Computing, Information Systems, Development Informatics & Allied Research Journal
Vol. 8 No. 1, March, 2017 - www.cisdijournal.net

Computational Analysis of Ondo Accents in the Pronunciation of English Words

Enikuomehin, A.O
Department of Computer Science
Lagos State University
Ojoo, Lagos State, Nigeria
toyinenikuomehin@gmail.com

Ameen A.O. & Sadiku J.S.
Department of Computer Science
University of Ilorin
Ilorin, Nigeria

ABSTRACT

A fundamental challenge with current research on speech science and technology is understanding and modeling individual variation in accents and dialects. It was observed that there are some English words often pronounced with strong accents by Ondo, Akure and Idanre speakers of English which are difficult for machine understanding. The aim of this research in man-to-machine communication in English Language by Ondo State Yoruba English speakers, is to analyze different accents of Ondo State English speakers from the three dominant tribes specifically Akure, Idanre and Ondo in other to highlight the differences in their speech parameters that could influence or affect Automatic Speech Recognition (ASR) system. The methodology used was Hidden Markov Model. A robust word recognizer was built which recognized key words that are difficult to pronounce for each of the three sampled tribes using computational analysis. The output of the system is a hypothesis transcription of the utterances. All the wrongly pronounced words were mapped with their respective correct pronunciation and this mapping was used as an input to the "Learning" module of the software programs. This would allow the program to replace the voice input matching those wrongly pronounced words with their corresponding correct spelling in the text output, thereby improving the efficiency and performance of the computer program in use. There are a big number of different approaches for the implementation of an Automatic Speech Recognition but for this study, the four major processing steps as suggested by Hidden Markov ToolKit were considered. The result forms a basic baseline for improved development of voice retrieval systems, voice to text models among others. Future research can be directed to more extensive language models.

Keyword: Automatic Speech Recognition, Received Pronunciation, Speech Signal, Task Grammar, Dictionary.

CISDI Journal Reference Format
Enikuomehin, A.O, Ameen, A.O. & Sadiju, J.S. (2017): Computational Analysis of Ondo Accents in the Pronunciation of English Words Computing, Information Systems, Development Informatics & Allied Research Journal. Vol 8 No 1. Pp 1-8.
Available online at www.cisdijournal.net

---

1. INTRODUCTION

During the last few years the problem of the standard of spoken English in Commonwealth Nigeria has received considerable attention. Because of the multiplicity of Nigerian languages, estimated at almost a thousand (Rice, 1961), English has become the lingua franca over large parts of the continent spoken as a mother tongue by only a tiny minority of Nigerians, English is now used as a second language by millions. For this reason, it is important that the type of English used should keep within certain norms, if speakers are to be mutually intelligible. If English should become so distorted - and there is some evidence that this is already happening (Stephen, 2010) - as to become incomprehensible both within and outside Nigeria, one of the main-purposes for which it has been learnt will have been thwarted.

In contact situation, natural languages behave in a manner quite unpredictable. (Ajani, 2007), for instance, remarks that when two or more languages and cultures come into contact, "different types of sociolinguistic chemistry take place": diaglossia, language shift, attrition, code-switching, creolization, pidginization, birth of a new language or even linguicide (death of an existing language). This situation has led some proponents of contact linguistics (Weinreich, 1968) to propose a phenomenal theory that explains the concept of language variation and change in language contact situations (Rofiq, 2016).). This idea also reflects the Sapir-Whorfian Hypothesis – that the culture of a people is bound to influence, in a major way the nature of language existing in the speech community (Chu 2016)). It is in the light of this that this study has decided to find out how much of the pronunciation problems exist among three sample dialects in Ondo tribe of Ondo state, Nigeria (Ondo, Akure, and Idanre) with respect to how this affects the performance of computer speech.

