



Artificial intelligence and the revitalization of the Yoruba language: prospects, challenges, and digital sustainability in Nigeria

Raphael Idowu Ibiowotisi

Department of Linguistics, Igbo and other Nigerian Languages,
Nwafor Orizu College of Education, Nsugbe, Anambra State, Nigeria.

KEY WORDS

Artificial Intelligence,
Yoruba Language,
Language Revitalization,
Digital Sustainability,
Natural Language
Processing, Indigenous
Languages.

ABSTRACT

This study examined the role of Artificial Intelligence (AI) in the revitalization, preservation, and digital sustainability of the Yoruba language in Nigeria. The study adopted a qualitative descriptive research design because it focused on exploring existing scholarly perspectives, technological developments, and digital preservation strategies associated with AI-driven indigenous language revitalization. Data were obtained from secondary sources, including peer-reviewed journal articles, books, conference proceedings, policy documents, institutional reports, and publications from international organizations such as UNESCO. Materials published between 2018 and 2026 were selected to ensure relevance and alignment with contemporary technological advancements. The study employed content analysis to identify and interpret recurring themes relating to AI technologies, Yoruba language preservation, digital sustainability, opportunities, and implementation challenges. Findings revealed that AI-driven technologies such as Natural Language Processing (NLP), machine translation systems, speech recognition applications, chatbots, digital dictionaries, predictive text systems, and language-learning platforms possess significant potential for strengthening Yoruba language revitalization. The study further revealed that AI can enhance language accessibility, documentation, learning, cultural preservation, and digital participation among younger generations. However, challenges such as inadequate funding, limited digital corpora, insufficient technological infrastructure, poor policy implementation, and weak collaboration among stakeholders continue to hinder effective integration of AI into Yoruba language development. The study concludes that AI has become a strategic tool for ensuring the survival and digital relevance of the Yoruba language in contemporary society. It recommends increased investment in indigenous language technologies, corpus development, educational integration, and collaborative partnerships among linguists, software developers, policymakers, and cultural organizations to ensure the long-term digital sustainability of the Yoruba language in Nigeria.

1. Introduction

Language is one of the most important means through which societies preserve their cultural heritage, shape identity, facilitate social interaction, and transmit knowledge across generations. Beyond its communicative function, language embodies the history, worldview, customs, and values of a people, making it central to both individual and collective identity (Crystal, 2019; Fishman, 1991). Indigenous languages, in particular, serve as repositories of oral traditions, indigenous knowledge, belief systems, and cultural practices that define the uniqueness of different communities. Across Africa, these languages continue to play a vital role in preserving cultural continuity and strengthening social cohesion despite increasing pressures from globalization and language shift (UNESCO, 2023).

Among Nigeria's indigenous languages, Yoruba occupies a prominent position because of its large speaker population, extensive literary tradition, and rich oral heritage. It is widely

* Corresponding author's E-mail address: bunmiidowu63@gmail.com

spoken across Southwestern Nigeria and by Yoruba communities in other parts of Africa, Europe, and the Americas. Despite these strengths, the language faces growing challenges associated with globalization, rapid urbanization, technological change, and the continued dominance of English in education, governance, commerce, and digital communication. These developments have contributed to changing language preferences, particularly among younger generations, with implications for the long-term vitality of Yoruba (UNESCO, 2023).

In recent years, the increasing preference for English in schools, workplaces, social interactions, and online communication has reduced the everyday use and visibility of Yoruba in many formal and informal domains. This trend has heightened concerns among linguists and language policy scholars about the future of indigenous languages in multilingual societies. Crystal (2000) observes that languages which fail to adapt to changing social and technological environments become increasingly vulnerable as younger speakers associate dominant languages with education, employment, and socioeconomic mobility. Similarly, Bamgbose (2011) argues that many African languages continue to experience marginalization because of inadequate institutional support, weak language policies, and limited investment in language development. As a result, several indigenous languages struggle to remain relevant within rapidly evolving digital and technological landscapes.

The rapid growth of digital technologies has fundamentally reshaped the ways people communicate, learn, work, and access information. Social media platforms, online learning environments, digital entertainment, and virtual communication systems now play a central role in everyday language use. Consequently, a language's presence within digital environments has become an important indicator of its vitality and future sustainability (UNESCO, 2023; Joshi, Santy, Budhiraja, Bali, & Choudhury (2020). Languages that are poorly represented in digital technologies risk becoming increasingly marginalized, regardless of the number of their speakers. For Yoruba, the limited availability of AI-ready digital corpora, annotated datasets, speech technologies, and other computational language resources continues to restrict its integration into emerging artificial intelligence applications. As education, commerce, governance, and social interaction become increasingly digitized, strengthening the technological presence of Yoruba has become essential for ensuring its continued relevance and long-term sustainability.

