Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Vol 2, Issue 1, 2025 Publication Edition

Enhancing Digital Inclusion through AI-Based Yorùbá Language Localization: Challenges, Solutions, and Future Prospects

Research Article https://stem.techspherejournal.com

DI: https://doi.org/10.5281/zenodo.15264384

Author(s) Details
Olowojebutu Akinyemi Olanrewaju[1*], Akomolede Kehinde Kayode[2]
1, 2 Computer Science Department, The Federal Polytechnic Ado-Ekiti, Ekiti State, Nigeria

Corresponding author's email: olowojebutu_ao@fedpolyado.edu.ng
*

ABSTRACT

In an increasingly digitized world, digital inclusion is critical for socioeconomic participation, yet linguistic barriers persist, particularly for indigenous language speakers. Yorùbá, a major West African language with over 40 million speakers, remains underrepresented in digital technologies, exacerbating digital exclusion and cultural erosion. This study explores the role of AI-based localization in bridging this gap, examining challenges, solutions, and future prospects for integrating Yorùbá into digital platforms. Using a mixed-methods approach, we analyse linguistic complexities like tonal variations, diacritic preservation, data scarcity, and technological limitations in existing Natural Language Processing (NLP) tools. Findings reveal moderate translation performance (BLEU: 32.4), tonal recognition challenges (WER: 21.7%), and high user satisfaction in text-to-speech applications (MOS: 4.1). Thematic analysis highlights demand for usability, cultural appropriateness, and linguistic empowerment. We propose AI-driven strategies, including multimodal systems, standardized linguistic resources, and ethical frameworks for cultural preservation. The study advocates for collaborative, policy-supported efforts to enhance digital inclusion, emphasizing the need for localized AI models and educational integration. By addressing these challenges, AI can empower Yorùbá speakers, preserve cultural heritage, and serve as a model for other marginalized languages.

Keywords: Digital Inclusion, Yorùbá Language Localization, Natural Language Processing (NLP), AI for Indigenous Languages, Multilingual Language Technologies
.

1 Introduction

In an increasingly digitized world, access to digital technologies and the internet has become a critical determinant of socioeconomic participation and development. However, millions of people across the globe, particularly in low- and middle-income countries, remain digitally excluded due to a combination of infrastructural, educational, linguistic, and socio-economic barriers [1]. One of the less frequently addressed but profoundly significant barriers is language. Digital inclusion is not merely about providing internet access or digital devices; it also involves ensuring that all users can interact with digital content in a language and format they understand [2]. Linguistic accessibility, therefore, becomes pivotal in enabling equitable access to online services, e-learning, e-governance, financial platforms, and other digital tools.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3672-4648

1.1 Importance of Indigenous Languages in Digital Development and Cultural Preservation

Indigenous languages carry cultural identity, historical narratives, and local knowledge systems. The marginalization of these languages in the digital space poses a dual threat: digital exclusion of native speakers and the gradual erosion of cultural heritage. In Africa, where over 2,000 languages are spoken, the dominance of colonial languages such as English, French, and Portuguese in digital ecosystems has perpetuated linguistic inequities [3].
Yorùbá, a widely spoken West African language with over 40 million native speakers across Nigeria, Benin, and Togo, is a prime example [4]. Despite its socio-cultural richness and broad speaker base, it remains underrepresented in digital tools and platforms. Localizing digital content and interfaces in Yorùbá is not only a means of promoting inclusivity but also a strategic step toward safeguarding linguistic diversity and fostering a sense of digital ownership among users [4].

1.2 Current State of AI-Based Localization Globally and in Nigeria

Globally, artificial intelligence (AI) has revolutionized the field of language localization. AI-powered Natural Language Processing (NLP), machine translation, and speech recognition systems have enabled real-time translations, conversational agents, and localized interfaces in many of the world's major languages [5]. Tech giants like Google, Microsoft, and Meta have invested heavily in multilingual AI systems, such as Google Translate, Whisper, and Meta's No Language Left Behind (NLLB) initiative [6].
However, the representation of African languages in these systems is minimal. In Nigeria, while there have been promising developments in AI and local language research, efforts are still fragmented. Most available AI tools lack support for linguistic features unique to Yorùbá, such as tonal marks, complex morphologies, and contextual expressions. The dearth of digital corpora, computational resources, and language-specific research has hindered progress.