2. STATE OF AUTOMATIC SPEECH RECOGNITION TECHNOLOGY AND ITS DESIGN IMPLICATIONS

The design of user interfaces for speech-based applications is dominated by the underlying Automatic Speech Recognition technology. More often than not, design decisions are based more on the kind of recognition the technology can support rather than on the best dialogue for the user (Mane, 2006). The type of design will depend, broadly, on the answer to this question: What type of speech input can the system handle, and when can it handle it? Word spotting and the ability to support more complex grammars opens up additional flexibility in the design, but can make the design more difficult by allowing a more diverse set of responses from the user. Some current systems allow a limited form of natural language input, but only within a very specific domain at any particular point in the interaction(Wong, 2015). Even in these cases, the prompts must constrain the natural language within acceptable bounds. No system allow unconstrained natural language interaction, and it is important to note that most human-human transactions over the phone do not permit unconstrained natural language either. Typically, a customer service representative will structure the conversation by asking a series of questions. According to Mane et al., 2006, with"barge-in" (also called "cut-through"), a caller can interrupt prompts and the system will still be able to process the speech, although recognition performance will generally be lower. This obviously has a dramatic influence on the prompt design, because when barge-in is available it is possible to write longer more informative prompts and let experienced users barge-in (Bloem, 2016).

3. METHODOLOGY

The goal of this study was to build a robust whole word recognizer which will recognize key words that are difficult to pronounce for each of the three dialects in Ondo State using computational analysis. That means it should be able to generalize both from speaker specific properties and its training should be more than just instance based learning. The paradigm used is HMM (Hidden Markov Models). To reduce the difficulties of the task, a very limited language model was used. Future research can be directed to more extensive language models. In Automatic Speech Recognition systems, acoustic information is sampled as a signal suitable for processing by computers and fed into a recognition process.

3.1 Data Preparation

The first stage of any recognizer development project is data preparation. Speech data is needed both for training and for testing. The training data is used during the development of the system. Test data provides the reference transcriptions against which the recognizer's performance can be measured and a convenient way to create them is to use the task grammar as a random generator. In the case of the training data, the prompt scripts will be used in conjunction with a pronunciation dictionary to provide the initial phone level transcriptions needed to start the HMM training paradigm. Before the data can be recorded, a phone set must be defined, a dictionary must be constructed to cover both training and testing and a task grammar must be defined.

3.2 The Task Grammar

The task grammar defines constraints on what the recognizer can expect as input. As the system built provides a voice operated interface for microphone input, it handles character strings.
The compiled grammar was:

```
#
#Task grammar
#
$digit=f|v|l|n|s|sh|c|f|v|p;
(SENT-START [$digit] SENT-END)
```

The above grammar can be depicted as a network as shown below:

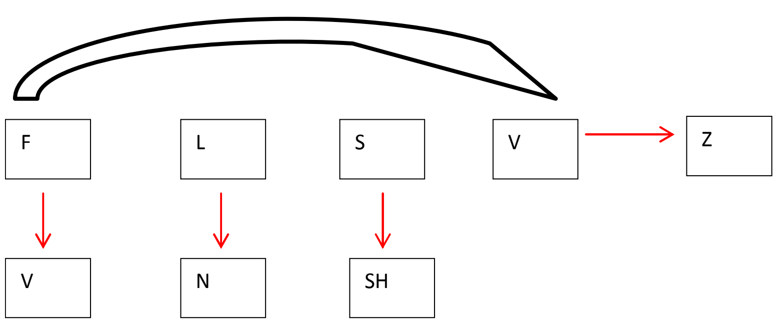

Fig. 1: Task Grammar dictionary network diagram

The above high-level representation of a task grammar is provided for user convenience. The recognizer actually requires a word network to be defined using a low level notation called HMM Tool Kit Standard Lattice Format (SLF) in which each word instance and each word-to-word transition is listed explicitly. This word network can be created automatically from the grammar above using the HParse tool, thus assuming that the file grammar contains the above grammar.