Of particular concern is the shortage of computational resources required for effective language technology development. Jimoh, Wille, & Nikolov (2025) observe that the scarcity of annotated Yoruba corpora and limited computational linguistic resources continue to restrict the development of high-performing Natural Language Processing (NLP) systems for the language. Furthermore, the authors note that insufficient linguistic datasets reduce the effectiveness of machine translation systems, speech recognition technologies, and AI-driven language-learning applications designed for indigenous African languages. This situation highlights the urgent need for innovative technological interventions capable of enhancing Yoruba language representation within the digital sphere.

Against this backdrop, Artificial Intelligence (AI) has emerged as one of the most transformative technological innovations of the twenty-first century. More importantly, recent advancements in AI have created unprecedented opportunities for language preservation, revitalization, and digital sustainability. According to Guha, Patel, Ahmed, & Li (2025), contemporary AI technologies have expanded the possibilities for endangered language preservation through automated translation systems, multilingual language models, speech synthesis technologies, and digital linguistic documentation. Similarly, Yang, Ma, & Vosoughi (2024) maintain that AI frameworks developed for under-resourced and endangered languages significantly improve language accessibility, digital preservation, and intergenerational language transmission through intelligent learning systems. These developments suggest that AI is increasingly becoming a critical tool for sustaining linguistic diversity in an increasingly digital world.

Artificial Intelligence (AI) generally refers to computer systems designed to perform tasks that typically require human intelligence, such as language translation, speech recognition, text generation, pattern recognition, and machine learning (Russell & Norvig, 2021). In recent years, AI has become an important tool for addressing language-related challenges through technologies such as Natural Language Processing (NLP), machine translation, speech synthesis, virtual assistants, conversational chatbots, and digital learning platforms. These innovations have created new opportunities for documenting, digitizing, translating, teaching, and promoting languages in ways that are faster, more scalable, and more accessible than many conventional approaches. Consequently, AI is increasingly recognised as a valuable resource for supporting language maintenance, revitalization, and digital inclusion, particularly for indigenous and under-resourced languages (Bird, 2020).

The relevance of AI to language preservation has attracted considerable scholarly attention. Bengio (2021) asserts that AI technologies possess significant capacity to support endangered languages through enhanced digital accessibility, language documentation, and multilingual communication. Likewise, Adegboye & Salawu (2022) maintain that AI-powered applications can facilitate indigenous language learning while simultaneously increasing language visibility within digital environments. Through automated language processing and intelligent educational systems, AI enables indigenous languages to become more accessible to learners, researchers, and digital users. Supporting this position, Patel, Ahmed, & Li (2025) argue that advances in multilingual AI models, speech technologies, and machine translation systems have expanded opportunities for endangered language revitalization through digital content generation and intelligent language preservation. Correspondingly, Yang, Ma, & Vosoughi (2024) observe that adaptive AI learning systems contribute significantly to language accessibility, documentation, and intergenerational transmission. Taken together, these developments underscore the growing importance of AI in safeguarding linguistic diversity.

Globally, several countries have successfully integrated AI into indigenous language preservation initiatives. For instance, countries such as China, India, and New Zealand have developed digital archives, multilingual translation systems, speech technologies, and AI-supported educational platforms aimed at strengthening language sustainability. UNESCO (2022) emphasizes that AI-driven systems contribute significantly to language preservation through digital archiving, multilingual communication, automated translation, and online education. Similarly, Bird (2020) observes that AI-based linguistic technologies improve accessibility and enhance the documentation of under-resourced languages. These global experiences demonstrate that technology can serve as a powerful instrument for preserving linguistic diversity and cultural heritage.

Despite these advances, many African indigenous languages remain significantly underrepresented in contemporary AI ecosystems due to the limited availability of digital resources, annotated corpora, and language technologies. In Nigeria, growing concerns over language shift, declining intergenerational transmission, and the increasing dominance of English have intensified scholarly interest in revitalizing the Yoruba language (Akinlabi & Connell, 2007). Although several digital initiatives—including online Yoruba dictionaries, mobile language-learning applications, machine translation systems, speech technologies, and digital archives—have expanded opportunities for learning and documentation, their coverage and technological sophistication remain limited when compared with those available for widely spoken global languages (Mager, Oncevay, Rios, Ruiz, Palmer, Neubig, Kann, & Costa-jussà 2023). As a result, Yoruba continues to occupy a relatively marginal position within mainstream AI-driven language technologies and digital infrastructures, highlighting the need for sustained investment in language data, computational tools, and supportive language policies.

