



ASHESI

ASHESI UNIVERSITY

**ASSESSING THE IMPACT OF (NIGERIAN, GHANAIAN,
ZIMBABWEAN, RWANDAN AND KENYAN) ACCENTS ON
ENGLISH-SPEAKING SPEECH RECOGNITION SYSTEMS**

UNDERGRADUATE THESIS

B.Sc. Management and Information Systems

Ekab-Osowo Tawo

2019

ASHESI UNIVERSITY

**ASSESSING THE IMPACT OF (NIGERIAN, GHANAIAN,
ZIMBABWEAN, RWANDAN AND KENYAN) ACCENTS ON
ENGLISH-SPEAKING SPEECH RECOGNITION SYSTEMS**

UNDERGRADUATE THESIS

Undergraduate Thesis submitted to the Department of Computer Science,
Ashesi University College in partial fulfilment of the requirements for the
award of Bachelor of Science degree in Computer Science

Ekab-Osowo Tawo

2019

DECLARATION

I hereby declare that this undergraduate thesis is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:

.....

Candidate's Name:

.....

Date:

.....

I hereby declare that preparation and presentation of this undergraduate thesis were supervised in accordance with the guidelines on supervision of undergraduate thesis laid down by Ashesi University.

Supervisor's Signature:

.....

Supervisor's Name:

.....

Date:

.....

Acknowledgement

I would like to thank and express my sincere gratitude to my supervisor Dennis Owusu of Ashesi University College for his time, patience and commitment as well as providing adequate feedback to the writing and completion of this Thesis. I would also like to thank my parents for their input towards the success of this study and to Tobi Adelaja for his support. Finally, I would like to thank the participants who took time out of their busy schedules to participate to the success of this research.

Abstract

The use of electronic devices such as computers and mobile phones have rapidly increased over time. Technological powerhouses have inculcated some form of speech recognition in their devices and as such, the increase in usage of this devices provides the increase in interaction between humans and speech recognition. Speech recognition involves adequately transcribing spoken sentences or words into text. This interaction is present all over the world, Africa inclusively, and the continent with its diverse languages comes with its diverse accents and pronunciation of words. A successful interaction between humans and speech recognition involves these systems adequately transcribing the spoken word to the correct text. However, with these transcriptions, errors are prone to occur due to the fact that, mismatch in accents can increase the error rate of a speech recognition system by more than 100% ^[16]. This Thesis explores to a very large extent the impact that certain African accents such as Nigerian, Ghanaian, Zimbabwean, Rwandan and Kenyan have on the accuracy and transcriptions made by English-speaking speech recognition systems. The findings made in this study states that to a very large extent, the various African accents used in the study does affect the accuracy of the speech recognition systems used for the study.

Table of Content

DECLARATION.....	i
Acknowledgements.....	ii
Abstract	iii
Table of Content.	iv
Chapter 1: Introduction	1
1.1: Background	1
1.2: Motivation	2
1.3: Research Objective	3
1.4: Research Approach	4
1.5: Research Question	5
Chapter 2: Literature Review	6
Chapter 3: Methodology.....	16
Chapter 4: Analysis and Presentation of Data.....	20
4.1: Experimentation.....	21
4.2: Analysis	22
4.2.1: Implementation	22
4.2.2: Results.....	24

4.3: Discussion	28
Chapter 5: Conclusion & Future Work	33
References	35
Appendix	vi
Appendix A: Sentences Used	vi
Appendix B: Google Transcriptions.....	.vi
Appendix C: IBM Transcriptions.....	xxvi

Chapter 1: Introduction

Today, the increase in the patronage of speech recognition systems have become quite evident as giants in the technology sector such as Apple, Google, Amazon, Microsoft, IBM to name a few have inculcated forms of speech recognition in their devices or have built or acquired speech recognition systems of their own. The growth and patronage of these brands are evident in the value and worth of these brands. An example involves the fact that for eight straight years, Apple has been the most valued brand and as of 2018, Google comes second ^[3]. Thus, the use of these brands keeps increasing furthermore increasing the use of the speech recognition systems provided by these companies through their devices. Africa, a diversified continent with numerous languages and accents, has also added to the growth and patronage of devices made by these brands. Thus, they are also at the helm of the exploration of this intriguing component, Speech recognition. Now, with this somewhat new component comes with its disadvantages

1.1 Background

Speech recognition as effective as it may seem that is; the ability to transcribe voice to text, also gives rooms for errors to occur during its transcription. Such errors come from the mispronunciation of words from the user, or misinterpretation of words by the system. The root of this problem is the fact that speech recognizers may be unable to understand specific accents that are spoken to it regardless of if these accents used are speaking English. As such, this results in a mismatch in accents, and this is a significant attribute that aids in an increase in error rates of speech recognition systems. The mismatch in accents can increase the error rate of a speech recognition system by more than 100% ^[16] as it may not understand the adequate pronunciation of words thus leading to misinterpretation. Though advancements in technology have subjected the significant causes of errors in speech recognition systems to noise and speaker variations, speaker accents also play a significant role in determining the progression

of these systems. Now, a problem that speech recognition systems have that may provide a reason as to why accent affects the credibility and reliability of these systems involves the training of speech recognizers, these training usually occur in the UK or US English accents [16]. Furthermore, though a large part of the African continent had England, as well as other European countries, colonizing it, the accents of Africans are unique and relatively different thus proving to be a hindrance in speech recognition systems recognizing African speech regardless of it being in English.

Speaker variability is a factor that affects the performance of automatic speech recognition (ASR). The attributes of this speaker variability involve gender and accents to name a few. Gender, on the one hand, has had a decent amount of research factored into it when understanding speech recognition. With regards to accents, little research has occurred. The accent issues have two related research, and these are accented adaptation through pronunciation modeling and accent identification. People with heavy accents are prone to make more pronunciation errors regarding standardized errors [4]. The African accent is a heavy accent as the pronunciation of English words varies from location to location. Even in similar geographical areas, the difference in mother tongue also affect the accents and thus change the pronunciation of words. Therefore, this is just enlightening as to how the speech recognition system though included in numerous devices seem to misinterpret and provide errors. The error rate as earlier stated is as a result of multiple factors, and accents are one of them, that then leads to the motivation and purpose of this research.

1.2 Motivation

The motivation for this research stems from the fact that using various speech recognition technologies, be it Siri, Google Voice or Cortana, errors are prone to occur. As earlier stated in the paper, accents have a substantial effect on the reaction from this recognition technologies and as such mistakes are likely to happen every day when accessing any of these

speech recognition technologies. The impact of accents on speech recognition systems are crucial in undergoing research because of firstly, the technological advancement that occurs every day around the world. With regards to this point, the prevalence of these systems would mainly increase because the world is moving towards an Internet of Things (IoT) and Artificial Intelligence (AI) as well as more speech recognition instruments to get about daily activities. Also, with the world becoming a global village, Africa would compete at almost the same level with other technological powerhouses, and these advancements and the use of technologies such as speech recognition systems would eventually be a norm in the continent.

1.3 Research Objective

Having looked at speech recognition systems and the various facets to it, as well as the importance of accents in understanding, acquiring credibility and reliability of these systems, **the objective of this research is to understand and assess to a considerable extent the impact of some African accents: Nigerian, Ghanaian, Zimbabwean, Rwandan and Kenyan, on English speaking Speech Recognition systems.** Moreover, there has been researching into the effects of differences in accents in speech recognition, the comparison of the UK and US accents in speech recognition, as well as the impact of the South African accent with regards to speech recognition. This paper seeks to look at some African accents from 3 major regions to work with a more diverse experimental scope. The significance of this research is the fact that it helps in understanding to what extent how some African accents affect English speaking speech recognition systems. Understanding to a considerable extent how African accents affect English speaking speech recognition is very important in the way Africans interact with such technologies as well as aiding African users with necessary understanding as to how and why errors exist when interacting with such techniques. Another importance of this research is the novelty aspect of it. The exploration of this research is relatively new that is, with regards to the scope and the accents in question. Thus, being able

to adequately understand the impact specific African accents have on English-speaking speech recognition can also be a foundation to building more research into this problem space. It can be the start of building African speech recognizers or even language speech recognizers, depending on the depth of the research being carried.

1.4 Research Approach

For this research, the methodological approach would be experimental as the use of human subjects would be considered. The participants would be college students precisely freshman from a diversified college in West Africa. These participants would be either randomly chosen or would have the opportunity to volunteer to the success of this research. Though the selection of the participants may be random, a more stratified approach would be considered, as the freshmen would first be grouped based on their nationalities before the randomization occurs. The focus of this research is around the following African nationalities; Nigerians, Ghanaians, Zimbabweans, Kenyans, and Rwandans. The exact number of participants from each group, ranges from 2-6 participants, with ages 17-21. Also, all forms of gender are allowed for the research. Furthermore, though all gender forms are applicable, the experiment which would be undertaken for this research would strive to have gender balance. The need for the gender balance is to strive for consistency during the project as gender is a possible factor for the occurrence in error rates by speech recognition systems. Aside from the participants, various speech recognizers would be made available, and these recognizers can either be trained for this research or not. However, it is highly unlikely that training of speech recognizers would be adequate for this research and as such, the considered speech recognizers would be used as they are for the experiment. For this research, the following speech recognizers include; IBM's Watson, Microsoft's Cortana, Google's Hello Google to name a few.

1.5 Research Question

- How does the Nigerian, Ghanaian, Zimbabwean, Rwandan and Kenyan African Accents affect English speaking speech recognition systems?

Chapter 2: Literature Review/Related Work

There is first a need to understand what accents are and if at all they affect speech recognition systems to aid the progress of this research. By doing this, as previously stated, it would assist in giving or guiding the purpose of this paper to a fruitful conclusion and the objectives of this research can be met.