3.3 Recording

In order to train and test the recognizer on the domain and on the voice of some selected people, a paragraph was automatically generated from the grammar with HMM Tool Kit's. Speech data of Eighteen (18) different English-Language speakers from Ondo State (6 Ondo, 6 Akure and 6 Idanre) of different age groups was recorded. Due to researcher's lack of access to a recording studio, the recordings were done in an office on a Sunday when there were no people in the office. As the toolkit does not require phoneme duration information for the training sentences(Chu, 2016), the (differences in) timing in the pronunciation of the training sentences is not important. The toolkit learns to recognize the words through fitting the word transcriptions on the training set. These transcriptions were used for all realizations of the same sentence, even though there might be variation between speakers relative to the transcription. The speakers were given a list with sentences which they had to read aloud. After about 5 sentences they took a short break, and drank a glass of water. The training corpus consisting of paragraph were recorded and labeled using the HTK tool HSLab.

After recording and labeling the training sentences, a test corpus was also created the same way as the training corpus. The difference noted in pronunciation between speakers (and their consequences) can be categorized as articulation variation. For example, some Idanre speakers replaced 'L' with 'N', some Akure speakers replaced 'S' with 'X', and some Ondo speakers replaced 'S' with 'SH'. Phonetic change degrades the quality of the training set, since the same phonetic transcription was used for all speakers (Fudge, 2015). These phonetic changes problems were solved by using isolated whole word models and having many different sentences such that at the end of the day, a speaker independent system was created (Franklin et al. 2016). Articulation variation on the other hand is of course a problem for recognition but if there was no articulation variation the task of recognizing would become an instance based learning problem(Gonçalves et al. 2015).

3.4 Encoding the Data

The speech recognition tools cannot process directly on speech waveforms. These have to be represented in a more compact and efficient way. This step is called" acoustical analysis": The signal is segmented in successive frames (whose length is chosen between 20ms and 40ms, typically), overlapping with each other. Each frame is multiplied by a windowing function for example, hamming function(Munro et al., 2015). A vector of acoustical coefficients (giving a compact representation of the spectral properties of the frame) is extracted from each windowed frame. In order to specify the nature of the audio data (format, sample rate, and so on) and feature extraction parameters (type of feature, window length, pre-emphasis, and so on.), a configuration file (config.txt) was created as follows:

```
#coding parameters SOURCEKIND =
waveform SOURCEFORMAT = HTK
SOURCERATE = 625 TARGETKIND =
MFCC 0 D A TARGETRATE = 100000.0
SAVECOMPRESSED = T SAVEWITHCRC
= T WINDOWSIZE = 250000.0
USEHAMMING = T PREEMCOEF = 0.97
NUMCHANS = 26 CEPLIFTER = 22
NUMCEPS = 12 ENORMALISE = F
```

To run a HCopy a list of each source file and its corresponding output file was created. The first few lines look like:

```
data/train/rimwe01.SIG data/MFC/rimwe.MFC data/train/rimwe02.sig
data/MFC/rimwe02.MFC                      data/train/rimwe03.sig
data/mfc/rimwe03.mfc
data/train/sil10.sig data/MFC/sil10.MFC
```

One line for each file in the training set. This file tells the tool kit to extract features from each audio file in the first column and save them to the corresponding feature file in the second column. The command used is:

```
HCopy -T 1 -C config.txt -S hcopy.scp
```

State 1
<Mean> 39

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

<Variance> 39

1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

<State> 2

<Mean> 39

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

<Variance> 39

1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

<State> 3

<Mean> 39

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

<Variance> 39

1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

<State> 4

<Mean> 39

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0

<Variance> 39

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0

<TransP>

0.0 0.5 0.5 0.0 0.0 0.0

0.0 0.4 0.3 0.3 0.0 0.0

0.0 0.0 0.4 0.3 0.3 0.0

0.0 0.0 0.0 0.4 0.3 0.3

0.0 0.0 0.0 0.0 0.5 0.5

0.0 0.0 0.0 0.0 0.0 0.0

<EndHMM>Models for each of the events were also constructed

3,5 Recognition

The recognizer is now complete and its performance can be evaluated. The recognition network and dictionary have already been constructed, and test data has been recorded.

