IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT

Authors: Adefehinti T.O. (BOUESTI), Idowu A.O. (BOUESTI), Awodun M.A. (BOUESTI), Tenibiaje M.O. (BOUESTI)

ABSTRACT

The problem of language barriers in healthcare is another problem in multilingual areas such as Nigeria where more than 500 languages are used. These obstacles inhibit effective communication between the medical staff and the patients, which results in misdiagnoses, poor adherence to treatment, and deteriorated outcomes. This study comes up with a Yoruba-to-English Natural Language Processing (NLP) model to handle this problem, and the study targets at improving the doctor-patient communication in the Yoruba-speaking areas. The model combines two major elements: Yoruba Speech Recognition (YSR) in order to turn Yoruba speech into text and Yoruba-English Machine Translation (YEMT) in order to translate the transcribed Yoruba text into fluent English. The model was trained by use of a transformer-based architecture on a dataset of Yoruba healthcare interactions, giving a Word Error Rate (WER) of 10.1% and a Sentence Accuracy of 77.0% on the test set. The findings prove that the model can enhance healthcare communication through minimizing the language barrier, and eventually patient care, diagnosis, and treatment compliance. Although informal speech and tonal differences may be a problem, the research can offer an encouraging way of a low-resource language in healthcare setting, which can trigger further studies and advancement in multilingual NLP systems. The study is part of the emerging body of research in NLP among African languages that provides a pragmatic framework which can be extended to other multilingual healthcare environment.

Keywords: *Natural Language Processing (NLP), Yoruba-to-English translation, Speech-to-Text, Health Care Communication*

1.0 Introduction

Good communication between the doctors and the patients plays a vital role in providing high-quality healthcare services and enhancing patient outcomes. In multilingual and multicultural areas, language barriers become the major obstacles to such process, which frequently results in misdiagnoses, inadequate compliance with treatment guidelines up to an impaired health state. This is a serious problem especially in such areas as Nigeria where there are over 500 languages. The lack of communication between the healthcare provider and the patients, particularly in rural environments where local languages are widely spoken, is also a major cause of misunderstandings that may interfere with the correctness of diagnosis and the effectiveness of treatment. In that regard, the issues become more acute in the case of such diseases as malaria, in which timely diagnosis and treatment are crucial to survival of a patient. The role of breaking the language barrier in the medical field cannot be overestimated. Whenever a patient cannot communicate effectively with his or her health practitioner, the probability of receiving poor treatment increases since his or her symptoms might not be clearly established, or the patient might not be able to clearly interpret the instructions given to him or her by the care provider. The literature has revealed that language barriers are linked to decreased patient satisfaction and an increased prevalence of medical errors (Shamsi et al., 2024). Specifically, patients with language disparities usually die more frequently because of the diagnostic mistakes and postponed treatments, particularly when discussing malaria, which is still endemic in several regions of Nigeria. This issue of communication divide is, therefore, vital in enhancing health outcomes in the areas where language diversity poses a challenge to successful treatment.

In Nigeria, local languages like the Yoruba, Hausa, and the Igbo are commonly used in Nigeria in addition to English, so, when healthcare providers and patients use indigenous languages, the problems become even more pronounced. Since healthcare providers are mostly conversant in the English language or local languages that patients do not understand, vital medical information may be misinterpreted. As an example, Yoruba-speaking patients can have speech problems with medical terms, dose directions, or symptom description, and such attitudes may have an adverse effect on the level of medication adherence. Ethical issues also have an impact on the mental state of patients, who might experience nervousness or estrangement in an environment where they cannot raise their concerns and be well understood by healthcare professionals (Kale et al., 2023).