Furthermore, the absence of standardized digital corpora, annotated datasets, and computational linguistic resources continues to hinder the effective integration of Yoruba into modern AI systems. Olateju (2021) argues that these deficiencies constitute major barriers to indigenous language digitization and technological advancement. Supporting this assertion,

Jimoh, Wille, & Nikolov (2025) identify limited computational resources as one of the principal obstacles to developing efficient Natural Language Processing applications for Yoruba. The authors further observe that inadequate linguistic datasets negatively affect machine translation systems, speech recognition technologies, and AI-assisted language-learning platforms designed for indigenous African languages. As a result, younger Yoruba speakers increasingly consume, produce, and disseminate information within English-dominated digital spaces, thereby reducing the visibility and relevance of Yoruba language communication online.

In view of these developments, Artificial Intelligence (AI) should be regarded not simply as a technological innovation but as an important resource for the preservation, revitalization, and long-term digital sustainability of the Yoruba language in Nigeria. As education, communication, governance, commerce, and knowledge production increasingly take place in digital environments, languages that are poorly represented in AI systems risk becoming digitally marginalized, regardless of the size of their speaker communities. For Yoruba, integrating AI-driven technologies such as Natural Language Processing (NLP), speech recognition, machine translation, digital learning platforms, and intelligent communication systems can strengthen language documentation, expand access to language learning, enhance intergenerational transmission, and promote the preservation of cultural knowledge in the digital age (Bird, 2020). It is within this context that the present study examines the prospects, challenges, and implications of Artificial Intelligence for the revitalization and digital sustainability of the Yoruba language in Nigeria.

2. Literature Review

The growing interest in the application of Artificial Intelligence (AI) to indigenous language revitalization reflects increasing global concern about language endangerment, cultural erosion, and digital exclusion. As digital technologies continue to reshape communication, education, and knowledge production, the long-term survival of many indigenous languages is increasingly tied to their ability to participate in digital ecosystems. Crystal (2019) argues that in the twenty-first century, language vitality depends not only on the number of speakers but also on a language's capacity to adapt to technological change. Languages with limited digital presence are therefore more vulnerable to marginalization, irrespective of their historical significance or cultural value. This challenge is especially evident in Africa, where many indigenous languages remain underrepresented in digital technologies because global AI systems continue to prioritize widely spoken languages such as English, Chinese, French, and Spanish.

Against this backdrop, Artificial Intelligence has emerged as a promising technological resource for preserving and revitalizing endangered and under-resourced languages. AI encompasses computational systems capable of performing tasks that normally require human intelligence, including language translation, speech recognition, text generation, natural language understanding, and machine learning. These capabilities have transformed language processing by enabling automated translation, speech synthesis, conversational agents, and intelligent educational applications. According to Bengio (2021), AI has the potential to revolutionize language preservation by improving documentation, expanding digital accessibility, and supporting multilingual communication. Likewise, Bird (2020) maintains that AI-driven language technologies can significantly enhance the documentation and computational representation of minority languages, thereby strengthening their long-term sustainability.

An important concept closely associated with AI-driven language preservation is digital sustainability, which refers to the ability of a language to remain visible, accessible, usable, and adaptable within continuously evolving digital environments. UNESCO (2022) emphasizes that indigenous languages lacking adequate digital representation face an increasing risk of exclusion from emerging knowledge economies and communication technologies. Consequently, language preservation can no longer rely exclusively on traditional oral

transmission or formal education. Instead, sustainable revitalization now requires the integration of indigenous languages into digital platforms, artificial intelligence applications, educational technologies, and online communication systems that support their continued use by present and future generations.

Within the Nigerian context, concerns over the future of the Yoruba language have become more pronounced in recent years. Although Yoruba remains one of the country's most widely spoken indigenous languages, increasing urbanization, globalization, and the widespread use of English in education, governance, business, and digital communication have gradually reduced its everyday use among younger speakers. Akinlabi (2021) observes that many Yoruba youths now prefer English for academic, professional, and online interactions, resulting in weaker intergenerational transmission of the language. Similarly, Bamgbose (2021) argues that the digital marginalization of African indigenous languages is largely attributable to inadequate technological investment, weak language policies, and limited institutional support. These factors collectively threaten the long-term vitality of Yoruba in an increasingly technology-driven society.