An input speech signal input signal is first transformed into a series of "acoustical vectors" (here MFCs) using the HTK tool HCopy, in the same way as what was done with the training data. The result was stored in a file known as test.scp (often called the acoustical observation). The input observation was then processed by a Viterbi algorithm, which matches it against the recognizer's Markov models using the HTK tool HVite as follows:
HVite -A -D -T 1 -H model/hmm3/hmmdefs.txt -irecout.mlf -w wdnetdict hmmlist.txt -S test.scp.
Where:

Hmmdefs.txt contains the definition of the HMMs. It is possible to repeat the -H option and list the different HMM definition files, in this case: -H model/hmm3/hmm 0.txt -H model/hmm3/hmm 1.txt and so on. but it is more convenient (especially when there are more than 3 models) to gather every definitions in a single file called a Master Macro File. For this study, this file was obtained by copying each definition after the other in a single file, without repeating the header information(Llanos, 2016).
The intention was to get the subjects read the data into a speech recognition program, which would be used for the analysis. The data was later loaded into the program, first for some acoustic investigation, and second, for perceptual analysis. Both quantitative and qualitative analysis was adopted for an insightful result. Again, the study made use of two groups of respondents. The first was the experimental group (EG) while the second was the Control Group (CG). The latter was employed as the standard to measure the performance of the former. The Control was made up of one native speaker of English born, bred and educated in Britain up to Master's Degree; this respondent admittedly speaks a variety of RP fluently. The EG comprised Nigerian graduates and indigenes of Ondo state, Nigeria: Ondo, Akure and Idanre. This arrangement was to enable the researcher have a wide coverage of the major linguistic groups so that the result could be adjudged as truly widespread. This way, generalization was possible.

4. IMPLEMENTATION

4.1 Software Implementation interface

The software package used for the actual voice training and implementation was Natural Reader's Vox Sigma (NRVS). The interface is a simple and easy to navigate graphics with features that make voice training, recording and playing easier. The interface supports voice output in MP3, WMA and AMR audio format.

The interface consists majorly of a menu bar, button/navigation bar and a text area for text input and output. The button/navigation bar consists of record, play, pause, next and previous button as well as a button to select tone and inflection of recorded speech.

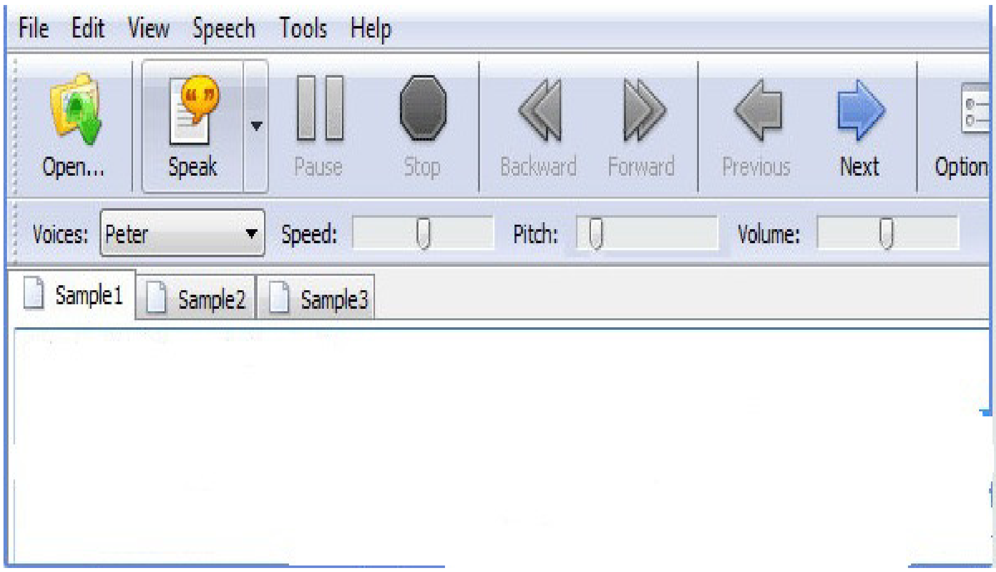

Fig. 2: Voice training toolkit interface

This refers to the act of collecting the voice input of the sample key speakers from the three sampled dialects in Ondo State, Nigeria (Ondo, Akure and Idanre). Each speaker was made to read the collection of text using a headset and a voice input device. The sample text used includes:

> *"The internationalization of Nigerian mother native language is the key factor contributing to ingenuity in our native languages. Paul and Shawn viewed this from a professional perspective in 1995. However, there is a general controversy regarding the Overview of what these native languages should look like".*

This was fed into the blank text area of voxsigma and used as a standard against which the pronunciations of the native speakers were measured. The sample text was stored as sample_01.ntr file.

4.2 Voice Input from the Ondo speaker

The Ondo speaker was requested to read the sample text as an input to the software program thereby allowing the program to process the input and produce a text equivalent as an output. The text equivalent of the pronunciation of the Ondo speaker is as follows:

> *"The **inter-la-tion-lani-sa-tion** of Nigerian mother **lative nanguage** is the key factor contributing to **ingeluity** in our **lative nanguages**. Paul and **Son**viewed this from a **profe-sso-nan** perspective in 1995. However, there is a general contro-**fa-**sy regarding the **O-fa-few** of what these native languages **sud** look like".*

The program represented wrongly pronounced words as bold and italicized

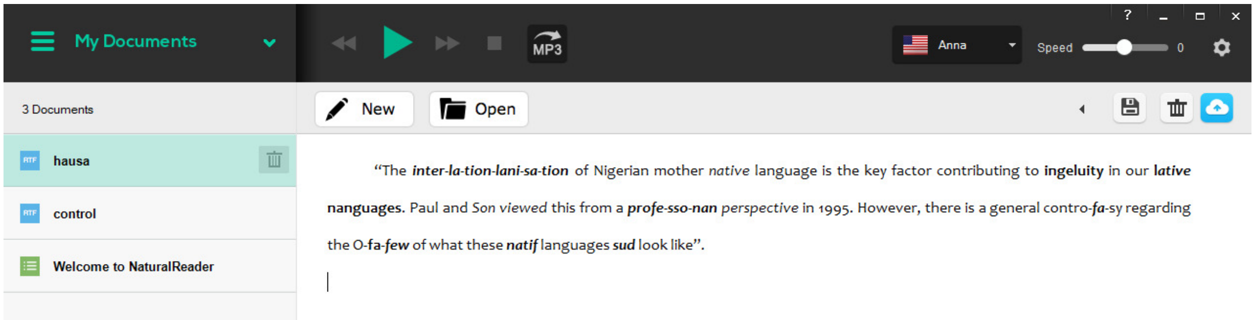

Fig. 3 Output interface for sample_01.ntr (Ondo speaker)

4.3 Voice Input from the Akure speaker

Like the Ondo speaker, the Akure speaker was also made to take the software through voice training by reading the same set of text into the software program. The program also produced a text equivalent of the voice input from the Akure speaker as follows:

> *"The **internationnaliza-shin** of Nigerian mother **native** language is the key **vactor contributing** to **ingeluity** in **awa native nanguages**. **Paul** and **Shin peewd** this **prom** a **frope-tion-nal fersfectip** in 1995.However, there is a general controversy regarding the overview of what these native languages **shild** look like".*

The input from the Akure speaker was represented as sample_02.ntr and the output interface is as shown below:

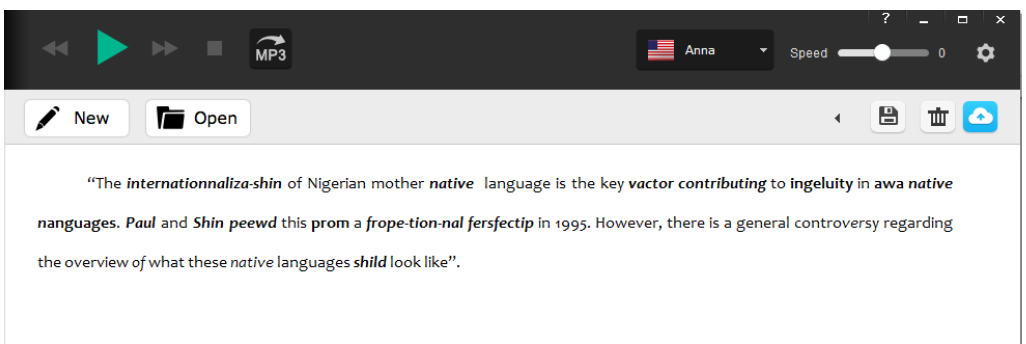

Fig. 4: Output interface for sample_02.ntr (Akure speaker)

The program represented wrongly pronounced words as bold and italicized

4.4 Voice Input from the Idanre speaker

Like the Ondo speaker and Akure speakers, the Idanre speaker was also made to take the software through voice training by reading the same set of text into the software program.

The program also produced a text equivalent of the voice input from the Akure speaker as follows:

> *"The **intenationnnizashon** of Nigerian mother native language is the key factor **conttributing** to ingenuity in our **natife languages**. Paul and Shawn view this from a **professional** perspective in 1995. However, **diar** is a **geleral** controversy **regarding** the overview of what these native languages should look like".*

The input from the Idanre speaker was represented as sample_03.ntr and the output interface is as shown below:

The program represented wrongly pronounced words as bold and italicized

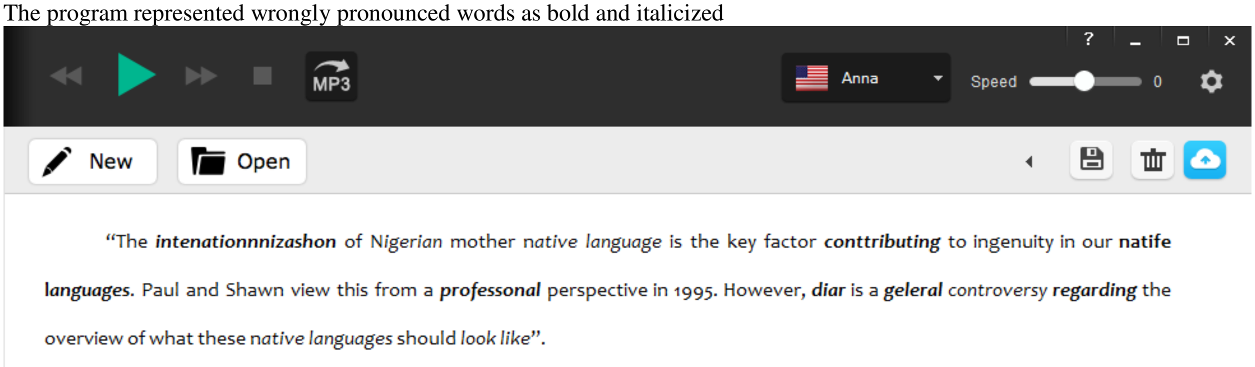

Fig. 5 Output interface for sample_03.ntr (Idanre speaker)

4.5 Program Pronunciation Learning (PPL)

In this process, all the wrongly pronounced words were mapped with their respective correct pronunciations. This mapping can then be used as an input to the "learning" module of software programs (Foote, et al. 2016). This will allow the programs to replace the voice input matching those wrongly pronounced words with their corresponding correct spelling in the text output thereby improving the efficiency and performance of the computer program in use.

Although, the learning module of the software program used in this study was not activated by the vendor of the program (Vocapedia research) even in paid versions, this functionality can be explored in the future since it is a part of the software documentation.

5. CONCLUSION

This paper has shown the use of Vitebri Algorithm as an assisting module for training and learning language from native speakers. Such approach will definitely improve the speech to text knowledgebase and also influence the building of effective voice retrieval systems across dialects. As observed, effective communication plays some vital roles in any given society thus the need to build multidialectal text-to-speech and voice-to-text computer programs that can automatically correct wrongly pronounced words. Our system, leveraging on the HMM Toolkit has been able to generate an acceptable report.