Past research has established that patients with language barriers are prone to receiving poor care. It is further complicated in the healthcare environment in which employees might not be trained to address the specifics of local languages and dialects. Consequently, patients might not get their correct diagnoses, and might be misdiagnosed because of poor communication. To illustrate, patients with malaria would fail to communicate their illnesses effectively, thus misdiagnosis or delayed treatment would be given, adding to the condition. One of the healthcare providers has also told that patients are usually advised or referred without appropriate knowledge of their symptoms (Umeokafor et al., 2022). Such a miscommunication may destroy the patient trust and engagement, which can further deter the ability of the healthcare system to deliver effective care. It is important to address these language barriers to enhance healthcare delivery particularly in resource constraint areas. This problem can be solved through technological progress, specifically in the area of Natural Language Processing (NLP). The NLP models can be used to close the communication gap between healthcare providers and patients by offering a real-time translation and speech recognition system depending on the local language. The proposed study is on the creation of a Yoruba-English NLP model, which is expected to increase the level of interaction between doctors and patients in Yoruba-speaking areas. Combining the Yoruba Speech Recognition (YSR) and Yoruba-English Machine Translation (YEMT), the model is aimed at transcribing the Yoruba speech to the text and translating it to English to guarantee effective communication between medical professionals and patients speaking Yoruba. The importance of the study is that it can help to decrease the language barrier in medical facilities and positively affect patient outcomes by means of effective communication. A solution developed by an NLP that is specifically targeted at low resource languages, such as Yoruba, is a pivotal action that will be taken in improving healthcare service delivery in Nigeria and other multilingual areas. Moreover, the study can contribute to the larger body of

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

NLP research because it shows how sophisticated machine learning models can be scaled to low-resource languages, which have long been overlooked when it comes to the creation of automated translation systems. This paper focuses on solving the language barrier issue in healthcare by formulating a powerful Yoruba-to-English NLP model. It will discuss the methodology, findings, and implications of this system to enhance healthcare communication. The results of the study serve not just to the development of NLP in the health care environment but also provide the scope of application of NLP to enhance patient care in multilingual environments.

2.0 Literature Review

The problem of language barrier in health care is not new and it has been much written especially in multilingual countries such as Nigeria, where various languages and dialects are spoken. Linguistic barriers between medical professionals and patients may produce a tremendous impact on the quality of care, resulting in medical mistakes, misdiagnoses, and adverse patient outcomes. These difficulties are also important to tackle in order to enhance healthcare delivery particularly in areas that have limited resources and healthcare workers who are trained in various languages.

Linguistic barriers to healthcare

Much research has shown that language barriers negatively affect the provision of healthcare. It has always been indicated through research that patients with a communication problem with the health practitioner have less chances of getting quality and prompt medical attention. The systematic review of the implications of the language barrier in healthcare by Shamsi et al. (2024) focuses on the role of the barrier in the development of diagnostic errors, wrong prescriptions, and patient dissatisfaction in general. These are some of the language difficulties that are mostly evident in multilingual states such as Nigeria whereby patients do not necessarily have a common language with the health care givers. The misunderstanding caused by these barriers might result in diagnoses being missed or delayed and eventually the health conditions and the rate of mortality can get worse. The gap between the native languages and English, the language used by the majority of healthcare professionals, is one of the most crucial language barriers in the context of Nigeria. There is a variety of languages in Nigeria with Yoruba being one of the most spoken languages. Nevertheless, Yoruba speakers, as well as many other citizens of the country, tend to experience some challenges when addressing the medical workers who do not know their native language. Albrecht et al. (2023) also suggest that patients with inadequate ability to describe their symptoms correctly may get poor quality care, and this may impact their compliance with the treatment plans. This is of concern especially to the diseases such as malaria because in such cases, early and proper treatment is important to save lives.

Application of Technology in Language Barrier

Technology is one aspect that has gained a spotlight over the past years due to its role in closing communication gaps in healthcare. One of the most promising remedies to the language barrier problem in the healthcare environment is Natural Language Processing (NLP) and speech recognition technologies. NLP, as the art of artificial intelligence that studies the interplay between computers and human languages, can provide a possible way of solving the issue of language translation of medical