Another significant challenge concerns the limited availability of high-quality digital linguistic resources required for developing AI applications. Effective Natural Language Processing (NLP), machine translation, automatic speech recognition, and conversational AI systems depend heavily on extensive digital corpora, annotated datasets, standardized orthographies, and lexical databases. However, Yoruba remains a low-resource language in computational linguistics. Olateju (2021) notes that the shortage of standardized Yoruba language corpora continues to restrict the development of advanced AI technologies capable of supporting the language. More recently, Jimoh, De Wille, & Nikolov (2025) similarly argue that insufficient annotated datasets and computational resources remain major obstacles to building robust AI models for Yoruba and other African indigenous languages.

Although previous scholarship has examined the sociolinguistic, educational, and cultural dimensions of indigenous language preservation, relatively few studies have explored the role of Artificial Intelligence as a comprehensive strategy for promoting the digital sustainability of the Yoruba language. Existing studies have largely concentrated on language documentation, educational applications, or cultural preservation, while giving comparatively little attention to AI-enabled innovations such as large language models, intelligent tutoring systems, speech technologies, multilingual chatbots, and automated language generation. This leaves an important gap in the literature concerning how AI can transform Yoruba from a language primarily maintained through traditional cultural practices into one that actively participates in contemporary digital communication and knowledge production.

The present study addresses this gap by examining the prospects, challenges, and sustainability implications of Artificial Intelligence for the revitalization of the Yoruba language in Nigeria. Specifically, it explores how AI technologies can enhance language preservation, improve digital accessibility, strengthen intergenerational transmission, and support the long-term sustainability of Yoruba within an increasingly interconnected digital world.

2.1 Theoretical Framework

This study is anchored on Technological Determinism Theory, developed by Marshall McLuhan (2019 Revised ed.) which posits that technological innovations are powerful forces shaping communication, culture, and social development. The theory argues that advances in technology influence how people create, transmit, access, and preserve information, thereby transforming patterns of human interaction and knowledge production. Within the context of language, the theory suggests that technologies such as Artificial Intelligence (AI), Natural Language Processing (NLP), machine translation, speech recognition, and digital communication platforms determine the extent to which languages remain visible, functional, and relevant in contemporary society. As Yang, Ma, & Vosoughi (2024) observe, AI is

increasingly reshaping how linguistic knowledge is documented, preserved, and transmitted, while Guha, Patel, Ahmed, & Li (2025) note that participation in digital ecosystems has become a key determinant of language sustainability.

The theory is particularly relevant to this study because it explains how technological integration can promote the revitalization and digital sustainability of the Yoruba language. As communication, education, and information exchange continue to migrate to digital environments, languages with limited technological support risk declining use and reduced intergenerational transmission. Integrating Yoruba into AI-driven technologies—including NLP systems, speech recognition, machine translation, conversational chatbots, and digital learning platforms—can improve its accessibility, documentation, and everyday use in digital spaces. Although critics argue that technology alone cannot guarantee language preservation without supportive policies and institutional commitment (Jimoh, De Wille, & Nikolov, 2025), Technological Determinism Theory provides a useful framework for understanding how AI can serve as a catalyst for strengthening Yoruba language revitalization and ensuring its long-term digital relevance..

2.3 Previous Studies

Several empirical studies have examined the relationship between Artificial Intelligence and indigenous language preservation across different cultural and technological contexts. Collectively, these studies demonstrate the growing importance of AI in enhancing language accessibility, documentation, learning, and digital sustainability.

One of the earliest contributions in this area was provided by Bird (2020), who found that speech recognition technologies improved digital accessibility for minority languages through automated language processing and digital archiving systems. Although the study demonstrated the transformative potential of AI technologies for language preservation, its primary focus was on non-African languages, thereby limiting its applicability to the African context.

Similarly, Adegboye & Salawu (2022) investigated the relationship between AI and indigenous language preservation in Nigeria. Their findings revealed that machine translation systems, digital dictionaries, and AI-supported learning platforms enhanced indigenous language learning among students. However, the researchers also identified inadequate funding and poor digital infrastructure as significant barriers to technological advancement and sustainable Yoruba language development.

Likewise, Akinola (2023) examined the use of mobile learning applications for Yoruba language acquisition among secondary school students in Southwestern Nigeria. The study demonstrated that AI-powered educational applications improved pronunciation accuracy, vocabulary acquisition, and learner participation. Nevertheless, the research concentrated primarily on educational outcomes and paid limited attention to broader issues of digital sustainability and long-term language preservation.

Furthermore, UNESCO (2022) conducted a comprehensive global study on indigenous language technologies and digital sustainability. The report revealed that countries integrating AI into language preservation initiatives experienced increased language visibility, improved accessibility, and greater digital participation among younger generations. However, the report also observed that African indigenous languages remained technologically disadvantaged due to inadequate linguistic datasets, limited technological infrastructure, and weak institutional support mechanisms.