Accents are patterns of pronunciation used by a speaker for whom English is the native language or, it involves this speaker patterns being distinctive to a community or social grouping. Furthermore, an accent is something which every speaker has, and it is something unique to that individual. Also, accents to a considerable extent are characteristics of people or a group of persons belonging to some geographical region, and belonging to a particular social class, and, it may even be representative to a speaker's sex, level of education or level of education. Around the world, the native languages of some continents, countries, and regions are not English, and as such the accents with which these nationals would speak it may be referred to as Foreign Accent. Foreign accents are pronunciation patterns seen as typical of speech of those who English is not their native language. These patterns would expect reflections from many phonological and phonetic characteristics of their mother tongue. An individual's accent is a potent tool as it helps as indicators of various aspects of that individual. Based on a person's pronunciation, the knowledge of where that individual comes from, where he/she grew up and, in some cases, where that individual lives may be observed or known. Accents also to some extent reveal a person sex and ethnicity. In the case of the sex of an individual, this isn't limited to the standard biological differences, but also how the individual sounds. For ethnicity, there exists an infusion of phonetic characteristics which are specific to their mother tongue ^[14].

Thus, since to a reasonable extent, understanding of accents have been established, the need to gain knowledge on what speech recognition systems are, how it works as well as

examples would be useful. Speech is a sequence of words encoded by a speaker into a continuous acoustic signal. The idea of speech recognition and the speech recognition problem involves a system, device or computer adequately identifying the specific words which are used by the speaker or more correctly, the words which were intended by the speaker. A basic overview of speech recognition is speaking to a device and having that device answer you or actively perform the task you want it to act or to interpret what has been said to it effectively. This form of technology has become increasingly needed; for example, an idea of speech recognition involves aiding deaf people with communication. This form of speech technology is about translating speech to text ^[8] however, as important as this may be, it doesn't connote what the paper wants to achieve or what this paper is trying to address. As earlier stated, this paper is trying to understand accents, African accents to be precise against English speaking speech recognition systems. That is, speaking to the speech recognition systems for them to trigger an effect or result in action.

Theories of acoustic-phonetics influenced the early attempts to develop speech recognition systems or automatic speech recognition systems (ASR). Acoustic phonetics describe phonetic elements of speech as well as trying to explain how the basic sounds of languages can be realized acoustically in the spoken utterance. Practically, this meant that for example in creating a vowel sound, the vocal cords need to have some form of vibration. The year 1952 was when Balashek, Davis, and Biddulph of Bell laboratories constructed a system for isolated digit recognition for a speaker. This system was only able to recognize spoken numbers from 1-9. Other systems built during the 1950's involved Olson and Belar of RCA Laboratories system which could identify ten syllables of a single talker. At MIT Lincoln Lab, Forgie made a speaker-independent ten vowel recognizers. In the 1960s, the Japanese laboratories showcased their skills in building some form of unique hardware to carry out speech recognition tasks. The vowel recognizer of Suzuki and Nakata at the Radio Research

Lab Tokyo was a notable recognizer built by the Japanese. Another early recognition system constructed by Fry and Denes of University College in England which involved their recognizer's ability to recognize four vowels and nine consonants; however, by teaching statistical information about allowable phoneme sequences in English, they were able to increase the overall phoneme recognition accuracy for words consisting of two or more phonemes.

In the 1970's, recognition of common vocabularies (order of 100-1000 words) using simple template-based, pattern recognition methods were developed. During this decade, the critical technology created was the pattern recognition models, the pattern clustering methods for speaker-independent recognizers to name a few. The 1980s brought about the tackling of speech recognition problems based on statistical methods that involve a wide range of networks handling language structures. The critical technologies introduced this period included the hidden Markov model (HMM) and the stochastic language model. These models together enabled powerful new methods for virtually handling any continuous speech recognition problems effectively and with high performances. The 1990s came with the ability to build large vocabulary systems with unconstrained language models as well as constrained task syntax models for continuous speech recognition and understanding. The last few years have brought about the introduction of extensive vocabulary systems with full semantic models, integrated with text-to-speech (TTS) synthesis systems and multi-modal inputs. During this period, the use of machine learning to improve speech understanding and speech dialogs as well as the introduction of mixed-initiative dialog systems to enable user control ^[2]. In no time the existence of speech recognition technology was being paved and now large technology companies today have their speech recognition systems. Amazon's Alexa, Google's Hello Google, Apple's Siri, Microsoft Cortana to name a few are examples of popular and mostly patronized speech recognition systems. Speech recognition technologies have entered the

marketplace benefiting these large technological powerhouses with revenue as well as innovative advancements.

The first step in beginning the work with speech recognition systems and accents is accented identification. Accent identification is an adequate tool that can help in training the recognizer to be aware of what kind of accent that is being made available to it. In their research^[13] worked on trying to identify foreign accents using an automatic identification accent system. Though the reasons for their study involved identifying accents from 6 European countries, the idea is still valid. From their work, it showed that it is preferable to train a recognizer with a mixture of accents as it would aid in accurate identification from speech recognizers. Their research also noted the fact that automatically identifying non-native accents is a difficult task as well as the fact that their system or analysis takes into consideration only pronunciation networks per word. This is a restriction in accent identification as it doesn't take account the multiple ways which an accent can be pronounced. For their current work as well as future work,^[13] are working on deriving various pronunciation networks and for the future, they intend to explore rhythmic cues to make a more effective accent identification.

After the identification of accents, the next research (Yan & Vaseghi), looked at the comparison between the UK and US English and how it affects the results of speech recognition systems. The experiments conducted in the research were performed doing a detailed study of the acoustic correlates of accents using intonation pattern and pitch characteristics. The paper further goes to explain the differences in accents and the elements of the differences between the UK and US accents. The early stages of their experiment were to do a cross accent recognition quantifying the accent effects between the British accent (BrA) and American accent (GenAm) on speech recognition. After training the specific speech recognizers based

on the national accents, the results from the first experiment is below.

Accent	British model	American model
British input	12.8	29.3
American input	30.6	8.8
Average	21.7	19.1

Per their beginning research, they recorded that the American English achieves 31% less error than the British English in a matched accent condition. However, an inconsistent accent of the recognition systems deteriorates the performance as the results got worse with 139% for recognizing British English with American models and 232% for identifying American English with British models. The next step in their experiments was to an acoustic feature of the two accents in question. This feature was done using duration, pitch characteristics, and prosody. These features helped determine the understanding that apart from phonetics, the difference in the slope of rising and fall exist in bringing about contrasting results when there is a mix of accents in speech recognition.

With the breakthrough and innovative advancements and development of speech recognition, speaker variability still affects the performance of these recognition systems. Part of the factors that affect speaker variability, accents is among the most important. The purpose of their research was to show the accent issues that occur in large vocabulary continuous speech recognition. Based on cross-accent experiments, it shows that accent problems are very dominant in speech recognition. The research areas that are related to accent issues in speech recognition involve accent adaptation through pronunciation and accent identification. Speakers with heavy accents tend to make more pronunciation errors regarding standard pronunciation. For their research, the looked at the impact of accents on speech in two views; the cross-accent speech recognition experiments were carried out and secondly, multivariate

analysis tools, PCA and ICA, were applied to confirm the importance of accent in speaker variability qualitatively.

The research carried out extensive research to show or to investigate the impact of accent on state-of-the-art automatic recognition systems. Furthermore, they examined some critical factors of speaker variability and how they correlate with each other. The whole experiment consisted of 980 speakers with 200 utterances per speaker, and they were from 2 accent areas, Beijing (BJ) and Shanghai (SH).

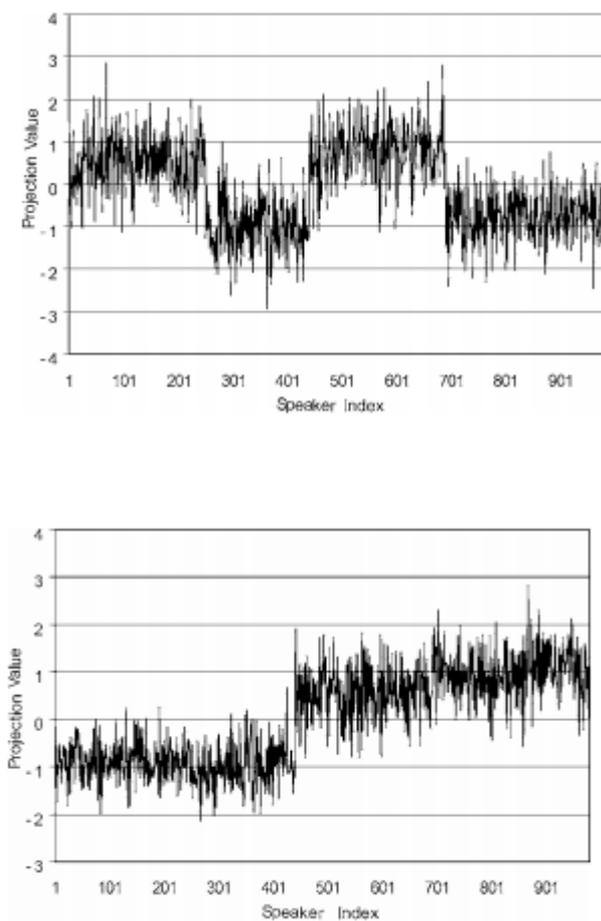


Figure 1. Projection of all the speakers onto the first independent component (The first block corresponds to the speaker sets BJ-F and SH-F, and the second block corresponds to the sets BJ-M and SH-M).

Table 4. Speaker distribution for speaker variability analysis.

	Beijing	Shanghai
Female	250 (BJ-F)	190 (SH-F)
Male	250 (BJ-M)	290 (SH-M)

The images above show the gender and accent distribution; it shows that the first independent component corresponds to gender characteristics of a speaker as projections from this component almost separate all speakers into gender categories. In the next figures, four subsets

occupy four blocks, and from that based on their findings, it is evident that the components have strong correlations with accents. Thus, from their research, they were able to establish the fact that both cross-accent experiments and speaker variability analysis show that accent is one of the most critical factors leading to fluctuating performance of ASR systems. Conclusively in their research, they were able to come up with the fact that yes speaker variability does affect speech recognition performance significantly as well as the fact that accent is one of the main factors that cause variability and should impact the recognition. The paper showed the fact that there was a 40-50% error increase from cross-accent speech recognition. Also, based on their experiment with PCA/ICA, they could qualitatively confirm the fact that accent is another dominant factor, adding to gender in speaker variability ^[9]. The second aspect of their research involved accent adaptation and identification. This aspect of the study is heavily linked to (Teixeira et al.). For this speech adaptation technique, the researchers used data available to sort of train the recognizers for them to see a substantial difference between the accents and for them to be able to classify these accents.