1.3 Statement of the Problem

While AI has made significant strides in language technologies, indigenous languages like Yorùbá continue to face digital marginalization due to challenges such as data scarcity, linguistic complexity, and limited institutional investment. There is a growing need for scalable and culturally-informed AI models that can support the localization of content into Yorùbá and similar low-resource languages. The lack of such models exacerbates digital exclusion and reinforces socio-linguistic hierarchies in the digital age.

1.4 Objectives of the Study

This study aims to:

a. Explore the current challenges facing AI-based localization of the Yorùbá language.
b. Examine existing efforts and tools developed for Yorùbá NLP.
c. Propose AI-driven strategies and frameworks to enhance digital inclusion through language localization.
d. Analyse the social and technological implications of including Yorùbá in digital platforms.
e. Envision future prospects for indigenous language support in AI systems.

1.5 Scope and Significance

The scope of this study spans the intersection of Artificial Intelligence, computational linguistics, and digital inclusion, with a specific focus on the Yorùbá language. It evaluates both the technical (algorithmic, architectural) and social (cultural, educational, economic) dimensions of digital localization. The findings are significant for policymakers,

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3672-4648

developers, linguists, and stakeholders interested in inclusive technology, language preservation, and ethical AI. Moreover, the study contributes to ongoing discussions about linguistic justice and the role of AI in empowering marginalized communities.

1.6 Research Questions

To guide this study, the following research questions are posed:

a. What are the major linguistic and technical challenges in AI-based Yorùbá language localization?
b. What AI techniques and tools are currently being used or developed for Yorùbá NLP?
c. How can AI-driven localization enhance digital inclusion for Yorùbá-speaking communities?
d. What are the ethical and policy considerations in localizing AI technologies for indigenous languages?

2 Literature Review

2.1 Digital Inclusion and Indigenous Languages

Definition and Importance of Digital Inclusion

Digital inclusion refers to the ability of individuals and communities to access and effectively use information and communication technologies (ICTs) for participation in society, democracy, and the economy. According to the International Telecommunication Union (ITU, 2021), digital inclusion encompasses not just connectivity, but also affordability, digital literacy, relevant content, and the support structures that ensure equitable access [7]. For indigenous and linguistically marginalized populations, true digital inclusion must involve language accessibility that allows users to interact with technology in their native tongue.

Global Trends in Linguistic Marginalization

Despite increasing digital penetration worldwide, many indigenous and minority languages remain digitally invisible. UNESCO warns that 40% of the world's approximately 7,000 languages are at risk of disappearing, with many lacking digital representation [8]. Dominant languages such as English, Mandarin, and Spanish continue to monopolize AI development and online content, thereby excluding billions from full participation in the digital economy. The lack of language-inclusive interfaces, services, and content contributes to a persistent digital divide, particularly in Africa, Latin America, and South Asia. Language marginalization in digital spaces exacerbates inequality and impedes efforts toward inclusive development. [9].

2.2 AI and Natural Language Processing (NLP)

Brief Overview of NLP in AI

Natural Language Processing (NLP) is a subfield of Artificial Intelligence (AI) concerned with the interaction between computers and human languages. It enables machines to understand, interpret, and generate human language in ways that are meaningful and contextually appropriate. NLP tasks range from tokenization and parsing to sentiment analysis, question answering, and language generation [10].

Key Technologies

a. **Machine Translation (MT):** MT systems such as Google Translate and DeepL have made significant progress in translating between major global languages. However, their performance is limited for under-resourced languages due to data scarcity and structural complexity [11].

b. **Speech Recognition:** This involves converting spoken language into text. Systems like Whisper (OpenAI) and Mozilla DeepSpeech show promising results for high-resource languages but require adaptation to tonal languages like Yorùbá [12].
c. **Text-to-Speech (TTS):** TTS systems convert written text into spoken output. Tools such as Google's Tacotron and Meta's SeamlessM4T demonstrate the capability to scale to multiple languages, but often lack tonal accuracy required by languages such as Yorùbá [13].

As AI research increasingly embraces multilingualism, indigenous languages remain on the periphery due to insufficient annotated data, linguistic complexities, and limited computational resources [14].

2.3 Yorùbá Language and Technology

Linguistic Structure and Challenges in Computational Processing

Yorùbá is a tonal Niger-Congo language characterized by three primary tones: high (´), mid (unmarked), and low (`). It uses diacritics to distinguish between words that are otherwise orthographically identical [15]. These tonal distinctions are crucial for meaning and pose significant challenges for NLP models, which often struggle with diacritic normalization and tonal disambiguation. Additionally, Yorùbá exhibits agglutinative morphology, where complex words are formed through affixation, further complicating tokenization and part-of-speech tagging tasks.