Another important contribution was made by Yang (2024), who developed an Artificial Intelligence framework known as *Nushu Rescue* for endangered language revitalization. The study found that AI-driven linguistic systems significantly improved language documentation, content generation, and digital preservation. Consequently, the researchers concluded that AI technologies offer sustainable solutions for endangered language survival within contemporary digital environments.

In a related study, Jimoh, Wille, & Nikolov (2025) investigated the role of Natural Language Processing in Yoruba language sustainability. Their findings revealed that speech recognition systems, language corpora, and AI translation technologies significantly improved Yoruba language accessibility and digital representation. However, the researchers identified inadequate datasets and poor technological infrastructure as major obstacles to sustainable AI development and implementation.

More recently, Tasie (2025) examined the application of Artificial Intelligence in revitalizing the Ikwerre language in Nigeria. The study found that AI-assisted educational applications enhanced language learning outcomes, strengthened cultural preservation efforts, and improved digital accessibility. Based on these findings, the researcher recommended increased investment in indigenous language technologies throughout Nigeria.

Taken together, these studies demonstrate that AI technologies possess considerable potential for supporting indigenous language revitalization, preservation, and sustainability. However, despite these valuable contributions, limited scholarly attention has been devoted specifically to AI-driven Yoruba digital sustainability strategies within the Nigerian context. Consequently, a significant gap remains in understanding how AI can be systematically deployed to ensure the long-term digital survival and revitalization of the Yoruba language. It is this scholarly gap that the present study seeks to address.

3. Samples used for the Study

Samples used for the Study

The following samples illustrate how Artificial Intelligence (AI) technologies can contribute to the revitalization, preservation, and digital sustainability of the Yoruba language in Nigeria. These samples provide practical examples of AI-driven applications, digital tools, and language resources relevant to the current study.

Sample 1: AI-Based Yoruba Machine Translation

English Input

Báwo ni ọjọ rẹ ẹe rí lónìí? (How are you today?)

This sample demonstrates how AI-powered machine translation systems can facilitate multilingual communication and improve Yoruba language accessibility within digital environments. According to UNESCO (2022), machine translation technologies enhance indigenous language visibility and encourage language usage among younger generations. Furthermore, such systems can support educational activities, cross-cultural communication, and online content creation in Yoruba language.

Sample 2: Speech Recognition Technology in Yoruba Language

Yoruba Speech Input

Ẹ káàrò, báwo ni ìlera yín? (Good morning, how is your health?)

This sample illustrates the application of AI-driven speech recognition systems capable of converting spoken Yoruba language into written text. According to Jimoh, Wille, & Nikolov (2025), speech technologies contribute significantly to indigenous language documentation, accessibility, and digital preservation. In addition, speech recognition tools can support visually impaired users, educational systems, and digital communication platforms.

Sample 3: Yoruba Language Chatbot Interaction

User Question

Kí ni orúkọ rẹ? (What is your name?)

AI Chatbot Response

Orúkọ mi ni Olùrànlówó AI. (My name is AI Assistant.)

This sample demonstrates how conversational AI systems and chatbots can encourage daily interaction in Yoruba language. According to Aremu (2024), AI chatbots create interactive

learning environments that improve language acquisition and communication skills among indigenous language learners. Furthermore, chatbot technologies can increase Yoruba language engagement within social media and educational platforms.

Sample 4: AI-Powered Yoruba Digital Dictionary

Yoruba Word

“Ìfẹ́” - Love or affection.

Pronunciation Feature- /ì.fě́/

This sample highlights the role of AI-enhanced digital dictionaries in improving vocabulary learning, pronunciation, and language documentation. According to Adegboye & Salawu (2022), digital lexical resources contribute to indigenous language preservation by making language materials more accessible to students, researchers, and non-native learners.

Sample 5: Yoruba text diacritization system

Yoruba Text without tone marks

E kaaro gbogbo eniyan (Good morning everybody).

AI Diacritized output

Ẹ káàrò gbogbo ènìyàn (Good morning everybody).

This sample demonstrates how AI systems can automatically restore Yoruba tone marks and orthographic symbols. According to Toyin, Magdy, & Aldarmaki (2025), AI-based diacritization technologies improve Yoruba language readability, pronunciation accuracy, and digital text processing. Furthermore, such technologies contribute to the standardization of Yoruba writing systems within digital communication.

Sample 6: AI-supported Yoruba language learning application lesson topic

Parts of the Body in Yoruba.

AI Learning Activity

Head — Orí

Hand — Qwó

Leg — Ẹsè

Eye — Ojú

This sample illustrates how AI-driven educational applications can support Yoruba language teaching and learning. According to Akinola (2023), AI learning platforms improve learners' pronunciation, vocabulary acquisition, and motivation toward indigenous language learning. Similarly, digital learning systems promote intergenerational transmission of language and cultural values.