Furthermore, ^[10] analyzed speaker variability. Speaker variability is an essential study in speech recognition as it highlights the various factors that affect the reliability and the interpretation which the speech recognition systems would provide based on the speech spoken to them. The paper highlights that gender and accent are significant factors in the analysis of speaker variability. In conducting their research, it involved two powerful statistical multivariate analysis. The plans include the principal component analysis (PCA) and the independent component analysis (ICA). The research yielded that, using ICA representation, they achieved about 6.1% and 13.3 error rates in gender and accent classification.

In their paper, ^[11] speaks about how the performance of a speech recognition gets degraded when the presented accent is different from the training set. Furthermore, the purpose of this paper involved using a much faster approach to accent classification using phoneme-class

model. The article also shows how the transformation of native accent pronunciation dictionary, can be changed to that for accented speech. The results from the research conducted show that the accent-adapted dictionary reduces recognition error rate by 13.5%. Kat and Fung used a mixture of feature-based and model-based discrimination, and since it was a small amount of data, they used phoneme-class HMMs. This phoneme consists of six classes which include; stops, affricates, fricatives, nasals, semivowels & glides and finally vowels. The paper further uses various features to show the differences between speaking styles and structures of two languages. The first feature used was energy. The figure below shows the differences between the main accents used in the research based on the energy feature.

Figure 1: Average mean energy of various phone classes

Phone classes	American	Cantonese
vowels	1035.1	506.65
nasals	601.70	252.28
stops	147.74	55.63
affricates	370.66	89.87
fricatives	224.79	117.99
semi-vowels & glides	1282.49	522.93

Another feature considered in the paper was Fundamental frequency. Beneath this feature, it is found that human perception indicates that listeners base their accent classifications partly on prosodic features such as pitch movements, rhythms and pausing. The results between the measured accents are below.

Figure 4: The pitch contour of the same utterance spoke by Cantonese speaker(top) and English speaker(bottom)

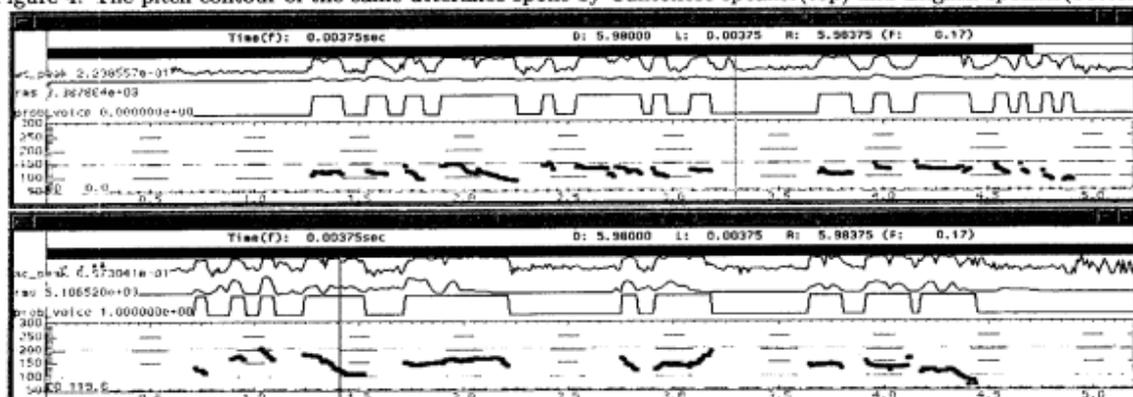


Figure 5: F0 information reduces classification error

parameter set	error rate
full set	14.52%
no dd(F0)	14.58%
no F0	14.65%
no d(F0)	14.67%
no F0 info.	15.33 %

The above images, the first shows the differences in pitch between the two accents used in the paper. The analysis done on the pitch per the research conducted shows that the reason for the higher pitch is that people tend to add their native or first language speaking tongue into foreign tongues. The second image shows that an increase in error rate by 5.6% occurs when ignoring the first derivatives and masking the second derivatives.

In their research ^[4], they give reviews into the understanding of automatic speech recognition and speech variability. The paper speaks about the specific barriers to flexible solutions and user satisfaction under some circumstances. Several factors are related to this issue. From the article, it states that such factors include environment or weak representation of grammatical and semantic knowledge. Also, in the review, it indicates that there are other deficiencies in the dealings of variations that are naturally present in speech. These deficiencies are part of the many factors that affect speech realization to name a few. The factors such as the speaker, gender, speaking rate, vocal effort, regional accent, speaking style to name a few, are variations not modeled well.

In their research, ^[6] looked at assessing and investigating the best possible way to combine speech data from five South African accents to improve the performance of speech recognition. For this paper, since there is difficulty in dealing with multiple accents in under-resourced environments, there are complications in working with automatic speech recognizers. In this article, three acoustic modeling approaches considered include; separate accent-specific models, accent-independent models obtained by pooling training data across accents, and multi-accent models. From the conclusion of their research, it shows that multi-accent models

help in offering a mechanism whereby speech recognition can have its performance optimized automatically as well as aiding in hard decisions involving which data to pool.

Word Error Rate (WER) is a method or a/the standard approach used in estimating or checking the performance of vocabulary speech recognition. This approach mainly involves manually transcribing data in order for the WER to be computed. In order to calculate the WER, the word sequence hypothesized by the speech recognition systems aligned with a reference transcript. Then, the errors are computed as follows;

- S – the number of substitutions,
- I- the number of insertions and,
- D – the number of deletions.

Now, if there are a total number of words in the reference transcription, i.e., N, then the WER can be computed with the following formula;

$$\text{WER} = \frac{I + D + S}{N} * 100$$

As stated just above, working with this approach involves some form of a manual transcription of the reference sentences which in turn becomes the hypothesis^[1].

Chapter 3: Methodology/Approach

An experimental approach with the use of human subjects would be ideal for adequately meeting the objectives this research paper seeks to achieve. The below chapter gives a detailed overview of the measures that would take in completing this research. The significant themes considered in this chapter involve; the materials used in the study, the samples gathered, the techniques used in sample preparations and sample gatherings, as well as every other important aspect of the experiment for this research. The experimental approach considered for this research involves recording the audios from a group of human subjects and then proceeding to test these audios on English speaking speech recognition systems. To examine if or how accents affect the interpretation of speech recognition systems. Firstly, the materials and equipment needed for this research involve audio recorders, a quiet area, and human subjects. The audio recorder is to record the participants adequately as well as play the recordings to the English speech recognition systems.

The sample techniques for this research are systematically carried out before adopting a stratified approach. This sample techniques occur in the selection and grouping of participants for the experiments carried out in this research. For the collection and grouping of the participants, the following activities are to be carried out. The first task involves identifying the nationals of the participants that would be adequate for the research. For this task, some African countries have been selected based on the geographical scope of the study to aid in this experiment. The selection of African countries was in no way random; it was instead done based on the availability and diversity of the countries present in the location of the participants. The countries for this research include Nigeria, Ghana, Rwanda, Zimbabwe, and Kenya. After the Identification of the following countries, the next task involves determining the age grades of the participants. Since the participants are freshmen college students, the average age of the students is between the ages of 17-21. The next step to take involves grouping and selection of

the participants needed for this research. Furthermore, participants are to sign consent forms stating that they are willing to cooperate with the success of this research. Following the selection of the African countries needed for this project, American accents would also be collected. The purpose of collecting American accents is to aid in the interpretation of the results after the experiment.

After the identification of several underlying bases of the experiment, the actual selection and grouping of the participants would be next. The gathering of participants would be a more stratified form of grouping as the participants would be grouped initially according to their nationalities. The next type of stratified grouping for this experiment involves grouping each participant based on gender. The gender-based grouping is for the existence of gender balance as gender also affects speaker variability. Then the participants are selected based on gender and nationality. Another form of selecting participants involves the volunteered selection. For the purpose and objectives of this research to be met, students who may be willing to be part of the study can volunteer as human subjects in the experiment. During this volunteering task, the students would have to come from the selected countries considered for the research as well as the existence of gender balance and the signing of consent forms before choosing.

The next phase of the experiment involves the accumulation of data. In this phase, sentences adequate for the test would be created for the participants to read and get recorded. Some examples of the sentences created for the sole purpose of this experiment include;

- I need a white envelope it's urgent.
- What play are we going to see in the theaters?
- Donuts covered in caramel are just the best.
- I like my bread with butter or jam rather than mayonnaise.
- What route is better if I want to go to the mall?

The motivation behind the structure, grammar, and wordings of the above sentences are influenced by firstly, identifying certain words that are known to bring out an individual's accents. Some examples of these words also underlined in the above sentences include; Caramel, Mayonnaise, Theater, Envelope, Mayonnaise, Route to name a few (Porzucki, 2015). Secondly, experiences with nationals of the considered countries in this research have shown that specific words pronounced by these individuals identify or show their accents.

After speaking, each participant is recorded, and these records saved as experimental data. The documents or the data would be collected based on each nationality and gender. That is, for each participant, his/her recordings are saved as **Subject A Male from Nigeria**. By doing this, it would aid in adequately understanding the kinds of accents subject to a particular gender and country. It would also help during interpretation of results as it would tell how the accents of this gender and this country affect English speaking speech recognition systems.