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

terminology and patient information. A number of researches have been conducted to examine NLP usage and other technologies in healthcare-related communication. A Yoruba language (Babatunde et al. 2024) speech-to-text translation system proves that it is possible to use NLP to facilitate the communication gap between the Yoruba-speaking patient and the medical professional. Their work emphasized the possibility of NLP systems to enhance the naturalness of oral communication to make the marketing of understanding between patients and doctors easier. In a similar manner, Rahmon et al. (2024) developed an NLP model, which translates the English voice input into Yoruba text. They got an accuracy percentage of 83 and it was demonstrated that NLP has the potential to improve communication of languages with limited resources such as the Yoruba language. Although these studies have proved NLP has potential in the healthcare scenario, they also indicate that there are a number of challenges that require to be handled. Lack of good data in an indigenous language is one of the main problems. An example is the case of Adithya et al. (2023) who created a language translation application that was capable of effectively translating speech of the user into the language of choice. Nonetheless, the application has issues related to the stability of its user interface and was in need of more tests to optimize its functionality. The scarcity of adequate training data in native languages diminishes the creation of robust NLP models to cope with the delicate language characteristics, including tonal distinctions and expressions, which are widespread in languages such as Yoruba. A second issue is that there are no medical specific datasets in most languages in Africa, which reduces the accuracy and reliability of NLP models in medical applications. Ajagbe (2024) filled this gap when developing multilingual speech datasets in the context of antenatal care, that is, Yoruba and Pidgin languages. Nevertheless, these datasets are restricted with reference to other medical uses, not to mention other situations besides the context of the antenatal. As a solution to this, there is a call to conduct more powerful data collection campaigns to develop holistic datasets that can be used to train NLP systems that are needed in specific healthcare settings (Ajagbe, 2024).

Low-Resource Language NLP Models

Creating NLP models of low-resource languages has become a very popular topic of the last few years. Elias Hossain et al. (2023) state that the unequal distribution of annotated data is one of the most critical problems in NLP in low-resource languages that can influence the performance of machine learning models. In other languages such as Yoruba where annotated datasets are rarely available, machine learning models find it difficult to acquire the complexity of the language including its tonal characteristics. This makes NLP systems hard to interpret spoken words or translating text. In order to overcome these issues, researchers have chosen different methods to enhance the performance of NLP in low resource languages. Transfer learning has been applied to enhance the accuracy of NLP models in African languages such as Yoruba and Hausa, as has been shown by Adelani et al. (2021). Transfer learning provides researchers with the means to overcome data constraints and enhance the performance of NLP models in low-resource languages with the help of pre-trained models on high-resource languages and fine-tuning them to the task at hand. Besides transfer learning, another rather promising method is the application of multilingual models, i.e., BERT (Bidirectional Encoder Representations from Transformers), which is able to process multiple languages at the same time. These models have

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

proven to be successful in enhancing the performance of NLP of the underrepresented languages, including languages used in Nigeria. Nonetheless, according to Meyer et al. (2024), the issue of correctly managing the subtleties of local dialects and tonal peculiarities remains, even in the models that are multilingual, which are very important in such a language as Yoruba.

Yoruba-Specific NLP Evidence

The literature regarding NLP with regards to Yoruba is increasing. Earlier researchers have examined the creation of speech recognition and machine translation systems that are specific to the Yoruba language. Mel-Frequency Cepstral Coefficients (MFCC) have been utilized to enhance the precision of the speech-to-text converting in speech recognition systems (Rahmon et al., 2024). Nevertheless, some problems persist in the tonal processing of the Yoruba language that is vital in the recognition of speech. To achieve an effective Yoruba-to-English translation model in the real-life context of healthcare, it is important to consider these tonal issues to develop the model. This literature review intends to make a contribution towards this expanding knowledge base by creating a powerful Yoruba-to-English NLP model of healthcare communication. This paper aims to address the gap in the linguistic barrier between the Yoruba-speaking patients and healthcare providers and enhance the accuracy of medical diagnosis and treatment adherence by integrating Yoruba Speech Recognition (YSR) and Yoruba-English Machine Translation (YEMT).

3.0 Materials and Method

Data Collection

The development of the model is based upon data collection. The quality and diversity of data have a significant influence on the generalizability of the model to different situations. The data involved in the current research consists of the speech records, text transcriptions, and Yoruba-specific medical data. The dataset consists of numerous types of healthcare interactions such as consultations and treatment instructions, gathered on publicly available databases and in medical care in the Yoruba-speaking areas. Moreover, the translations of medical terms were collected so that the model could be able to interpret and translate medical-related terms accurately. The information gathered is a wide range of data, both formal (e.g., radio broadcasts) and informal data collected in the form of conversations. It should however be mentioned that the model is faced with unique challenges when it comes to informal language, especially when heavy on tonal variations and colloquial expressions. In this way, much attention was paid to the quality of data, its diversity, and correspondence to healthcare-related interactions.