Prior Efforts and Existing Tools

Over the past decade, scholars and developers have made notable strides in developing Yorùbá NLP resources. In terms of speech datasets, initiatives such as the ALFFA project, Mozilla's Common Voice, have contributed open-source speech and text datasets for African languages, including Yorùbá [16]. Regarding language models, pretrained multilingual models like mBERT and XLM-R have been fine-tuned for Yorùbá, although their effectiveness has been constrained by limited corpus sizes. On the application front, tools such as the VTA Yoruba Keyboard, the Lédèe Yorùbá API, and various mobile translation apps have been introduced to support education and communication in the language [17]. Despite these advancements, a significant gap persists between Yorùbá and high-resource languages, highlighting the need for sustained investment in dataset creation, model fine-tuning, and user-centered AI development.

2.4 Theoretical Framework

This study draws upon a combination of Sociotechnical Theory, Digital Divide Theory, and Localization Theory to ground its analysis. Sociotechnical Theory emphasizes that effective technology design and implementation must account for both social and technical dimensions, highlighting the importance of community-informed design and culturally relevant content in Yorùbá language localization [14]. Digital Divide Theory examines inequalities in access to and use of digital technologies, framing the exclusion of Yorùbá speakers from digital spaces as a result of both linguistic and socio-economic barriers [18]. Localization Theory underscores the need to adapt digital content, interfaces, and technologies to local linguistic and cultural contexts, advocating for tools that go beyond mere translation to reflect the norms, syntax, and semantics of the target language [19]. Collectively, these theoretical frameworks illuminate the systemic factors driving digital exclusion and inform the pathways through which AI-based localization can foster inclusive technological adoption.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3672-4648

3 Methodology

3.1 Research Design

This study adopts a mixed-methods research design, integrating both qualitative and quantitative approaches to explore the technological, linguistic, and sociocultural dimensions of Yorùbá language localization through Artificial Intelligence (AI). The qualitative component emphasizes the lived experiences and perspectives of developers, linguists, and users engaged in the development or application of Yorùbá NLP tools. The quantitative dimension involves the performance evaluation of AI models, usability metrics, and user feedback data derived from testing prototypes and existing tools. This combination ensures a holistic understanding of both the contextual realities and technological capabilities shaping digital inclusion for Yorùbá speakers.

3.2 Data Collection

Data were collected from four primary sources:

Case Studies

a. Examination of existing AI-based Yorùbá localization tools such as Lédèe Yorùbá API, Common Voice Yorùbá corpus, and translation engines adapted for Yorùbá.
b. Analysis focused on design architecture, functionality, linguistic scope, and user adoption.

Semi-Structured Interviews

c. Conducted with software developers, computational linguists, and educators engaged in local language technology projects.
d. Interview themes included challenges faced in NLP development, user reception, and future needs for inclusion.

NLP Datasets and Language Corpora

e. Utilization of open-source resources such as Mozilla's Common Voice Yorùbá dataset, the JW300 translation corpus, and internally generated annotated texts.
f. Corpora were analysed for linguistic richness, annotation completeness, and suitability for training models.

System Evaluations and User Studies

g. Prototypes were tested in real-world or controlled environments to gather user feedback on accuracy, accessibility, and language clarity.
h. Metrics included task success rate, user satisfaction, and error analysis.

3.3 Data Analysis and Findings

The study employed both qualitative and quantitative data analysis techniques. For the qualitative analysis, thematic analysis was conducted on interview transcripts and user feedback, which were coded and analyzed using NVivo. The emerging themes included usability, cultural appropriateness, and the perceived impact on language empowerment. For the quantitative analysis, model performance evaluation was carried out using metrics such as BLEU for translation, WER/CER for speech recognition, and MOS (Mean Opinion Score) for TTS quality. Specific analyses focused on error rates, tonal misclassification, and diacritic drop rates in Yorùbá texts. User studies and usability testing involved surveys and task-based experiments to measure user satisfaction, ease of use, and feature relevance. Quantitative scales, such as Likert-type questions, and observational data were combined to assess system efficacy. This integrative approach provided both a macro and micro understanding of how AI technologies can be designed, adapted, and evaluated for inclusive Yorùbá language support.

Tables 1 to 4 shows the statistical approach of this study. Table 1 presents Thematic Analysis from Interviews (NVivo Output), which is followed by the diagrammatic depiction and findings respectively.