The next phase of the research involves identifying the speech recognition systems needed. For this phase, some English-speaking speech recognition systems such as Google API, IBM's Watson, Karl D to name a few. After the actual selection of these speech-recognizers, the experiment would begin. The experiment involves playing each recording for each speech recognizer then have the speech recognizers transcribe the data played to it. Once the transcription is done, the sentences and transcribed sentences are stored. The data stored is in the appendix which is situated at the end of the paper. Following this is the computation of the WER for each of the transcribed sentences. The WER which is the word error rate is the metric used to check the accuracy of speech recognition systems. This metric can be used to help meet the objective of this research which involves assessing the impact of African accents on English speaking speech recognizers. The reason as to why the computation of the WER is necessary for this research is due to the fact that, firstly, the word error rate is the standard metric used to calculate or check the accuracy for speech recognizers. As such, it is important

that to understand how accents affect speech recognizers or to check the impact of speech recognition, WER is essential. Secondly, the WER per its formula caters for the number of words, inserted, deleted or substituted during the transcription made by the speech recognizers. By this formula, when computing the WER, it is essential to actually devotedly identify the words which the speech recognizers failed to accurately transcribe ^[1]. Furthermore, in computing the WER, the average of each country, each gender, and each speech recognizer is essential to this research, as such they would be computed.

The next step in the research involves analyzing and interpreting the results obtained based on the WERs computed. In order to understand the impact of African accents on speech recognition, the results (averages) of each country, gender and speech recognizer would be analyzed. To achieve the objective of this research, analysis and interpretation of the results obtained would be done in two ways. Firstly, the WERs of each country according to the transcription by the speech recognizers available would be compared against the official WER of the speech recognizers used. The next comparison would be done between the WERs of the African countries against WERs of American accents. These comparisons would be done to adequately identify how the African accents considered in this research impact the transcriptions made by these English-speaking speech recognizers as compared to the transcriptions made on the American accents collected. The final stage of this research involves concluding with the analysis made on the results obtained as well as answering the research question posed during the course of the study.

Chapter 4: Analysis and Presentation of Data

The data for this research were successfully collected and collated. The participants read the sentences accurately and were successfully recorded. The software used for recording the participants was **Audacity** which enables voice recordings as well the formatting of saved audios in any particular format was used to record the sentences. The recording of the data collected was collated and kept in .WAV form. This was due to the fact that WAV format saves uncompressed audios using the RIFF bit stream format method of storing data. Thus, WAV being able to store uncompressed sounds retains about 100% or the original audio quality. Also, the WAV format is easy to edit ^[7]. During data collection processes, certain limitations occurred which led to adjustment and changes to the methodology of this research. During data collection, only two speech recognition systems were considered; these systems were Google Cloud Platform and IBM's Watson (online version). This was due to the lack of accessibility to other systems. Another change that occurred during data collection was the participants who were selected. The first change that occurred involved the year group of the participants. The original year group identified for this research were first-year students who had just come from their countries. However, due to unforeseen circumstances such as the location (University), had not admitted male students from Kenya or Rwanda having just one male student. These occurrences would have created room for outliers in the data collection processes. Further problems arose amongst the age group of the participants. Amongst some of the first-year students from countries such as Rwanda, Zimbabwe, and even Kenya, the possible participants were above the particular age-group of 17-21 years. As such, this contributed to reasons as to why the methodology initially designed for this research had to be re-adjusted. The adjustments include;

- The participants were chosen from any year group. – This was random selection of participants from the year groups present in the institution.

- Secondly, the participants were between the ages of 17-25. – This was because it was unlikely that any of the considered participants would be above this age range.

Furthermore, into the research processes, after the data collection, each participant was recorded and these recordings saved, the transcriptions of this data were also done using the two speech recognizers (Google Cloud and IBM's Watson). These transcriptions were saved and below is a tabular representation showcasing the collected data as well as the transcribed sentences.

4.1 Experimentation

The data being collected, that is the recordings of the participants after each subject had read the sentences were transcribed using the two speech recognizers considered for this. The sentences read by each participant is located at Appendix A at the end of the paper. Each participant both male and female, from the considered countries had their recordings played to the speech recognizers and the data was transcribed. The data after being transcribed can be seen in Appendix B and C which is located at the end of the paper. Appendix B consists of the transcriptions done by Google cloud speech transcription platform, while Appendix C involves the transcriptions done by IBM Watson. The process of having the data transcribed was that, firstly, the Google cloud speech recognizing platform as the name implies is on the web and as such, each subject's recorded file was uploaded before transcription occurred. This activity was similar to that of IBM data transcription. IBM's recognizer (Watson) is also an online tool that is used for voice to speech transcription. The transcription as stated was done by uploading the recorded data before transcription occurs. Furthermore, an important element of speech recognition and transcription is the period at which the transcriptions occur. By period, referring to the particular time which the transcription was done, this is because speech recognizers undergo modifications on a daily basis. Thus, during the experimentation, the transcription done with the Google cloud platform was specifically done on the 10th of February

2019. For IBM's transcription, this was done the following day, 11th February 2019. Also, for the transcription of the American data, the Google and IBM transcriptions were done on the 10th of April 2019.

4.2 Analysis

The tables in the appendix show the transcribed data which was obtained during the experimentation. Furthermore, to achieve the goal of this research, which involves assessing the impact of these specific African accents on English speaking-speech recognition systems, the Word Error Rate (WER) of each sentence needs to be calculated. These calculations would be done on the transcribed data collected.

4.2.1 Implementation and Results

The computation of the WER was done with Python programming language due to the fact that, with Python programming language, it is easy to compute mathematically-based solutions for problems that may arise. The language, which is in English, provides programmers with; the ability to add, subtract, multiply and divide. The square root, as well as other seemingly complex mathematical problems, can easily be computed and solved using python as the language is relatively easy to understand. Below are lines of code from the computation of the WER for each country and each speech recognizer.

```

def WER(referenceText, hypothesisText):
    brokenReferenceText = referenceText.split()
    brokenHypothesisText = hypothesisText.split()

    insertion = len(brokenHypothesisText) - len(brokenReferenceText)
    deletion = 0
    if (insertion < 0):
        deletion = abs(insertion)
        insertion = 0
    i = 0
    substitution = 0
    while(i < len(brokenHypothesisText) and i < len(brokenReferenceText)):
        if (brokenReferenceText[i] != brokenHypothesisText[i]):
            substitution += 1
        i += 1

    wer = ((insertion + deletion + substitution)/(len(brokenReferenceText)))*100
    return wer

```

Figure 1.0

Figure 1.0 above, is the python function used to calculate the WER of the sentences. The code caters for each of the components of the WER formula. It checks the insertion, substitution and deletion and finally uses this to calculate the WER for each sentence. As part of the implementation, forty-eight (48) text files with the subject's name, country, gender, reference sentence and its corresponding hypothesis were created. The creation of the forty-eight (48) text files was because, per the research, twenty-four (24) participants were selected (4 each) from the considered African countries and America. Thus, since two English-speaking speech recognizers (Google Cloud API and IBM's Watson) were considered for this experiment, two (2) sets of data are required. In other words, twenty-four (24) for Google and twenty-four (24) for IBM totaling forty-eight (48) data text files. After this, the python code was computed in such a way that it reads each line in the forty-eight (48) text files then adequately calculates each participants WER. This involves each sentence WER, as well as the average WER of that individual. The program also computes the average WER of each country identified for this research, as well as the overall average WER of the two speech recognizers; and IBM's Watson. The resulting mechanism works in a particular way. After the python code reads the text files and the WER is calculated for each participant, country and speech recognizer, the

results are then written to an output.txt file. Figures 1.1 and 1.2 below show the exact calculations of the average WER for each country and also WER based on gender.

```

googleTests = "Google"
IBMTests = "IBM"
services = [googleTests, IBMTests]
googleWER = 0
IBMWER = 0
countryGoogle = [0, 0, 0, 0, 0, 0]
countryIBM = [0, 0, 0, 0, 0, 0]
maleCountryGoogle = [0, 0, 0, 0, 0, 0]
femaleCountryGoogle = [0, 0, 0, 0, 0, 0]
maleCountryIBM = [0, 0, 0, 0, 0, 0]
femaleCountryIBM = [0, 0, 0, 0, 0, 0]
outputFile = open('output.txt', 'w+')

```

Figure 1.1.

```

for i in range(48):
    totalWER = 0
    inputFile = open(str(i+1) + '.txt', 'r')
    refText = ''
    hypoText = ''
    subject = inputFile.readline()
    country = inputFile.readline()
    gender = inputFile.readline()
    outputFile.write('\n' + subject+country+gender)
    for j in range(10):
        refText = inputFile.readline()
        hypoText = inputFile.readline()
        wer = WER(refText, hypoText)
        totalWER = totalWER + wer
        outputFile.write("Statement " + str(j) + " WER: " + str(wer) + "% or in Decimal " + str(wer/100) + '\n')
    outputFile.write("Average WER : " + str(totalWER/10) + '\n')
    #Google Calculations
    if i < 24:
        googleWER = googleWER + totalWER/10
        print(i//4)
        countryGoogle[i//4] = countryGoogle[i//4] + (totalWER/10)
        if (gender == "M\n"):
            maleCountryGoogle[i//4] = maleCountryGoogle[i//4] + (totalWER/10)
        if (gender == "F\n"):
            femaleCountryGoogle[i//4] = femaleCountryGoogle[i//4] + (totalWER/10)
        if (i%4 == 3):
            outputFile.write('\n')
            outputFile.write("Country Male WER for Google: " + str(maleCountryGoogle[i//4]/2) + '\n')
            outputFile.write("Country Female WER for Google: " + str(femaleCountryGoogle[i//4]/2) + '\n')
            outputFile.write("Country WER for Google " + country + ": " + str(countryGoogle[i//4] / 4) + '\n')
    #IBM calculations
    else:
        IBMWER = IBMWER + totalWER/10
        print(i//4 - 6)
        countryIBM[i//4 - 6] = countryIBM[i//4 - 6] + (totalWER/10)
        if (gender == "M\n"):
            maleCountryIBM[i//4 - 6] = maleCountryIBM[i//4 - 6] + (totalWER/10)
        if (gender == "F\n"):
            femaleCountryIBM[i//4 - 6] = femaleCountryIBM[i//4 - 6] + (totalWER/10)
        if (i%4 == 3):
            outputFile.write('\n')
            outputFile.write("Country Male WER for IBM: " + str(maleCountryIBM[i//4 - 6]/2) + '\n')
            outputFile.write("Country Female WER for IBM: " + str(femaleCountryIBM[i//4 - 6]/2) + '\n')
            outputFile.write("Country WER for IBM " + country + ": " + str(countryIBM[i//4 - 6] / 4) + '\n')
    #print(gender)
    inputFile.close()

```

Figure 1.2

4.2.2 Results.

Below are the results gotten during the research. The results are as follows, Googles Average WER for each African country, after which Google's Average WER based on gender and the same for USA. Finally, Google's Total WER for the African Countries and overall WER with USA inclusive. Results achieved with IBM Watson are also displayed in the same format.