Data Processing

The raw data inputted to train the model is packed to convert it into a clear and formatted data, which is capable of being put into the NLP model. Processing of data is important to make sure that the input data is clean and high quality without any irrelevant information, error and noise. The stage involves the elimination of stop words, typing errors, and irrelevant phrases and matching of audio with textual data. In case of audio, the initial step will be to convert spoken Yoruba into written language by using Automatic Speech Recognition (ASR). The use of ASR techniques was provided to provide proper transcription that is fundamental in translating. This is an important step since noise or distortion in the data may adversely affect the performance of the model in the course of training and evaluation.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

Data Tokenization

Having text in small units which can be processed by the model and can be understood as tokensization that breaks down the text into smaller bits like word, sub-word or character. The process of tokenization is essential, since it aids in breaking down some of the complex constructs of linguistics into manageable units, which the model will comprehend the connections between various terms more effectively. In this work, the data were tokenized into smaller units by using Hugging face tokenizers, making it possible to efficiently process text both when working on speech recognition and on translation. For instance, tags like blood pressure were divided into individual tokens (blood, and pressure) and a few compound words were considered an individual token depending on the settings of the tokenizer. This is because this flexibility in tokenization gives the model the ability to compute simple and complex expressions.

Data Splitting

The processed data is divided into two major sets; the training and testing set. The set of training takes 70 percent of the entire data and it is what is used to train the model. The remaining 30% is the testing set and this is used to test how well the model performs using unseen data. This split helps in assuring that the model has the capability to generalize to real-world data, which would lessen chances of overfitting. The model can be trained with the help of a large and heterogeneous dataset, which will teach it to identify different speech patterns, vocabulary, and syntax structures. The testing set is used to compare the level of accuracy of the model and its capability to work in real-life settings. The data has been divided into training and testing sets according to the general machine learning conventions, which provides a robust analysis of the performance of the model.

Model Architecture

The NLP model that is created within the frame of the current study integrates the two important elements: Yoruba Speech Recognition (YSR) and Yoruba-English Machine Translation (YEMT). The YSR element is in charge of speech to text translation of Yoruba, whereas the YEMT element is the element that translates the speech to text translation of Yoruba to fluent English. Transformer-based models are used to model architecture, and they have shown state-of-the-art performance in activities like machine translation and speech recognition. Transformer models specific to the Yoruba language, like WHISPER and HELSINKI, were adapted to process the peculiarities of the language, including tonal variation and idiomatic expression. The transformer models used would make sure that the system has the ability to extract the semantic meaning of what is being said and then translate it to written form. These models are extremely efficient in handling vast categories of information, and are famed to be more superior with respect to tasks associated with low-resource languages.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