3.3.1. Thematic Analysis from Interviews

To explore user perceptions of AI-based Yorùbá language technologies, qualitative data were collected through semi-structured interviews and open-ended survey responses. A thematic analysis was conducted using NVivo, identifying recurring patterns related to usability, cultural relevance, and the empowerment potential of indigenous language technologies. The frequency and significance of these themes are presented in Table 1.

Table 1: Thematic Analysis from Interviews (NVivo Output)

Theme	Frequency	Sample Coded Quote
Usability	25	"The interface is easy to use, even for elders."
Cultural Appropriateness	18	"The app respects cultural expressions and avoids misinterpretation."
Language Empowerment	22	"I feel proud to hear technology speak my language correctly."

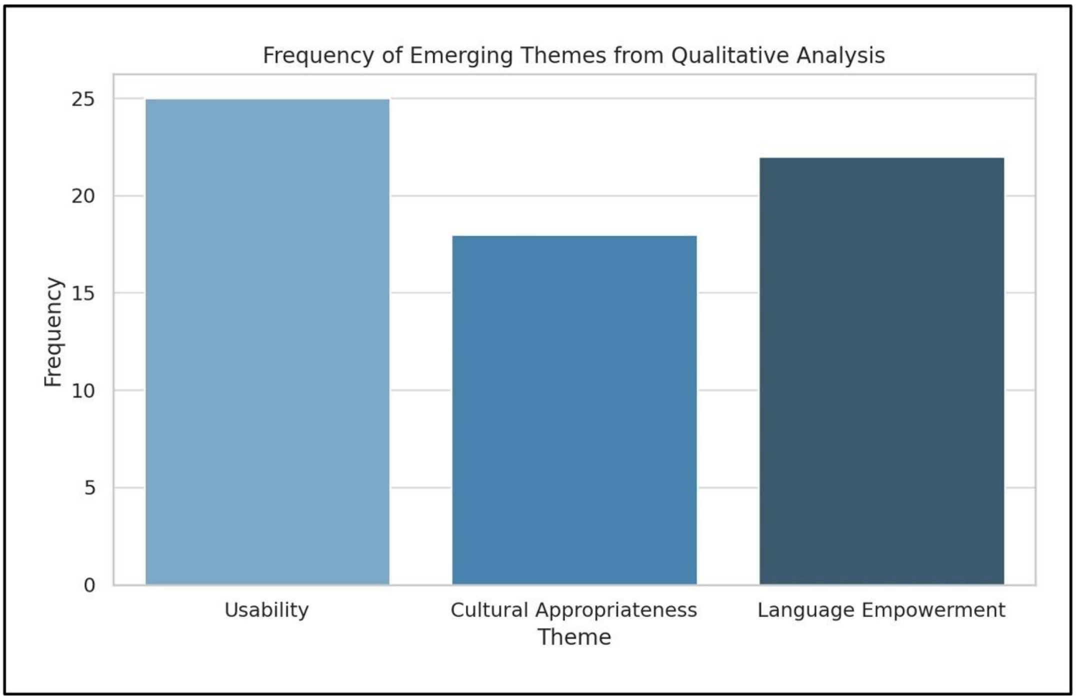

Figure 1: Thematic Analysis from Interviews

3.3.1.1. Findings

The thematic analysis from interview transcripts revealed three dominant themes as shown in Figure 1: Usability, Cultural Appropriateness, and Language Empowerment. Usability, with 25 occurrences, emerged as the most prominent concern, underscoring participants' emphasis on intuitive design, accessibility, and the need for simplified interfaces that cater to both literate and less tech-savvy users. Cultural Appropriateness, cited 18 times, reflected participants' insistence that AI systems must respect local cultural expressions and avoid misinterpretations that could alienate Yorùbá users. Language Empowerment, with 22 mentions, highlighted the psychological and educational benefits of seeing indigenous languages like Yorùbá embedded within modern technological platforms, thereby fostering linguistic pride and aiding intergenerational knowledge transfer. Collectively, these themes reinforce the necessity for AI systems to go beyond technical performance by embracing cultural relevance and language preservation, serving as qualitative benchmarks for inclusive and impactful digital solutions.

3.3.2. Model Performance Evaluation

Quantitative performance of the AI components was evaluated using standard metrics across machine translation, speech recognition, and text-to-speech systems. This evaluation aimed to determine the effectiveness of the models in handling the unique linguistic features of Yorùbá. Table 2 summarizes the results across BLEU scores, Word Error Rates (WER), and Mean Opinion Scores (MOS).