Results of Google Cloud Platform's WER for African Countries.

Country	Subject	AWER for each Subject	AWER for each Country
Ghana	A	26.90%	
	B	17.30%	
	C	31.26%	
	D	40.67%	
Total WER for Ghana			29.03%
Nigeria	A	43.30%	
	B	10.51%	
	C	27.18%	
	D	29.66%	
Total WER for Nigeria			27.66%
Zimbabwe	A	29.55%	
	B	16.26%	
	C	23.37%	
	D	27.13%	
Total WER for Zimbabwe			24.07%
Rwanda	A	22.42%	
	B	34.68%	
	C	71.48%	
	D	19.10%	
Total WER for Rwanda			36.92%
Kenya	A	17.48%	
	B	33.20%	
	C	38.52%	
	D	25.83%	
Total WER for Kenya			28.76%

Figure 1.3

Google Cloud Platform WER for USA

Country	Subject	AWER for each Subject	AWER for Country
USA	A	13.45%	

	B	14.36%	
	C	6.02%	
	D	21.44%	
Total WER for USA			13.82%

Figure 1.4

Google Cloud Platform WER for African Countries based on Gender

Country	Gender	AWER for Male Subjects	AWER for Female Subjects
Ghana	F		26.90%
	F		17.30%
	M	31.26%	
	M	40.67%	
Country WER for Genders		37.46%	22.01%
Nigeria	M	43.30%	
	F		10.51%
	F		27.18%
	M	29.66%	
Country WER for Genders		36.48%	18.85%
Zimbabwe	F		29.55%
	F		16.26%
	M	23.37%	
	M	27.13%	
Country WER for Genders		25.25%	22.90%
Rwanda	F		22.42%
	F		34.68%
	M	71.48%	
	M	19.10%	
Country WER for Genders		45.29%	28.55%
Kenya	M	17.48%	
	F		33.20%
	F		38.52%
	M	25.83%	
Country WER for Genders		21.66%	35.86%

Figure 1.5

Google Cloud Platform WER for USA based on Gender

Country	Subject	AWER for Male Subject	AWER for Female Subject
USA	M	13.45%	
	F		14.36%
	F		6.02%
	M	21.44%	
Country WER for Genders		17.45%	10.19%

Figure 1.6

Total Google Cloud Platform WER for the study

WER for Google System (African Countries)	29.29%
Total WER for Google System (USA) inclusive	26.84%

Figure 1.7

Results of IBM WER for African Countries

Country	Subject	AWER for each Subject	AWER for each Country
Ghana	A	61.07%	
	B	40.87%	
	C	83.87%	
	D	71.18%	
Total WER for Ghana			64.25%
Nigeria	A	60.01%	
	B	74.92%	
	C	51.67%	
	D	73.25%	
Total WER for Nigeria			64.96%
Zimbabwe	A	70.27%	
	B	87.50%	
	C	57.55%	
	D	63.78%	
Total WER for Zimbabwe			69.78%
Rwanda	A	62.65%	
	B	88.41%	
	C	102.75%	
	D	50.31%	
Total WER for Rwanda			76.03%
Kenya	A	62.28%	
	B	69.37%	
	C	66.06%	
	D	63.59%	
Total WER for Kenya			65.32%

Figure 1.8

Result of IBM's WER for USA

Country	Subject	AWER for each Subject	AWER for Country
USA	A	17.01%	
	B	24.00%	
	C	19.02%	
	D	28.84%	
Total WER for USA			22.22%

Figure 1.9

Result of IBM's WER for African Countries based on Gender

Country	Gender	AWER for Male Subjects	AWER for Female Subjects
----------------	---------------	-------------------------------	---------------------------------

Ghana	F		61.07%
	F		40.87%
	M	83.87%	
	M	71.18%	
Country WER for Genders		77.52%	50.97%
Nigeria	M	60.01%	
	F		74.92%
	F		51.67%
	M	73.25%	
Country WER for Genders		66.63%	63.29%
Zimbabwe	F		70.27%
	F		87.50%
	M	57.55%	
	M	63.78%	
Country WER for Genders		60.67%	78.89%
Rwanda	F		62.65%
	F		88.41%
	M	102.75%	
	M	50.31%	
Country WER for Genders		76.53%	75.53%
Kenya	M	62.28%	
	F		69.37%
	F		66.06%
	M	63.59%	
Country WER for Genders		62.93%	67.71%

Figure 2.0

Result of IBM's WER for USA based on Gender

Country	Subject	AWER for Male Subject	AWER for Female Subject
USA	M	17.01%	
	F		24.00%
	F		19.02%
	M	28.84%	
Country WER for Genders		22.93%	21.51%

Figure 2.1

Total IBM WER for the Study

WER for IBM (African countries)	68.06%
WER for IBM (USA inclusive)	60.43%

Figure 2.2

Total WER based on Gender

Speech recognizer	Total Male WER for Speech recognizer	Total Female WER for Speech recognizer
Google	22.93%	21.51%

IBM (Watson)	61.20%	59.65%
Total Gender WER	45.90%	41.36%

Figure 2.3

4.3 Discussion.

The results obtained during the research has been highlighted in the above section. These results tend to answer the question posed at the beginning of the research which is “how does certain African accents affect English-speaking speech recognizers”. Thus, to delve into how these results answer this question, this section seeks to discuss the results acquired. The first set of results to be discussed is the Average WER of Google cloud platform for the African countries considered during the study. From the results obtained, the WER of Ghana is about 29.03 percent, that of Nigeria is 27.66 percent, the WER of Zimbabwe is 24.07 percent, that of Rwanda is 36.92 percent and finally Kenya’s WER is 28.76 percent. To adequately understand the meaning of WER, it can be recalled from previous chapters that, the lower the WER, the more effective or better the accuracy of transcription was. Thus, from this understanding, it can be said that, the country with the best average WER happens to be Zimbabwe. The two West African countries used in this research have quite impressive averages alongside Kenya. Rwanda, however, happens to be the only country to have an average WER of above 30 percent. This is due to the fact that the individual WER for subjects in Rwanda were quite high as compared to other participants from other countries. This is seen in the individual WER of Subject C who recorded a WER of 71.48%. The next set of results needed to analyze Figure 1.3 is Figure 1.5. The results in Figure 1.5 looks at the WER of the countries based on Gender. From the table (Figure1.5), it is quite evident that, apart from Kenya, the Female subjects of the other countries have fairly better WER than their male counterparts.

To accurately identify the impact that the African accent have on the used English-speaking speech recognizer, a comparison between the results conducted and the official word

error rates of the speech recognizers used would be done. Firstly, Google speech recognition technologies had achieved a 4.9 percent word error rate as of May 2017, and they stated that they transcribed every 20th word incorrectly [5]. However, Google's testing would have been done with a dataset full of American accents, thus, an experiment was conducted using the American accents available in the geographical location where the study was conducted. From Figure 1.4, it can be seen that, the WER for American accents was roughly 13.82 percent which is almost 7 times the WER recorded by Google. Also, based on the computation of the WER based on Gender, the female subjects once again tend to have better WER averages than their male counterparts. Now per the research conducted, the average WER of Google speech recognizer on the African countries considered was about 29.29 percent which is lower than the 4.9 percent recorded by the company. Also, comparing the average WER's of the countries that partook in this research, for Ghana, the WER was 29.03 percent which is relatively higher than that of the reported WER for Google. Nigeria records 27.66 percent, Zimbabwe, 24.07 percent, Rwanda, 36.92 percent and Kenya marks a WER of 28.76 percent. These are relatively higher WERs than what was recorded by Google. Also, adding the WER of the American dataset, the Average WER of Google which can be seen in Figure 1.7 is 26.84 percent. With the results presented, there are clear signs as to the claim made in the research that African accents do affect how English-speaking speech recognizers transcribe, is quite evident from the presentation of the results.

The next speech recognizer, IBM's Watson has a poorer WER or accuracy level as compared to Google's speech recognizer. the WERs for the African countries which are in Figure 1.8 are as follows; for Ghana, the WER is 62.25 percent, Nigeria records 64.96 percent, Zimbabwe 69.78 percent, Rwanda records 76.03 percent and Kenya 65.32 percent. The results of IBM's Watson show a balance in the average WER of each participant. The WER's are rather high thus proving to be very poor or inaccurate transcription made by the speech recognizer.

Also, in the results above, Rwanda has a very high WER average, and the participants under that stratum recorded extremely high individual WER averages. Such high averages include 102 percent and 88 percent recorded by some participants. The average WER of each country under this category all recorded about 60 percent WER. That is more than double the average WER marked from the transcripts made by Google. Also, with respects to the WER of the African countries based on Gender, the WERs are balanced amongst the countries considered based on gender. This can be seen in Figure 2.0.

With regards to IBM's Watson, the company reported that it had achieved an error rate of 5.5 percent as of March 2017. The recorded WER was a highly impressive milestone in the speech recognition industry at that time ^[12]. Once again, these tests would have been conducted on an all-American accent dataset and as such transcriptions with American accents were done. Per Figure 1.9, the average WER for America under IBM was about 22.22 percent which is much better than the WER of the African countries. Now, comparing what IBM reported to the data set that was present in this research, IBM's Watson or IBM's speech recognizer recorded a relatively poor WER. The average WER of IBM's Watson was 68.09 percent this is by far worse than the WER reported by the company as of March 2017. Even adding the American WER to the Average all-African WER, the average WER of IBM's Watson now becomes 60.43 percent. Relatively better than the all-African WER but very poor as compared to what the company reports.

