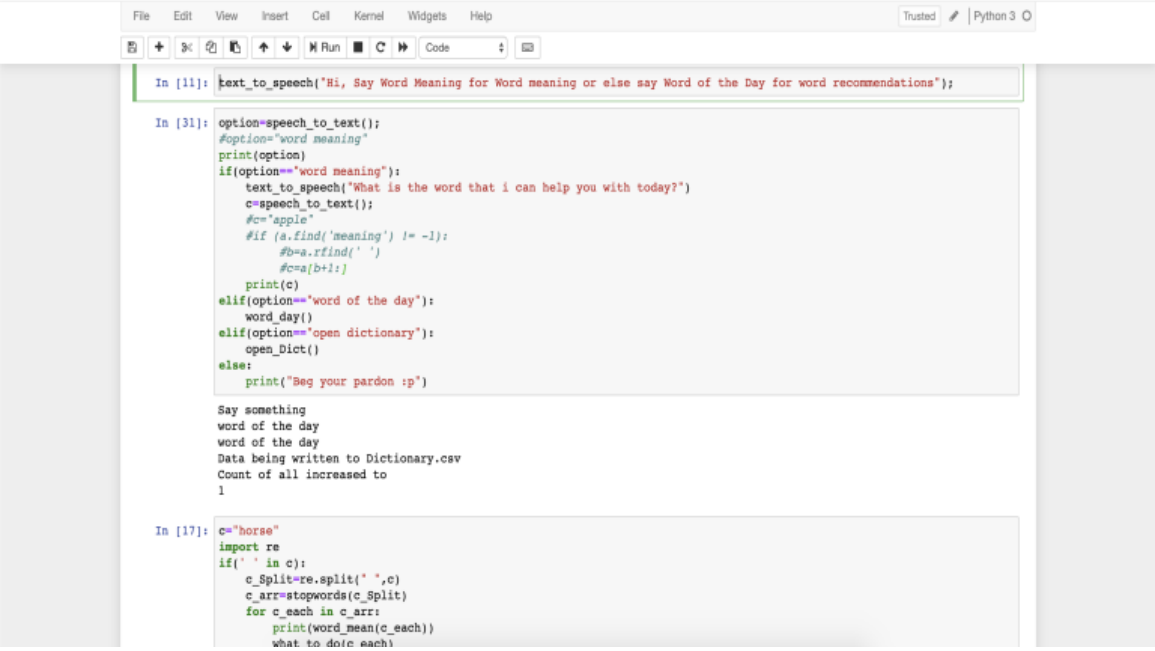

Fig. 2.Implementation Example

VI. CONCLUSION

This program can be implemented in various languages, with the help of a translation API, this can be used to convert any language to English and hence fetches it's meaning in both languages. Also, we can implement this further to make a Mobile App and a web interface and hence making it easily accessible. This will help us increase the reach of our program to the specified audience. We can also add image fetch for the specific words, and hence giving a visual representation of the specific word, and hence making for children to understand it easily. Hence, with some more additional features, we'll able to target all age groups and people from all descents.

REFERENCES

1. https://www.fluentu.com/blog/english/importance-of-english-language/
2. https://www.forbes.com/sites/realspin/2014/11/06/the-problem-with-the-english-language-in-india/#49e3dd21403e
3. https://www.irrodl.org/index.php/irrodl/article/view/3746/4721
4. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1033&context=nebeducator
5. https://www.researchgate.net/publication/261246911_Effectiveness_of_Using_English_Vocabulary_Mobile_Applications_on_ESL's_Learning_Performance
6. https://sisaljournal.org/archives/sep18/ludwig/
7. https://web.uri.edu/iaics/files/05-Tetsuya-ENOKIZONO.pdf
8. Kashyap, Abhishek, "Developments in the linguistic description of Indian English: State of the art. Linguistics and the Human Sciences",2014.

9. Jon-Chao Hong, Ming-Yueh Hwang, Kai-Hsin Tai, Yi-Ling Chen , “Using calibration to enhance students' self-confidence in English vocabulary learning relevant to their judgment of over-confidence and predicted by smartphone self-efficacy and English learning anxiety”, November 2013.
10. Saloni Krishnan, Elise Sellars, Helena Wood, Dorothy V. M. Bishop and Kate E. Watkins , “The influence of evaluative right/wrong feedback on phonological and semantic processes in word learning.”, September 2018.
11. Marco Roccetti, Paola Salomoni, Mathieu Loiseau, Monica Masperi, Virginie Zampa, Andrea Ceccherelli, Cristiana Cervini, Antonella Valva “On the design of a word game to enhance Italian language learning”.

AUTHORS PROFILE

Rishabh Chanana is currently pursuing his B.Tech in computer from Vellore Institute of Technology, Vellore. His research areas include Artificial Intelligence, Machine Learning, Deep Learning, AI in robotics. He is also good at developing android apps and solving algorithmic problems through code. He has done various internships and also attended training programs at NUS, Singapore in the field of ANN.

Sanya Taneja is currently pursuing B.Tech in Computer Science from Vellore Institute of Technology, Vellore. She has worked on : Spot-A-Spot- Efficient car parking system which is an Artificial Intelligence project and is to be published in Springer journal. She has done various technical internships and worked on projects involving latest technologies like deep learning, natural language processing, blockchain.

Ayushi Dang is currently pursuing B.Tech in Computer Science from Vellore Institute of Technology, Vellore. She has done various technical internships and worked on projects involving latest technologies like deep learning, natural language processing, blockchain. Her research areas include Artificial Intelligence, Machine Learning, Deep Learning.

Dr. Anand Prem Rajan is a senior assistant faculty at VIT vellore. He is also a Faculty coordinator for Students, Project Officer - Algal Research, CO2 & Green Technology Center, Anti Ragging Coordinator SBST and also Exam Coordinator for M.Tech SBST. He has over 30 publications.