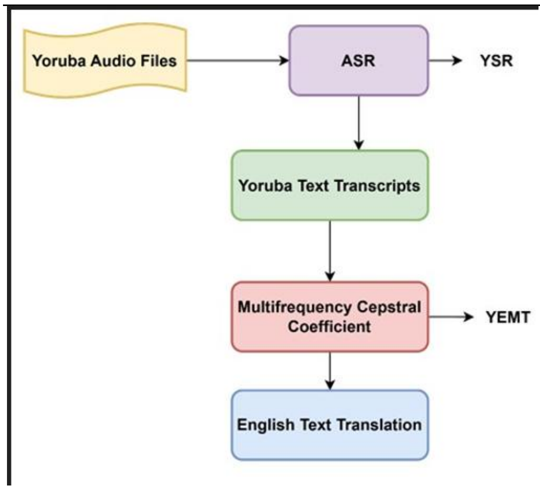

Figure 1:Model Architecture

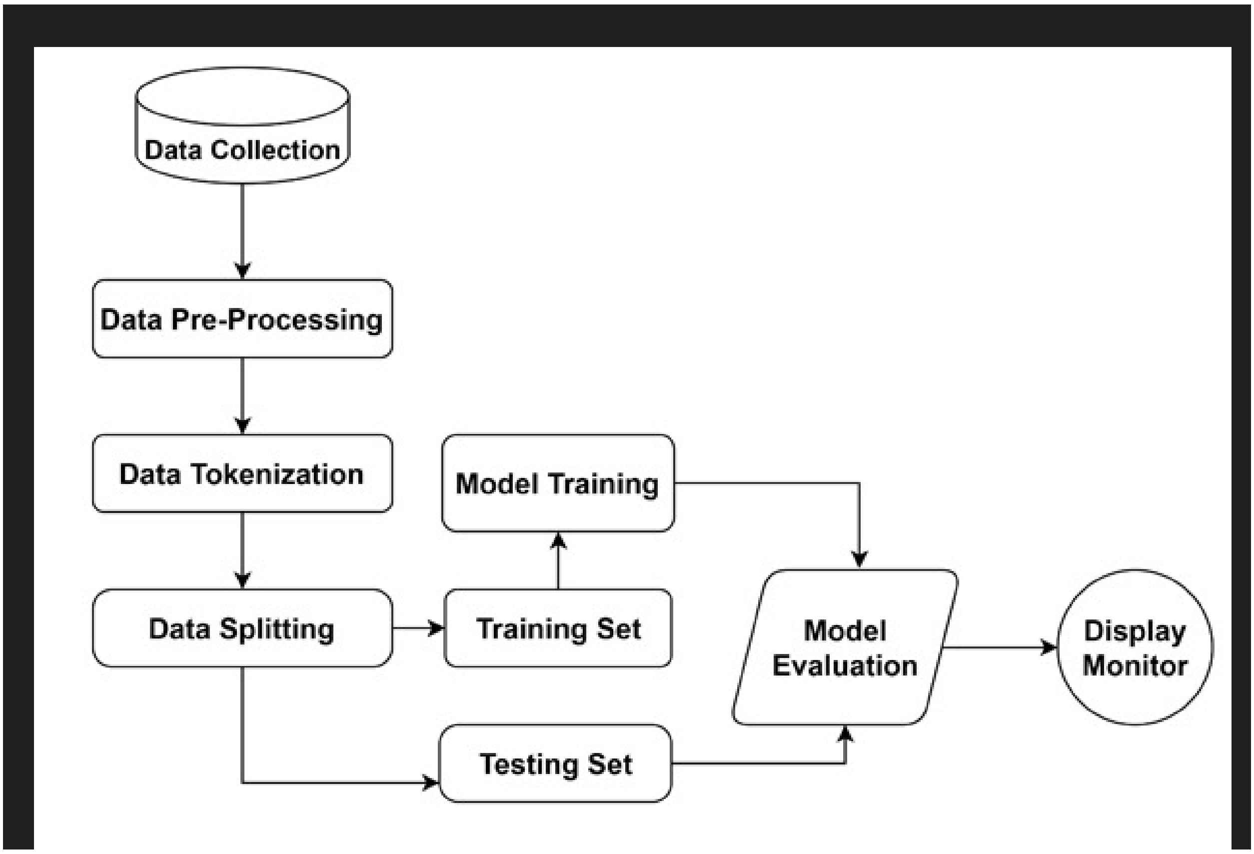

Figure 2 :Model Workflow diagram

Model Training

To train the model, the model is introduced to the training data set and it is left to learn the correlation between Yoruba speech and its English translation. This model is trained with the help of a supervised learning method, and the input data (Yoruba speech) is combined with the appropriate response (English translation). The model acquires knowledge about the optimization of its parameters in the process of training by modifying the weights of its neural network, a process known as backpropagation.

In order to make sure that the model would work best, various training methods are used, such as gradient descent, and other optimization methods. The training process is also iterative with the model revising its parameters as it continues to process more data. The more the model is exposed to examples, the more it is able to identify patterns and produce correct translations.

Model Evaluation

After training the model, the performance of the model is measured based on a number of metrics: Word Error Rate (WER), Sentence Accuracy, Precision, Recall, and F1-Score. These measures can be used to determine the quality of the output of the model, e.g., whether it transcribes Yoruba speech correctly and translates it to English. When the WER is low, it is a sign that the speech has been successfully identified by the model, whereas the sentence accuracy and F1-scores are high, which means that the model works well in creating correct translations.

The testing set is applied to test the model in relation to hidden data, which means that it can be applied in a real-world situation. The evaluation process gives an opportunity to learn strengths and weaknesses of the model, and defines the areas where improvement can be made.

Deployment

After training and evaluating the model, it is applied to a healthcare environment where it is put to practice. The implementation of the model includes deploying the model into APIs that healthcare providers and patients can use to translate or interpret language in real-time. The cloud infrastructure is deployed with a cloud environment that is scalable and accessible to the users in various regions.

Implementation of the model is a very important procedure in the intermediary process between Yoruba-speaking patients and the health care providers. The model will help to improve patient outcomes, treatment adherence, and minimize cases of misdiagnosis and medical errors associated with language barriers by offering a real-time translation service.