There are numerous factors that can determine the accuracy of speech recognizers, some of these factors are uncontrollable while some are. An example of uncontrollable factors may include; the type of device used, the number of frequency filters of possibly the recorder, and also the bits/words in reference patterns to name a few. This are uncontrollable factors as such this could have largely affected the transcriptions made by the speech recognizers during this research. Also, controllable factors exist when determining the accuracy level seen in the

transcriptions made by the speech recognizers. Examples of controllable factors or variables include; the speaking rate, the number of speakers, the speaker skill, the dialect, the sex or gender and even the physical and psychology state of the speaker to name a few. These are factors that largely affect the performance of speech recognizers^[15]. Speaker accent and gender are the major controllable factors considered in this research. As such, the results are largely based on the speaker accent affecting the speech recognizer's accuracy as gender also plays a significant role. Thus, from the results especially the WER based on gender in Figure 2.3, it is clear that the female gender does relatively better as compared to the male gender. As such, since gender is also a factor that can affect the transcription of speech recognizers, one thing to point out is the fact that possibly since the male gender has a more deeper vocal tone as compared to the female gender, the speech recognizers may have found it difficult adequately understanding what the male participants said.

Finally, referring back to the results recorded, IBM records a much poorer WER as compared to Google. Reasons as to why IBM speech recognizer recorded very poor WER as compared to that of Google is not entirely understood, especially as the same data were used for the two speech recognizers. However, reasons as to why this is so could be as a result of the speaker variability or the number of speakers. Where Google speech recognizer probably was active in adequately listening and transcribing the data, IBM fell short. Though even with Google WERs that disparity between genders is evident though fairly close, it can be said that Google did a better job at transcribing as compared to IBM. It can also be said that; gender affects the how speech recognizers transcribe data. Also, and to answer the question being answered in this research, African accents do affect English-speaking speech recognizer accuracy.

Chapter 5: Conclusion/Future Work.

Speech recognition has become a vital component in the interaction of humans with various devices. Thus, as further stated, this research was conducted to, **to understand and assess to a considerable extent the impact of some African accents: Nigerian, Ghanaian, Zimbabwean, Rwandan and Kenyan, on English speaking Speech Recognition systems.** From the research conducted in this thesis and the result obtained, the factors as to why speech recognizers fail to transcribe data adequately is due to some factors. This research focused on accent being one of the elements as to how the transcriptions made by speech recognizers work. Therefore, to state the impact that African accents have on the English-speaking speech recognizers used, the WER of these speech recognizers as well as the participants suggest that the speech recognizers found it somewhat difficult to understand and adequately transcribe what each participant said. From sections 4.2.2-4.3, it was deduced based on the results and based on comparisons between the WERs of the English-speaking speech recognizers in this research and the WERs of the English-speaking speech recognizers stated by the creators of these recognizers that African accents do affect how English-speaking speech recognizers transcribe. Also, analyzing the WERs based on gender, gender also plays a role in speech recognition transcription. Thus, this research was able to answer the question posed at the beginning as well as meeting its objective.

Therefore, after achieving the objective of this study, more research needs to be done primarily in the dynamics of African accents and speech recognition. Also, using a larger dataset and possibly more African accents, this is a facet in the speech recognition field that would need more research. Also, there is need to look at probably creating or developing a speech recognizer(s) for African accents so as to restrict the inaccurate transcriptions made by English speaking speech recognizers. Another thing to consider is the need to try and understand other metrics that are available to check the accuracy of speech recognition systems.

Finding out or computing WER cannot just be the only largely acceptable metric or the most suitable metric for testing the accuracy of speech recognizers.

References

- [1] Ali, A.; Renals, S. Word Error Rate Estimation for Speech Recognition: E-WER. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*; Association for Computational Linguistics: Melbourne, Australia, 2018; pp 20–24.
- [2] Juang, B.-H.; Rabiner, L. R. *Automatic Speech Recognition - A Brief History of the Technology Development*; 2004.
- [3] Badenhausen, K. The World's Most Valuable Brands 2018: By The Numbers <https://www.forbes.com/sites/kurtbadenhausen/2018/05/23/the-worlds-most-valuable-brands-2018-by-the-numbers/> (accessed Mar 28, 2019).
- [4] Benzeghiba, M.; De Mori, R.; Deroo, O.; Dupont, S.; Erbes, T.; Jouviet, D.; Fissore, L.; Laface, P.; Mertins, A.; Ris, C.; et al. Automatic Speech Recognition and Speech Variability: A Review. *Speech Communication* **2007**, *49* (10–11), 763–786. <https://doi.org/10.1016/j.specom.2007.02.006>.
- [5] Google's Speech Recognition Technology Now Has a 4.9% Word Error Rate. *VentureBeat*, 2017.
- [6] Kamper, H.; Muamba Mukanya, F. J.; Niesler, T. Multi-Accent Acoustic Modelling of South African English. *Speech Communication* **2012**, *54* (6), 801–813. <https://doi.org/10.1016/j.specom.2012.01.008>.
- [7] Harris, J. 5 Reasons Why to Choose WAV Over MP3 and WMA. *Info | Remo Software*, 2017.
- [8] Houde, R. Prospects for Automatic Recognition of Speech. *American Annals of the Deaf* **1979**, *124* (5), 568–572.

- [9] Huang, C.; Chen, T.; Chang, E. Accent Issues in Large Vocabulary Continuous Speech Recognition. *International Journal of Speech Technology* **2004**, 7 (2), 141–153. <https://doi.org/10.1023/B:IJST.0000017014.52972.1d>.
- [10] Tao, C. H.; Huang, C.; Chen, T.; Li, S.; Chang, E.; Zhou, J. *Analysis of Speaker Variability*.
- [11] Fung, and P. Fast Accent Identification and Accented Speech Recognition. In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*; 1999; Vol. 1, pp 221–224 vol.1. <https://doi.org/10.1109/ICASSP.1999.758102>.
- [12] Reaching new records in speech recognition <https://www.ibm.com/blogs/watson/2017/03/reaching-new-records-in-speech-recognition/> (accessed Mar 31, 2019).
- [13] Teixeira, C.; Trancoso, I.; Serralheiro, A. Accent Identification. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*; 1996; Vol. 3, pp 1784–1787 vol.3. <https://doi.org/10.1109/ICSLP.1996.607975>.
- [14] Wells, J. C. Accents of English by J. C. Wells [/core/books/accents-of-english/A9495FCF7CC13D1D11FADB3DE2F335F6](https://www.corebooks.com/books/accents-of-english/A9495FCF7CC13D1D11FADB3DE2F335F6) (accessed Apr 22, 2019). <https://doi.org/10.1017/CBO9780511611766>.
- [15] What Causes Speech Recognizers to Make Mistakes? In *ICASSP '82. IEEE International Conference on Acoustics, Speech, and Signal Processing*; 1982; Vol. 7, pp 2030–2033. <https://doi.org/10.1109/ICASSP.1982.1171849>.
- [16] Yan, Q.; Vaseghi, S. A Comparative Analysis of UK and US English Accents in Recognition and Synthesis; 2002; Vol. 1, pp I–413. <https://doi.org/10.1109/ICASSP.2002.5743742>

Appendices

Appendix A

Below are the sentences used during the experiments which was conducted during this research.

- What route is better if I want to go to the mall?
- In my country buses are red, and the cars are green.
- The greatest gift is a passion for reading.
- Look at that lady over there, she has a purple eye.
- I like my bread with butter or jam and not mayonnaise.
- What play are we going to see in the theaters?
- What do you call it when rain falls while the sun shines?
- Donuts covered in Caramel are just the best.
- I need a white envelope its urgent.
- What is the largest telecommunication network in West Africa?

Figure 1.1

Appendix B

Google cloud platform

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Ghana	A	Female	What route is better if I want to go to the mall?	What's which is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and causing a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter and jam and what's my name
			What play are we going to see in the theaters?	What's play are we going to spend a few houses?
			What do you call it when rain falls while the sun shines?	What do you call it when the rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel, I just a bit
			I need a white envelope its urgent.	I need a white envelope the edges.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	B	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading
			Look at that lady over there, she has a purple eye.	Look at that lady whether she has a purple eye
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter and jam and not my name.

			What play are we going to see in the theaters?	What play are we going to see in the theater?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel I just the bed.
			I need a white envelope its urgent.	I need a white and loop is edging.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	C	Male	What route is better if I want to go to the mall?	What's wrong is better, I don't want to go to the mall
			In my country buses are red and cars are green.	In my county buses are red and cause a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for eating.
			Look at that lady over there, she has a purple eye.	Lucca Deli Dover de she has a Pitbull
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam or not my niece
			What play are we going to see in the theaters?	What's play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines?

			Donuts covered in Caramel are just the best.	Donuts covered in Caramel adjusted list
			I need a white envelope its urgent.	I need a white and blue legends.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country roses are red in casa green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Lucas valley urology
			I like my bread with butter or jam and not mayonnaise.	I like our bridge with bats our jam. I'm not my enemies.
			What play are we going to see in the theaters?	What's play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	When do you call it when it rains while the sun shines?
			Donuts covered in Caramel are just the best.	Do not cover in caramel. I just did this.
			I need a white envelope its urgent.	I need a White Sands Loop it surgeons

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication as WIC in West Africa?

Figure 1.2

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Nigeria	A	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	My country buses are red and cause a green
			The greatest gift is a passion for reading.	Greatest gift the passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not marinades
			What play are we going to see in the theaters?	What's play are we going to see the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when the rain falls while the sun shine?
			Donuts covered in Caramel are just the best.	Do not covered in Caramel at Joe's the best
			I need a white envelope its urgent.	I need to win enveloped its Origins

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	B	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are read and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there, she has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam or not my name.
			What play are we going to see in the theaters?	What's play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel at Joe's the best.
			I need a white envelope its urgent.	I need w white envelope sergeant.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	C	Female	What route is better if I want to go to the mall?	What service is better if I want to go to the mall?

			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Lucas that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter and jam and mayonnaise.
			What play are we going to see in the theaters?	What day are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when the rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel or just the best.
			I need a white envelope its urgent.	I need a words envelope for surgeons.
			What is the largest telecommunication network in West Africa?	What's the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country but is a red and cause a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mine is
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? When are we supposed sunshine?
			Donuts covered in Caramel are just the best.	Do nuts covered in caramel adjust the best.
			I need a white envelope its urgent.	I need a white envelope. Its urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?