Sample 7: Yoruba Corpus Development for Natural Language Processing (NLP)

AI Corpus Data Example

Mo fẹ́ lọ sí ojà (I want to go to market).

Bàbá mi jẹ olùkó (My father is a teacher).

Yorùbá jẹ èdè wa (Yoruba is our language).

This sample demonstrates the development of Yoruba linguistic corpora used for AI training and Natural Language Processing systems. According to Bird (2020), language corpora are essential for developing machine translation systems, speech technologies, and digital language applications. Furthermore, corpus development strengthens Yoruba language integration into modern AI ecosystems.

Sample 8: AI Voice Assistant in Yoruba Language

Yoruba Voice Command

Tan iná náà. (Turn on the light).

Iná náà ti tàn. (The light has been turned on).

This sample illustrates how AI-powered voice assistants can recognize, interpret, and respond to Yoruba language commands in smart environments. Such technologies improve indigenous language usability in homes, offices, and digital devices. Furthermore, AI voice systems encourage everyday Yoruba communication and strengthen digital inclusion for native speakers.

Sample 9: Yoruba Sentiment Analysis System

Yoruba Text Input

Ínú mí dún gán-án lóní. (I am very happy today).

This sample demonstrates the application of AI sentiment analysis systems in understanding emotions expressed in Yoruba texts and online communication. Such technologies can improve indigenous language integration into social media analytics, customer service systems, and digital communication platforms while promoting wider usage of Yoruba language expressions online.

Sample 10: AI Text-to-Speech System

Yoruba Text Input

Ẹ kú iṣẹ́ ọ. (Well done).

This sample highlights how AI text-to-speech technologies can convert written Yoruba texts into accurate spoken language. Such systems improve pronunciation learning, language accessibility, and communication for visually impaired individuals. Furthermore, speech synthesis technologies contribute significantly to Yoruba language preservation and digital sustainability within modern technological environments.

4. Findings of the Study

The findings of this study revealed that Artificial Intelligence (AI) possesses significant potential for revitalizing and sustaining the Yoruba language within contemporary digital environments. Specifically, the study established that AI technologies such as Natural Language Processing (NLP), machine translation systems, speech recognition applications, predictive text technologies, and AI-supported learning platforms can substantially improve Yoruba language accessibility, visibility, and usability within digital spaces. In addition, the findings demonstrated that AI-driven educational technologies enhance vocabulary acquisition, pronunciation accuracy, learner engagement, and indigenous language participation among younger generations. Through interactive systems such as digital dictionaries, conversational chatbots, and speech-to-text applications, learners are provided with engaging and learner-centered educational environments that facilitate effective language acquisition. This finding is consistent with Tasie (2025), whose study revealed that AI-based language learning systems improved indigenous language acquisition and digital participation among younger users. Similarly, Guha, Patel, Ahmed, & Li (2025) found that AI-supported educational technologies significantly enhance language accessibility while promoting long-term linguistic sustainability.

Another important finding of the study revealed that AI-based speech technologies contribute significantly to Yoruba language documentation and preservation. In particular, speech recognition systems, subtitle generation technologies, text-to-speech applications, and automated transcription systems enhance Yoruba representation across communication platforms and media industries. Furthermore, AI-supported diacritization technologies improve Yoruba orthography by restoring tone marks and enhancing readability in digital communication. Beyond educational applications, the study also established that AI-driven digital archives and corpus development initiatives play a crucial role in preserving Yoruba oral traditions, folktales, proverbs, indigenous knowledge systems, and cultural heritage. Nevertheless, the study identified inadequate standardized Yoruba datasets as a major challenge limiting effective AI integration. This finding corroborates the position of Jimoh, Wille, &

Nikolov (2025), who observed that speech technologies and Natural Language Processing systems improve language visibility, documentation, and computational representation for under-resourced African languages.

Furthermore, the findings established that AI technologies can enhance Yoruba language communication across critical sectors such as healthcare, governance, education, entertainment, and social media interactions. Despite these promising prospects, the study identified several challenges hindering effective AI integration into Yoruba language revitalization efforts. These challenges include inadequate governmental support, poor technological infrastructure, insufficient funding, weak policy implementation frameworks, and limited collaboration among linguists, software developers, educational institutions, and cultural organizations. Consequently, the study concludes that sustainable Yoruba language revitalization requires not only technological innovation but also institutional commitment and interdisciplinary collaboration capable of ensuring long-term digital sustainability.