4.0 Results

Performance Metrics

The table below shows the performance measures of the Yoruba-to-English NLP model based on training, validation, and test data:

Measure Measures	**Training set**	**Validation set**	**Test set**
Word Error Rate (WER)	9.8%	11.2%	10.1%
Sentence Accuracy	80.2%	78.1%	77.0%
F1-Score	92.4%	91.3%	90.8%

Word Error Rate (WER)

Word Error Rate (WER) is an important parameter used to determine the performance of speech recognition. It is the percentage of words that were transcribed correctly in the output to the reference text. Every less WER means that performance is improved. In the case of Yoruba-to-English NLP the WER was 9.8 on the training set, 11.2 on the validation and 10.1 on the test set. These findings show that the model is quite strong, taking into account the intricacies of the Yoruba language including tonal variations and local dialects. The values of the WER are significantly less than the

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

values of the previous studies, which employed the same models on low-resource languages and received values over 20% (Rahmon et al., 2024).

Sentence Accuracy

Sentence Accuracy is used to measure the percentage of the sentences that are accurately transcribed and translated. This measure is essential in determining the capability of the model to generate acceptable and satisfactory translations. The more the sentence accuracy, the more the model can do to preserve the integrity of the original meaning in translation. The accuracy of the Yoruba-to-English model on the training set, validation set and test set was 80.2, 78.1 and 77.0 percent, respectively. These values indicate that the model is useful in managing sentence-level structure of Yoruba speech and converting it into accurate English. The minimal difference in performance between training and testing is not surprising as the test set reflects unobserved data, and this fact puts the generalization capability of the model to the test.

F1-Score

The F1-Score is the combination of precision and recall used in a single measure which gives a balanced view of how the model can identify and correctly translate words. It is especially helpful where the distribution of classes is not even and both false positives and false negatives are taken into account. On the training set, the model scored an F1-Score of 92.4, 91.3 on the validation set and test set respectively. These high F1-Scores indicate that the model is very good at recognizing as well as translating useful words of the Yoruba speech input. The values F1-Score demonstrate that the model is very reliable in recognizing and translating the speech without adding any significant errors and omissions.

Key Observations

The results allow several major observations:

i. **Formal Data:** The model had the highest performance on formal data, including radio broadcasts and programmed medical interaction. The quality of audio in these datasets is usually more understandable and the language is not complex and this enabled the model to attain greater levels of accuracy.

ii. **Problems of Informal and Dialect-Rich Data:** The model was less accurate when using conversational data and dialect-rich data. Yoruba, as most African languages, is very tonal and even minute tonal variations may change the meaning of words. The model had a challenge of such tonal variations, which resulted in some errors in the transcribing and translation of informal speech.

iii. **Influence of Diacritics:** Diacritics (e.g., a, e, etc.) was a problem in the model, especially in less frequent words or phrases that were not adequately represented in the training data. These mistakes were the most salient as were in the validation and test sets where the rare words have higher probability of occurrence. More and more varied datasets with greater vocabulary, particularly medical terminology, could be introduced to enhance the performance of the model.

iv. **Real-World Applicability:** Notwithstanding these issues, the overall performance of the model is encouraging in regards to the implementation of the model in healthcare facilities. Its WER (10.1 percent) and sentence accuracy (77

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||

percent) indicate that the model can be practically implemented during the translation of doctor-patient interactions in Yoruba, specifically in a formal healthcare setting.

Comparative Analysis

The outcomes of the present research are similar to the past research on NLP of Yoruba and other low-resource languages. As an example, Rahmon et al. (2024) found a WER of 21% of a comparable system, whereas the present model had a much lower WER (10.1). The increase in the performance is credited to the application of transformer-based models that have demonstrated the best performance in managing the complexities of the language with low resources. This paper identifies the usefulness of transformer architectures in addressing the problems of tonal languages such as Yoruba. Although the model has good prospects of enhancing healthcare communication, it has gaps that can be filled in future. In particular, it will be essential to increase the performance of the model on the data of informal and dialects-contamination to increase its overall generalizability to various situations, such as rural regions where communication is less formal.