Figure 1.3

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Zimbabwe	A	Female	What route is better if I want to go to the mall?	What road is the type of want to go to the mall?
			In my country buses are red and cars are green.	In my country, but he's already in Casa Grande.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread is bought out Jim and not my life.
			What play are we going to see in the theaters?	What movie are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donut covered in Chrome or just debris
			I need a white envelope its urgent.	I need a white envelope agent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	B	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and it caused a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not my name.

			What play are we going to see in the theaters?	What planet are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel with just the best
			I need a white envelope its urgent.	I need a white envelope it's a Junt.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	C	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red in color green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She is in People eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or Jim and not my own eyes.
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?

			Donuts covered in Caramel are just the best.	Ganache covered in caramel adjusted the best.
			I need a white envelope its urgent.	I need a white envelope. It's urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What road is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses that are 18 cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady of a day are the people I
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not my own eyes.
			What play are we going to see in the theaters?	What clay are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel, I just the best.
			I need a white envelope its urgent.	I need a white envelope is Agent.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
--	--	--	---	---

Figure 1.4

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Rwanda	A	Female	What route is better if I want to go to the mall?	What route if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cards are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not my nails
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when the rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel or just the best
			I need a white envelope its urgent.	I need a white envelope Archer

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	B	Female	What route is better if I want to go to the mall?	What's frogs is better if I want to go to the mall
			In my country buses are red and cars are green.	In my country versus a red and cause a green
			The greatest gift is a passion for reading.	The great excuse to the passion for reading
			Look at that lady over there, she has a purple eye.	Look up. Cynthia over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam in
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? When rain falls where the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel or just the bed
			I need a white envelope its urgent.	I need a white envelope. Its urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?

	C	Male	What route is better if I want to go to the mall?	What's Lucy's pizza if I want to go to the mall?
			In my country buses are red and cars are green.	My country buses are red and cars are green. Kaiser wins
			The greatest gift is a passion for reading.	Did you test the gifts in fashion for reading
			Look at that lady over there, she has a purple eye.	Look at that when you over there approval in.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam. I know its Matt Martinez.
			What play are we going to see in the theaters?	But where are we going to see any theaters
			What do you call it when rain falls while the sun shines?	What do Curry's when it rains five while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel adjust the best
			I need a white envelope its urgent.	I need the whites. I'm broke it is origin.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country has a red and cause a green.

			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	Like my bread with Beretta. Oh jam and not mayonnaise.
			What play are we going to see in the theaters?	What place are we going to see if in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? When rain falls while the sun shine?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel are just the best.
			I need a white envelope its urgent.	I need a white envelope is Argent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?

Figure 1.5

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Kenya	A	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country but it's a red and cars are green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	Look at that lady over there, she has a couple
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and what's my name?
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel are just the best
			I need a white envelope its urgent.	I need a white envelope with agent
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication maker in West Africa?
	B	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	My country versus the red and cause a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there, she has a couple art.
			I like my bread with butter or jam and not mayonnaise.	I like my bread is by salt or drum and wash my face.

			What play are we going to see in the theaters?	What play are we going to see in the catalyst?
			What do you call it when rain falls while the sun shines?	What do you call me the rain falls on the sunshine?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel at just the best
			I need a white envelope its urgent.	I need a light envelope is Genesis.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	C	Female	What route is better if I want to go to the mall?	What truck is better if I want to go to the mall?
			In my country buses are red and cars are green.	FEMA country buses and red in cars and green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there, she has a purple rain
			I like my bread with butter or jam and not mayonnaise.	Lick my belly button jam and not mayonnaise
			What play are we going to see in the theaters?	What's play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls for sunshine?

			Donuts covered in Caramel are just the best.	Connect covered in Caramel adjustable
			I need a white envelope its urgent.	I need a white envelope is urgent
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not my niece.
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? When rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel or just the best
			I need a white envelope its urgent.	I need a white envelope.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication Network in USA?

Figure 1.6

Country	Subject	Gender	Read Sentences	Transcribed Sentences
USA	A	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mayonnaise.
			What play are we going to see in the theaters?	What player are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel or just the best
			I need a white envelope its urgent.	I need a white envelope. Its urgent.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication maker in West Africa?
	B	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mayonnaise.
			What play are we going to see in the theaters?	What player are we going to see in the catalyst?
			What do you call it when rain falls while the sun shines?	What do you call it? When rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel or just the best
			I need a white envelope its urgent.	I need a white envelope. Its urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?

	C	Female	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mayonnaise.
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it? When rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel are just the best.
			I need a white envelope its urgent.	I need a white envelope with urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	D	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall?
			In my country buses are red and cars are green.	Am my country buses are red and cars are green

			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mayonnaise.
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in caramel adjust the best
			I need a white envelope its urgent.	I need a white envelope. Its urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication Network in USA?

Figure 1.7

Appendix C

IBM's Watson

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Ghana	A	Female	What route is better if I want to go to the mall?	What's new is is this a if I want to go to the mall.

			In my country buses are red and cars are green.	In my country buses are read in because it of green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at the leading over the shows the papal I'd.
			I like my bread with butter or jam and not mayonnaise.	Like my bread the bustle John and that's mine.
			What play are we going to see in the theaters?	Was the only good thing that the essence.
			What do you call it when rain falls while the sun shines?	What do you call it when they're in full as well the sunshine?
			Donuts covered in Caramel are just the best.	Donuts covered in Carmel adjusted this.
			I need a white envelope its urgent.	I need a life on you agents.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?
	B	Female	What route is better if I want to go to the mall?	What truth is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses I read in class green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for media.

			Look at that lady over there, she has a purple eye.	Look at that needed let me she has a purple I.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and that's my needs.
			What play are we going to see in the theaters?	Let's play are we going to see in the theater.
			What do you call it when rain falls while the sun shines?	What do you call it when reign fulls while the sun shines.
			Donuts covered in Caramel are just the best.	Doing that's covered in Carmel at just the best.
			I need a white envelope its urgent.	I need a white envelope is that Jim.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication that's wake in West Africa.
	C	Male	What route is better if I want to go to the mall?	With the roads is been tied phones glued to the more.
			In my country buses are red and cars are green.	In my county buses or anything because I.
			The greatest gift is a passion for reading.	Greece's gives some questions for you.
			Look at that lady over there, she has a purple eye.	You can only do Monday she has a purple or.
			I like my bread with butter or jam and not mayonnaise.	I like my brain but those John Lennon's means.

			What play are we going to see in the theaters?	Let's be are we going to see the up this.
			What do you call it when rain falls while the sun shines?	New call ins when removals while the sun shine.
			Donuts covered in Caramel are just the best.	Discovered in Carmel adjusted based.
			I need a white envelope its urgent.	I need a white and blue its engines.
			What is the largest telecommunication network in West Africa?	What's in the largest city communication it's working West Africa.
	D	Male	What route is better if I want to go to the mall?	With roots is but if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are rude and because of.
			The greatest gift is a passion for reading.	The greeters give this a question for you.
			Look at that lady over there, she has a purple eye.	Because of the need you on the she hasn't.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with but so John notes monies.
			What play are we going to see in the theaters?	What's plea only going to see in the defense's.
			What do you call it when rain falls while the sun shines?	What do you call it it's when it's means what the sunshine.

			Donuts covered in Caramel are just the best.	Do not score than in Carmel and just that this.
			I need a white envelope its urgent.	I need a way to end the decisions.
			What is the largest telecommunication network in West Africa?	What is the largest city communication as we can stuff.

Figure 1.8

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Nigeria	A	Male	What route is better if I want to go to the mall?	What's what's the better if I want to go to the mall.
			In my country buses are red and cars are green.	My country buses I read on because I green.
			The greatest gift is a passion for reading.	Good does give you the partial for.
			Look at that lady over there, she has a purple eye.	Look at the lady over there she has a poor boy.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam a not to my.
			What play are we going to see in the theaters?	What are we going to see in the theater.
			What do you call it when rain falls while the sun shines?	What do you call it when to reinforce what the sun shine.
			Donuts covered in Caramel are just the best.	Do not call waiting Carmona Joe's the best.
			I need a white envelope its urgent.	I need to wait include its origins.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa.
	B	Female	What route is better if I want to go to the mall?	What's the best route if I want to go to the mall?
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there, she has a purple eye.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not mayonnaise.
			What play are we going to see in the theaters?	What play are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines?
			Donuts covered in Caramel are just the best.	Donuts covered in Caramel are just the best.
			I need a white envelope its urgent.	I need a white envelope its urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa?

	C	Female	What route is better if I want to go to the mall?	What's yours is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are rude and cars and green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Because the the deal with the she has a purple are.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter and jam and that's Manny's.
			What play are we going to see in the theaters?	Was clearly going to see in the food service.
			What do you call it when rain falls while the sun shines?	What do you call it's when the rain falls while the sun shines.
			Donuts covered in Caramel are just the best.	Doing it's covered in Carmel or just the best.
			I need a white envelope its urgent.	I need a word. Urgent.
			What is the largest telecommunication network in West Africa?	What was the largest telecommunications work in West Africa.
	D	Male	What route is better if I want to go to the mall?	What's Ruth is but that if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are rated on because agree.

			The greatest gift is a passion for reading.	The greatest gift is a partial for reading.
			Look at that lady over there, she has a purple eye.	Look at the lady who was a she has the pope will buy.
			I like my bread with butter or jam and not mayonnaise.	I like my braid we booked up would jump on notes in my knees.
			What play are we going to see in the theaters?	What are we going to see in the theater.
			What do you call it when rain falls while the sun shines?	What do you call it square root foes but the sun shine.
			Donuts covered in Caramel are just the best.	Do not covered in Carmel are just the best.
			I need a white envelope its urgent.	I need a white and blue it's urgent.
			What is the largest telecommunication network in West Africa?	What is the largest in the communication network in West Africa.