6. RECOMMENDATION

Based on the finding of this study and the challenges experienced during implementation, it is hereby imperative to make the following recommendation.

i. The intelligibility of Nigerian English must be improved so as to minimize the influence of mother tongue in the usage of English. This can be achieved by the Government ensuring that the teachings of English Language at all levels of Education must conform to international standards.
ii. The use of text-to-speech and speech-to-text computer programs should be highly adopted by fairly educated users of English who are computer literate. This will draw individual's attention to his/her grey areas of English words pronunciation.
iii. Speech Recognition Programs should be improved so that they can be adapted for learning of "seemingly" difficult English words for speakers whose mother tongue are not English Language.

REFERENCES

1. Adesanoye, F. (1973). “A study of varieties of written English in Nigeria”. Unpublished Ph.D.thesis, University of Ibadan, Ibadan.
2. Adetugbo, A. (1977). “Nigerian English: fact or fiction?”. Lagos Notes and Records: A Journal of African Studies. Vi: 128-141.
3. Ajani, T. (2007). “Is There Indeed A Nigerian English‟. Retrieved October 17, 2009 from http//www.scientific journals.org/journals2007/articles/1084.htm.
4. Aladeyomi, S.A. and Adetunde, A. K. (2007). “Errors of Segmental Phonemes in the Spoken English of Nigerian English Newscasters”. In Medwell Journal: the Social Sciences, 2(3). Pp. 302-306.
5. Alptekin, Cem (2002). “Towards Intercultural Communicative Competence in EIT”. In: English Language Teaching Journal, 56 (1). Pp. 57-64.
6. Awonusi, V. O. (1985). “Sociolinguistic Variation in Nigerian English”. UpublishedPh.D Thesis. University of London.
7. Banjo, A. (1982). English language studies in a multilingual setting in Nigeria. Text of lecture delivered at theUniversity of Pennsylvania, U.S.A. 29. 1-16.
8. Banzina, E., Dilley, L. C., & Hewitt, L. E. (2015). The Role of Secondary-Stressed and Unstressed–Unreduced Syllables in Word Recognition: Acoustic and Perceptual Studies with Russian Learners of English. Journal of psycholinguistic research, 1-19.
9. Bloem, J., Wieling, M., & Nerbonne, J. (2016). Automatically identifying characteristic features of non-native English accents. The future of dialects: Selected papers from Methods in Dialectology XV, 1, 155.
10. Bernardez, E. (2008). “Collective cognition and individual activity: Variation, language and culture”. In Frank, R,Dirven, R., Zienike, T.C., &Bernadez, E. (eds.). Body, language and mind: Sociocultural situations. 2. NewYork. 138-147.
11. Bobda, S. A. (1995). The phonology of Nigerian English and Cameroon English. In Bamgbose, A.,Banjo, A., andThomas, A.(eds.). New Englishes. A West African perspective. Ibadan: Mosuro. 248-268.
12. Brann, C.M.B. (2006). “Macro-sociolinguistics in Africa: A tentative introduction”. In Language in education &society: An anthology of selected writings of C.M.B.Brann (1975-2005). Maiduguri/ Jos: Fab Education Books.31-48.
13. Brann, C.M.B. (2006). “Mother tongue, other tongue and further tongue”. In Language in education and society: Ananthology of selected writings of C.M.B.Brann (1975-2005). Jos: Fab Education Books. 2-21.
14. Brann, C.M.B. (2006). “The delectable field of socio-linguistic variety differentiation”. In Language in education &society: An anthology of selected writings of C.M.B.Brann (1975-2005). Jos: Fab Education Books. 63-66.
15. Chu, C. H. (2016). In-service Teachers’ Intelligibility and Pronunciation Adjustment Strategies in English Language Classrooms. English Language Teaching, 9(4), 30.
16. Ekong, P. A (1980). “Investigation into the Intelligibility of a Possible Standard Model for Nigerian Spoken English”. In Uno, S. O. (ed). Journal of Language. Arts and Communications 1(i): 1-11.
17. Ekong, P. A (1982). “On the Use of Indigenous Model for Teaching English in Nigeria”. World Language English. 1.2. Pp 87-92.
18. Fakuade, G. (1998). “Varieties of English in Nigeria and Their Social Implications”. In: Fakuade, G. (ed.). Studies in Stylistic and Discourse Analysis. Yola: Paraclet Publishers. Pp. 123-136.
19. Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2016). Pronunciation teaching practices in communicative second language classes. The Language Learning Journal, 44(2), 181-196.
20. Franklin, A., & McDaniel, L. (2016)). Exploring a Phonological Process Approach to Adult Pronunciation Training. American Journal of Speech-Language Pathology, 25(2), 172-182.
21. Giles, H. and Coupland, N. and Coupland, J. (1991). Contexts of Accommodation Developments in Applied Linguistics. Cambridge: Cambridge University Press. Gut, U. (2002). “Nigerian English A Typical West African Language?” Proceedings of TAPS, Bielefield. Pp. 56-67.
22. Gonçalves, A. R., & Silveira, R. (2015). Frequency Effects On The Intelligibility Of English Words With High Front Vowels/Efeitos Da Frequência Na Inteligibilidade De Palavras Do Inglês Contendo Vogais Altas Anteriores.Organon, 30(58), 127.
23. Jenkins, J. (2000). The Phonology of English as an International Language: New Models, New Norms, New Goals. Oxford: Oxford University Press Jibril, M. M. (1979). “Regional Variation in Nigerian Spoken English.
24. Kruse, J. (2016). Accent variation reflected in the standard writing system of English. The Routledge Handbook of the English Writing System, 175.
25. Olaniyi, O.K. (2006). Approximation to standard British English: Nigerian English suprasegmental. In Journal ofNigeria English Studies Association (JNESA) Book of proceedings. 109-123.