5.0 Conclusion

This research was able to make a Yoruba to English NLP model that would facilitate better communication between doctors and patients in the medical facilities, especially those in the Yoruba-speaking areas. The model uses a modern-day transformer architecture to overcome language barriers by combining Yoruba Speech Recognition (YSR) and Yoruba-English Machine Translation (YEMT). The performance metrics, Word Error Rate (WER), Sentence Accuracy, and F1-Score show the promising potential of the model in the real-life application, though there are certain difficulties in the fact that the Yoruba language is tonal and dialect-rich.

The study findings demonstrate that the model does an exceptional job with formal data, with a WER of 10.1% on test set and a sentence accuracy of 77.0%. These findings suggest that this model can transcribe Yoruba speech and translate it into English with relatively high rates of accuracy, which is why it could become a useful tool that can be used in healthcare facilities in Nigeria and other multilingual areas. The values of F1-score varying between 90.8 to 92.4 also make it clear that the model has the capacity to detect and translate the relevant words, preserving the original meaning of the word in the translation. The performance of the model, however, on informal and dialect-laden data, like conversational speech, was not as impressive. The tonal variation and informal terms that are used in the informal Yoruba speech were a problem, making some mistakes in transcription and translation. This problem serves to emphasize the fact that additional development of the model should be undertaken, in order to address the intricacies of tonal languages. Moreover, there were instances of the spikes in WER due to the presence of diacritics within some words, which implied that the more diversified and inclusive datasets, especially the ones devoted to medical terms, would contribute to the enhanced performance of the model in general. In spite of these shortcomings, the formal data performance of the model implies that there is a great potential of the model to enhance communication in the healthcare environment where effective and prompt translations are critical in

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

patient care. This paper adds to the emerging literature on NLP in low-resource languages, specifically in healthcare settings, by showing how transformer models can be used to manage the linguistic complexity of Yoruba. The model can also be used to enhance patient satisfaction, treatment compliance, and ultimately, health outcomes, especially with malaria, where early diagnosis and treatment play an important role. In this study, the need to construct language models specific to low-resource languages such as Yoruba, in which traditional NLP methods did not tend to perform well because of the absence of annotated data and the language specifics, is also highlighted. This study is a success that opens the path to additional developments in NLP of the African languages and it can be applied in medicine, and in other areas like education, government and business.

References

Albrecht, U., V; Behrends, M., Schmeer, R., Matthies, H.K., & Von Jan, U. (2023). Usage of Multilingual Mobile Translation Applications in Clinical Settings. Journal of Health Information Science.

Ajagbe, S.A. (2024). Developing Nigeria Multilingual Languages Speech Datasets for Antenatal Orientation. In: Florez, H., Astudillo, H. (eds) Applied Informatics. ICAI 2024. Communications in Computer and Information Science, vol 2237. Springer, Cham.

Albrecht, J., et al. (2023). Language Barriers in Healthcare. Journal of Healthcare Communication.

Babatunde, Akinbowale Nathaniel, et al. (2024). Speech-to-Text Hybrid English to Yoruba SMS Translator. Innovative Computing Review, 4.1, 15-36.

Devlin, N.J., Shah, K.K., Feng, Y., Mulhern, B., van Hout, B. (2018). Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England. Health Economics, 27, 7-22.

Kale, R., Ojo, J., & Eze, T. (2023). Cultural Communication in Healthcare: Language Barriers and Patient Satisfaction. African Journal of Health Careers.

Kale, A., et al. (2023). Implications of Language Barriers for Healthcare. Oman Medical Journal, 35(2), e122. doi: 10.5001/omj.2020.40.

Meyer, J., et al. (2024). Challenges and Opportunities in NLP for Underrepresented Languages in Healthcare. Journal of AI and Healthcare.

Rahmon, M., et al. (2024). Development of a Yoruba-to-English NLP Model for Healthcare Communication. International Journal of NLP and Healthcare Research.

Shamsi, H., Almutairi, A., Mashrafi, S., & Kalbani, T. (2024). Implications of Language Barriers for Healthcare: A Systematic Review. OMAN Medical Journal.

Umeokafor, N., Eze, T., & Alabi, A. (2022). Addressing Language Barriers with NLP in African Healthcare Systems. Nigerian Medical Journal.

Adelani, D., et al. (2021). Named Entity Recognition for African Languages. Transactions of the Association for Computational Linguistics.

Hossain, E., et al. (2023). Natural Language Processing in Electronic Health Records in Relation to Healthcare Decision-Making: A Systematic Review. Computers in Biology and Medicine.