Figure 1.9

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Zimbabwe	A	Female	What route is better if I want to go to the mall?	Watch movies but it's like going to school to the mall.
			In my country buses are red and cars are green.	In my country but is there anything because the green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	Look at the lady but she has a lot.
			I like my bread with butter or jam and not mayonnaise.	I like my bread and butter jam and not mine.
			What play are we going to see in the theaters?	Let me going to see in the.
			What do you call it when rain falls while the sun shines?	What do you call the twin read full for the function.
			Donuts covered in Caramel are just the best.	Donna's governing.
			I need a white envelope its urgent.	I need a white. Agent.
			What is the largest telecommunication network in West Africa?	What is the like this to the communication network in West Africa.
	B	Female	What route is better if I want to go to the mall?	One who is been to and if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses and read and causing.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	You can maybe whose name she has a formal I.
			I like my bread with butter or jam and not mayonnaise.	Thank member in his buttonhole jammed and knows my needs.

			What play are we going to see in the theaters?	And we're going to see in the theater.
			What do you call it when rain falls while the sun shines?	Would you be willing to in reading falls well this and James.
			Donuts covered in Caramel are just the best.	June is committing caramel just based.
			I need a white envelope its urgent.	I need to win and lose its engine to.
			What is the largest telecommunication network in West Africa?	What's his images to the communication and what can we expect god.
	C	Male	What route is better if I want to go to the mall?	What troops is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses. And cause a green.
			The greatest gift is a passion for reading.	The greatest gift is shown for reading.
			Look at that lady over there, she has a purple eye.	Look in that lady over there she is if people I.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam into my own 9.
			What play are we going to see in the theaters?	What are we going to see in the theater.
			What do you call it when rain falls while the sun shines?	What do you call in sick when rain falls while the sun shines.

			Donuts covered in Caramel are just the best.	Do not provide interim and adjusted the best.
			I need a white envelope its urgent.	I need to go white envelope its agents.
			What is the largest telecommunication network in West Africa?	What is the longest telecommunication network in West Africa.
	D	Male	What route is better if I want to go to the mall?	What should this be to if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses that aid in class I agree.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading
			Look at that lady over there, she has a purple eye.	Look at that lady but they hear the purple I.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and not to buy a nice.
			What play are we going to see in the theaters?	What play we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you call a 2 in the rain falls where the sun shines
			Donuts covered in Caramel are just the best.	Donuts covered in could I mean I guess the best.
			I need a white envelope its urgent.	I need to wait in lobbies agent.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication it's sick in West Africa.

Figure 2.0

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Rwanda	A	Female	What route is better if I want to go to the mall?	What routes he's decide if I want to go to the.
			In my country buses are red and cars are green.	In my country but since I read it cards agree.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that maybe there she has a couple.
			I like my bread with butter or jam and not mayonnaise.	Unlike my bread with butter or jam in the maintenance.
			What play are we going to see in the theaters?	What are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you call it when reign schools where the sun shines.
			Donuts covered in Caramel are just the best.	And it's covered in Carmel or just the best.
			I need a white envelope its urgent.	I need a white envelope arching.

			What is the largest telecommunication network in West Africa?	Which is the largest of communication networks in West Africa.
	B	Female	What route is better if I want to go to the mall?	What's wrong is been safer once because of the mold.
			In my country buses are red and cars are green.	America entry roses are red and cars are green.
			The greatest gift is a passion for reading.	The richest uses a passion for anything.
			Look at that lady over there, she has a purple eye.	Look at that they deal with it she has our.
			I like my bread with butter or jam and not mayonnaise.	And they were bred with John and that's minus.
			What play are we going to see in the theaters?	We'll plan on seeing the kids and.
			What do you call it when rain falls while the sun shines?	With new colleagues when when fulls with the sunshine.
			Donuts covered in Caramel are just the best.	Gone is covered in Carmel unchanged.
			I need a white envelope its urgent.	I need to wait on hold it's 18.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication this work in West Africa.

	C	Male	What route is better if I want to go to the mall?	Let's move his big time if I want to put them all.
			In my country buses are red and cars are green.	Muskoxen buses and read on the kinds of green.
			The greatest gift is a passion for reading.	We tend to give to the partial for reading.
			Look at that lady over there, she has a purple eye.	Look I've done we do over there. She's Hans a coupon in.
			I like my bread with butter or jam and not mayonnaise.	Unlike my brain B. but some old John I know it's much maintenance.
			What play are we going to see in the theaters?	What's been going to see an anti itself.
			What do you call it when rain falls while the sun shines?	Once you Cory's when rains files when a sunshine.
			Donuts covered in Caramel are just the best.	No no it's covered in economy adjusted this.
			I need a white envelope its urgent.	I need someone to come through destroyed in.
			What is the largest telecommunication network in West Africa?	What is the noun is the communication it's working with stuff we can.
	D	Male	What route is better if I want to go to the mall?	What route is beta if I want to look.

			In my country buses are red and cars are green.	In my country basses irate and cause a green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading
			Look at that lady over there, she has a purple eye.	Look at that lady. Over there. She has a purple and.
			I like my bread with butter or jam and not mayonnaise.	I am like my parade with button. Town and not mine is.
			What play are we going to see in the theaters?	What are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you quality when the rain falls well the sun shines.
			Donuts covered in Caramel are just the best.	Donna covered in Karama. Our chest the best.
			I need a white envelope its urgent.	I need a quiet and about its agent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa.

Figure 2.1

Country	Subject	Gender	Read Sentences	Transcribed Sentences
Kenya	A	Male	What route is better if I want to go to the mall?	What is better if I want to go to the more.

			In my country buses are red and cars are green.	In my country buses that raid on Gaza and green.
			The greatest gift is a passion for reading.	The greatest gift is not Boston for reading.
			Look at that lady over there, she has a purple eye.	Look and that lady over there she has a couple.
			I like my bread with butter or jam and not mayonnaise.	I let me bring it back to John and I'm with my knees.
			What play are we going to see in the theaters?	What are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you do in rainforest when the sanctions.
			Donuts covered in Caramel are just the best.	Donuts government gone in and just the best.
			I need a white envelope its urgent.	I need a right and look at it.
			What is the largest telecommunication network in West Africa?	What is the latest telecommunication network in this difficult.
	B	Femal e	What route is better if I want to go to the mall?	What is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country boxes the red and because we.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.

			Look at that lady over there, she has a purple eye.	No balcony over about she has a problem.
			I like my bread with butter or jam and not mayonnaise.	I like my bread is impossible John and what's more he's.
			What play are we going to see in the theaters?	What are we going to see in the status.
			What do you call it when rain falls while the sun shines?	What do you call me to the rain falls on the sunshine.
			Donuts covered in Caramel are just the best.	Covered in Carmel of chests and backs.
			I need a white envelope its urgent.	I need a large envelope Arjun's.
			What is the largest telecommunication network in West Africa?	What is the largest in the communication aspect in this topic on.
	C	Female	What route is better if I want to go to the mall?	What to do his best to if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses and rate and because and green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at the lady over there she has a couple.
			I like my bread with butter or jam and not mayonnaise.	I like my body but don't jump in the to my needs.

			What play are we going to see in the theaters?	Which mail we going to see if you had to.
			What do you call it when rain falls while the sun shines?	What do you call the tune rain falls for this time trying.
			Donuts covered in Caramel are just the best.	The nets covered income and I just.
			I need a white envelope its urgent.	And you know what envelope I jeans.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication it's working on stuff will come.
	D	Male	What route is better if I want to go to the mall?	Is that if I want to go to the more.
			In my country buses are red and cars are green.	In my country buses are red and because green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a problem with.
			I like my bread with butter or jam and not mayonnaise.	Hey like my bread with but and I know to my knees.
			What play are we going to see in the theaters?	What are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What's the you would need to in rain forests where the sun shines.

			Donuts covered in Caramel are just the best.	Donuts covered the condiment adjusted based.
			I need a white envelope its urgent.	They need to avoid envelope I didn't.
			What is the largest telecommunication network in West Africa?	What is the largest in the communication and talking to Mr.

Figure 2.2

Country	Subject	Gender	Read Sentences	Transcribed Sentences
USA	A	Male	What route is better if I want to go to the mall?	What route is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are red and cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there. She has a purple I.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam and non mayonnaise.
			What play are we going to see in the theaters?	What player are we going to see in the theaters?
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines.

			Donuts covered in Caramel are just the best.	Donuts covered in Caramel or just the best
			I need a white envelope its urgent.	I need a white envelope it's urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication maker in West Africa.
	B	Female	What route is better if I want to go to the mall?	Well it is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are read and colors are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purple I.
			I like my bread with butter or jam and not mayonnaise.	And like my bread with butter or jam and nominees.
			What play are we going to see in the theaters?	What player are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls on the sun shines.
			Donuts covered in Caramel are just the best.	Donuts converting caramel or just the.
			I need a white envelope its urgent.	I need a white envelope it's urgent.

			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa.
	C	Female	What route is better if I want to go to the mall?	What ground is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are red cars are green.
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purpose.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter or jam in not meaning.
			What play are we going to see in the theaters?	What played are we going to see in the theaters.
			What do you call it when rain falls while the sun shines?	What do you call it when rain falls while the sun shines.
			Donuts covered in Caramel are just the best.	Donuts governing care mail are just the best.
			I need a white envelope its urgent.	I need a white envelope it's urgent.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication network in West Africa.

	D	Male	What route is better if I want to go to the mall?	What we're it is better if I want to go to the mall.
			In my country buses are red and cars are green.	In my country buses are red and cars are green
			The greatest gift is a passion for reading.	The greatest gift is a passion for reading.
			Look at that lady over there, she has a purple eye.	Look at that lady over there she has a purple high.
			I like my bread with butter or jam and not mayonnaise.	I like my bread with butter and jam and not man is.
			What play are we going to see in the theaters?	What player are we going to see in the theater.
			What do you call it when rain falls while the sun shines?	What do you call it when the rain falls while the sun shines.
			Donuts covered in Caramel are just the best.	Donuts covered in caramel it just the best.
			I need a white envelope its urgent.	I need a white envelope Spurgeon.
			What is the largest telecommunication network in West Africa?	What is the largest telecommunication Network in Africa.

Figure 2.2