26. Olaniyi, O. K. (2011). "Nigerian Spoken English: A Phono-Sociolinguistic Study". In: Udofot, I. M. &Udoudom, J. (eds.). English Usage in Nigeria Since 1842: Patterns and Change. A Festschrift for Prof David Eka.Uyo: Devconsort Services Ltd.
27. Llanos, F., & Francis, A. L. (2016). The Effects of Language Experience and Speech Context on the Phonetic Accommodation of English-accented Spanish Voicing. Language and Speech, 0023830915623579.
28. Munro, M. J., Derwing, T. M., & Thomson, R. I. (2015). Setting segmental priorities for English learners: Evidence from a longitudinal study.International Review of Applied Linguistics in Language Teaching, 53(1), 39-600)
29. Preston, D. (1989). Sociolinguistics and second language acquisition. Oxford: Blackwell Publishers.
30. Rofiq, A. (2016). A PHONOLOGICAL OUTLOOK ON THE DIFFICULTIES OF LEARNING ENGLISH AND ITS IMPLICATION ON THE TEACHING ENGLISH AS A FOREIGN LANGUAGE.
31. Schmied, J. (1989). Linguistics in the service of Africa. Bayreuth: Bayreuth African Studies Series 18-23.
32. Selinker, L. (1972). Interlangauge. IRAL, 10, 209- 231. http://dx.doi.org/10.1515/iral.1972.10.1-4.209.
33. Trudgill, P. (1986). Dialects in Contact. Oxford: Blackwell Publishers.
34. Tiffen, Brian W. (1974). "The Intelligibility of Nigerian English". Ph.D. Dissertation, University of London.
35. Udofot, I M, (2011). "The Rhythm of Standard Nigerian English". In: Udofot, I. M. &Udoudom, J. (eds.). English Usage in Nigeria Since 1842: Patterns and Change. A Festschrift for Prof David Eka.Uyo: Devconsort Services Ltd
36. Wong, S. W., Mok, P. P., Chung, K. K. H., Leung, V. W., Bishop, D. V., & Chow, B. W. Y. (2015). Perception of native English reduced forms in Chinese learners: Its role in listening comprehension and its phonological correlates. TESOL Quarterly.