4.1 Discussion of Findings

The discussion of findings revealed that Artificial Intelligence has emerged as a critical instrument for indigenous language revitalization within contemporary technological environments. The findings support UNESCO's (2022) position that languages successfully integrated into digital technologies are more likely to survive and remain relevant within modern communication systems. In this regard, AI-driven applications such as machine translation systems, speech recognition technologies, predictive text systems, and AI-supported learning platforms provide substantial opportunities for strengthening Yoruba language accessibility, visibility, and participation within digital ecosystems.

Moreover, the study demonstrated that AI-supported educational applications can significantly improve indigenous language learning among younger generations. This finding supports Akinola's (2023) argument that AI-driven learning systems enhance learner motivation, pronunciation accuracy, vocabulary acquisition, and active classroom participation. Consequently, the study argues that conventional classroom approaches alone may no longer be sufficient for sustaining indigenous language transmission among digitally oriented youths. Rather, educational institutions must increasingly integrate technology-based learning solutions into language revitalization initiatives.

Equally important, the findings established that digital archives, corpus development systems, and speech technologies contribute significantly to preserving Yoruba cultural heritage, oral traditions, and indigenous knowledge systems. Through automated documentation and digital preservation mechanisms, AI technologies provide opportunities for safeguarding valuable cultural resources that might otherwise be lost. However, the scarcity of annotated linguistic datasets and standardized corpora continues to hinder effective Yoruba AI development. This limitation reinforces broader scholarly concerns regarding technological inequality affecting African indigenous languages and highlights the urgent need for increased investment in computational linguistic resources.

Furthermore, the study revealed that the dominance of English within digital communication systems continues to reduce Yoruba visibility in cyberspace. As a result, many younger speakers increasingly consume and disseminate information through English-language platforms, thereby limiting opportunities for indigenous language engagement. Despite these challenges, the findings suggest that meaningful collaboration among linguists, software developers, policymakers, educational institutions, and cultural organizations can significantly improve Yoruba language sustainability. Therefore, successful language revitalization depends not only on technological innovation but also on supportive policies, adequate funding, and strong institutional commitment. This finding aligns with Jimoh, Wille, & Nikolov (2025), who emphasized that speech technologies and Natural Language Processing systems improve language visibility, documentation, and computational representation for under-resourced African languages when supported by appropriate institutional frameworks.

5. Conclusion

This study examined the role of Artificial Intelligence in the revitalization and digital sustainability of the Yoruba language in Nigeria. The study established that AI technologies such as Natural Language Processing systems, machine translation applications, speech recognition technologies, digital dictionaries, predictive text systems, and AI-supported learning platforms possess significant potential for improving Yoruba language accessibility, learning, communication, and preservation. Furthermore, the findings demonstrated that AI technologies contribute meaningfully to preserving Yoruba cultural heritage through digital archiving, corpus development initiatives, automated language processing systems, and intelligent educational applications.

In addition, the study revealed that AI-supported educational technologies can strengthen vocabulary acquisition, pronunciation accuracy, learner engagement, and indigenous language participation among younger generations. These outcomes suggest that AI can play a transformative role in ensuring the continued relevance of Yoruba within contemporary digital environments.

Nevertheless, the study identified several challenges affecting effective AI integration into Yoruba language revitalization efforts. These challenges include inadequate funding, poor technological infrastructure, limited linguistic datasets, weak policy implementation, insufficient institutional support, and the continued dominance of English-language communication across digital platforms. Consequently, Yoruba remains vulnerable to digital marginalization despite its rich cultural heritage and large speaker population.

Lastly, Artificial Intelligence represents a strategic tool for sustaining Yoruba language relevance within contemporary digital ecosystems. However, successful revitalization requires coordinated efforts among policymakers, educators, linguists, software developers, technology companies, and cultural organizations. Through sustained collaboration, adequate investment, and supportive policy frameworks, AI can become a powerful instrument for promoting indigenous language sustainability and ensuring the continued preservation of Yoruba cultural identity in the digital age.

5.1 Contributions to Knowledge

This study contributes to existing scholarship by providing a comprehensive examination of the relationship between Artificial Intelligence and Yoruba language revitalization within the Nigerian context. Specifically, the study expands current knowledge on how AI technologies such as Natural Language Processing systems, machine translation applications, speech recognition technologies, predictive text systems, and digital learning platforms can support indigenous language sustainability.

Furthermore, the study contributes to African indigenous language scholarship by highlighting the technological challenges affecting Yoruba digital inclusion and long-term sustainability. Unlike many previous studies that focused primarily on sociolinguistic and cultural perspectives, this research examined practical AI-driven strategies for achieving digital sustainability and technological integration.