Meyer, R., et al. (2024). Improving NLP for Code-Switched Speech in Under-Resourced Languages Using Out-of-Domain Data. Journal of Natural Language Processing.

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||
"ENHANCING HEALTHCARE COMMUNICATION: A YORUBA-TO-ENGLISH NLP MODEL FOR DOCTOR-PATIENT INTERACTIONS USING MACHINE TRANSLATION AND SPEECH-TO-TEXT"

Saskia Locke, et al. (2021). Natural Language Processing in Medicine: Understanding the Societal Impacts. Information, Communication & Society.

William Leeson, et al. (2019). Natural Language Processing (NLP) in Qualitative Public Health Research: A Proof-of-Concept Study. International Journal of Qualitative Methods.

Shamsi, H., et al. (2024). Sentiment Analysis in Healthcare: Enhancing Patient Care through Language Processing. OMAN Medical Journal.

Vieira, L. N., O'Hagan, M., & O'Sullivan, C. (2020). Understanding the Societal Impacts of Machine Translation: A Critical Review of the Literature on Medical and Legal Use Cases. Information, Communication & Society, 24(11), 1515-1532.

DECLARATIONS

Ethics Approval and consent to participate: Not Applicable

Consent for publication: Not applicable

Availability of data and Materials: datasets are available from corresponding author upon reasonable request

Competing Interest: Not applicable

Funding: Not applicable

Acknowledgments: The Patients who chose to be anonymous during the gathering of the Yoruba dataset are appreciated.

Appendix

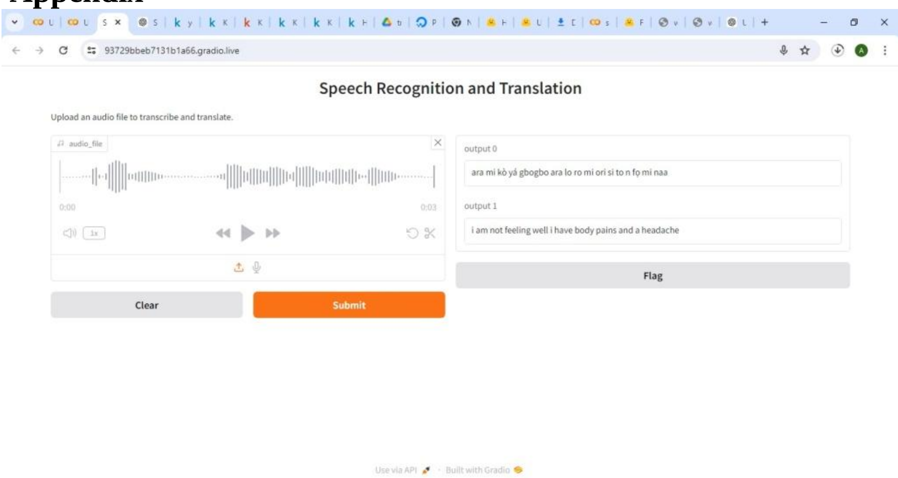

Figure 3: Desktop user interface of the model

IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN 2814-1881) Adefehinti T.O *
https://ijojournals.com/ Volume 09 || Issue 05 || May, 2026 ||

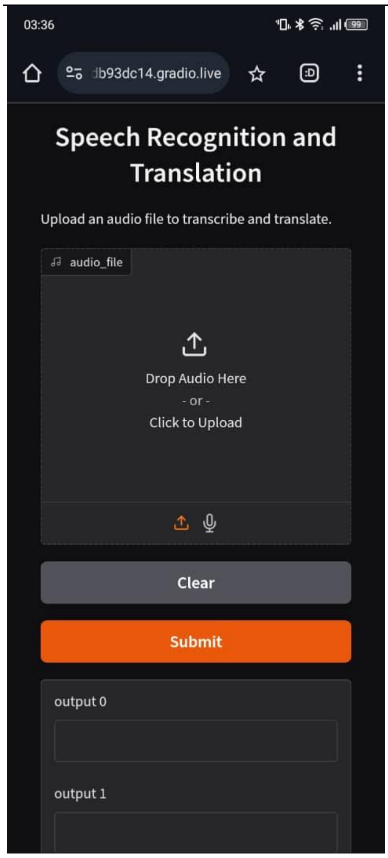

Figure 4: Mobile user interface of the model