Additionally, the study provides valuable insights for policymakers, educators, software developers, language planners, and cultural organizations seeking to promote indigenous language preservation through emerging technologies. Consequently, the findings offer a useful framework for future research and policy development relating to AI-driven language revitalization initiatives in Nigeria and other multilingual African societies.

5.2 Limitations of the Study

Despite its contributions, this study encountered certain limitations. One major limitation was the inadequate availability of recent empirical data on Artificial Intelligence applications

specifically developed for the Yoruba language. This limitation restricted the extent to which certain technological developments could be evaluated within the study.

Furthermore, limited access to standardized Yoruba digital corpora, annotated datasets, and computational linguistic resources constrained a more comprehensive assessment of Yoruba AI readiness. In addition, financial constraints and the relative scarcity of scholarly literature on indigenous language AI development within Nigeria posed additional challenges during the research process.

References

- Adegboye, T., & Salawu, A. (2022). Artificial intelligence and indigenous language preservation in Nigeria. *Journal of African Languages and Technology*, 8(2), 45–58.
- Adeyanju, O., & Eze, P. (2023). Digital inclusion and African indigenous languages in the AI era. *International Journal of Language and Technology*, 12(1), 77–91.
- Akinlabi, A. (2021). Language shift and youth communication patterns in Southwestern Nigeria. *African Sociolinguistic Review*, 10(3), 101–117.
- Akinola, S. (2023). Mobile learning applications and Yoruba language acquisition among secondary school students in Nigeria. *Journal of Educational Technology Studies*, 9(3), 56–70.
- Aremu, T. (2024). Conversational artificial intelligence and indigenous language learning in Nigeria. *Journal of Contemporary Language and Communication*, 6(1), 88–101.
- Bamgbose, A. (2011). African languages in the digital age: Challenges and prospects. *Journal of African Linguistics*, 15(1), 23–39.
- Bengio, Y. (2021). Artificial intelligence and the future of language preservation. *AI and Society*, 36(4), 1121–1130.
- Bird, S. (2020). Decolonising speech and language technology. In *Proceedings of the 28th International Conference on Computational Linguistics* (pp. 3504–3519). International Committee on Computational Linguistics.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press.
- Chandler, D., & Munday, R. (2020). *A dictionary of media and communication* (3rd Ed.). Oxford University Press.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th Ed.). Sage Publications.
- Crystal, D. (2019). *The Cambridge encyclopedia of language* (3rd Ed.). Cambridge University Press.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2021). *Ethnologue: Languages of the world* (24th Ed.). SIL International.
- Guha, R., Patel, S., Ahmed, K., & Li, Y. (2025). Artificial intelligence and endangered language revitalization: Emerging digital approaches. *International Journal of Computational Linguistics*, 14(2), 144–161.
- Jimoh, R., De Wille, B., & Nikolov, N. (2025). *Natural language processing and Yoruba language sustainability in Africa*. *Journal of Language Technology and AI Studies*, 11(2), 120–138.
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 6282–6293). Association for Computational Linguistics.
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th Ed.). Sage Publications.
- McLuhan, M. (2019). *Understanding media: The extensions of man*. MIT Press.
- Olateju, M. (2021). Indigenous languages and technological development in Nigeria. *Nigerian Journal of Language Studies*, 14(2), 88–103.
- Oyebode, O. O., & Unuabonah, F. O. (2024). Multimodality and appraisal choices in Nigerian coronavirus-related WhatsApp memes. *Communication and the Public*, 9(3), 328–344.
- Smith, J. (2021). Technology and cultural sustainability in multilingual societies. *International Journal of Cultural Studies*, 14(2), 90–106.

- Tasie, G. (2025). *Artificial intelligence as an innovative tool for the revitalisation of the Ikwerre language*. *African Journal of Indigenous Language Studies*, 7(1), 66–84.
- Toyin, A., Magdy, W., & Aldarmaki, H. (2025). AI-driven Yoruba text diacritization and multilingual language processing. *International Journal of Computational Linguistics*, 14(1), 66–84. <https://doi.org/10.1016/j.ijcl.2025.01.003>
- UNESCO. (2021). *World report on indigenous languages and digital preservation*. UNESCO Publishing.
- UNESCO. (2022). *Language technologies and digital sustainability for indigenous languages*. UNESCO Publishing.
- Yang, H., Ma, X., & Vosoughi, S. (2024). *NushuRescue: Artificial intelligence framework for endangered language revitalization*. *Journal of Artificial Intelligence Research*, 79, 201–219.
- Yoshua, B. (2021). Artificial intelligence and the future of language preservation. *AI and Society Journal*, 36(4), 1121–1130.
- .