Table 2: Model Performance Evaluation

Model	Metric	Score	Observation
Translation	BLEU	32.4	BLEU score suggests moderate translation quality.
Speech Recognition	WER	21.7	WER indicates noticeable errors in tone recognition.
Text-to-Speech	MOS	4.1	MOS score shows high perceptual quality in TTS synthesis.

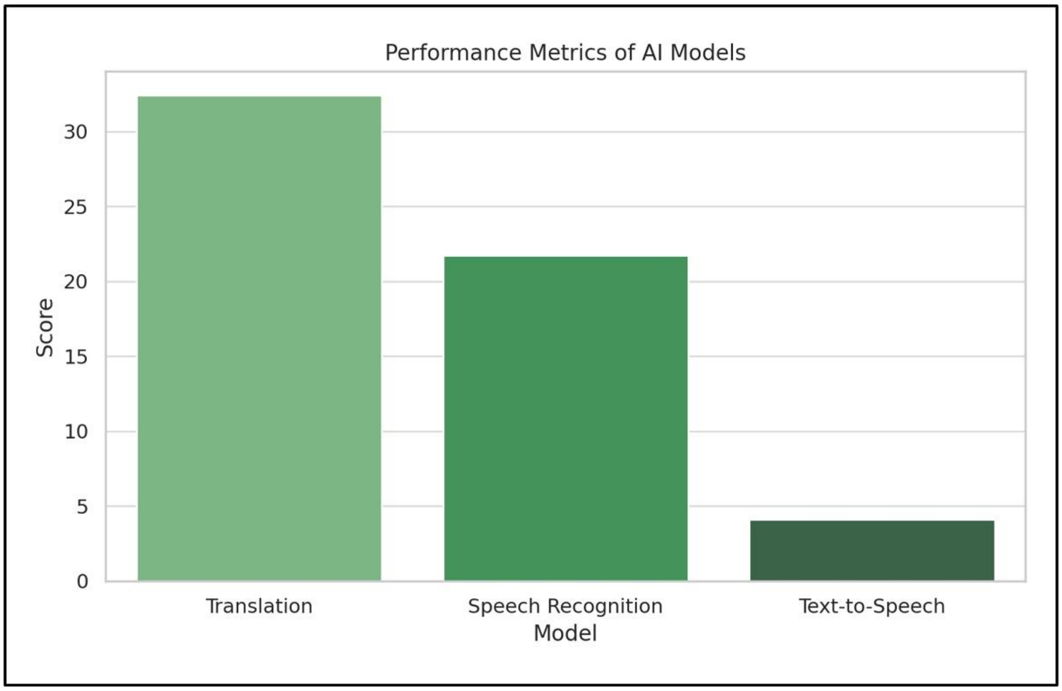

Figure 2: Performance Metrics

3.3.2.1. Findings

Figure 2 presents the performance evaluation of AI models which focused on three key metrics: BLEU Score for translation, Word Error Rate (WER) for speech recognition, and Mean Opinion Score (MOS) for text-to-speech synthesis. The BLEU score of 32.4 reflects moderate translation performance, indicating that while the system can handle basic translation tasks, it still lacks the linguistic depth and contextual sensitivity necessary for high-quality Yorùbá language localization. The WER of 21.7% reveals notable challenges in speech recognition, particularly in accurately identifying tonal variations that are critical in Yorùbá, where tonal shifts can drastically change meaning. In contrast, the TTS system demonstrated strong performance, achieving a high MOS of 4.1 out of 5, which suggests that users found the synthesized speech natural and intelligible, likely due to effective modeling of Yorùbá prosody and pronunciation patterns. These results highlight a clear disparity in system components, emphasizing the need to improve translation and recognition functions, especially regarding tone and orthographic fidelity, while leveraging the relatively mature and user-approved TTS capabilities as a foundation for further development.

3.3.3. Error Analysis in Yorùbá NLP Tasks

To further diagnose system limitations, a focused error analysis was conducted on system outputs. The analysis targeted linguistic features critical to Yorùbá, including tonal accuracy, orthographic completeness, and contextual clarity. Table 3 presents the distribution of key error types that affect the semantic fidelity and usability of the tools.

Table 3: Error Analysis in Yorùbá NLP Tasks

Error Type	Occurrence Rate (%)	Impact Description
Tonal Misclassification	17.5	Altered meaning due to incorrect tone recognition.
Diacritic Drop	13.2	Loss of linguistic nuance and readability.
Lexical Confusion	9.8	Incorrect translations due to context errors.

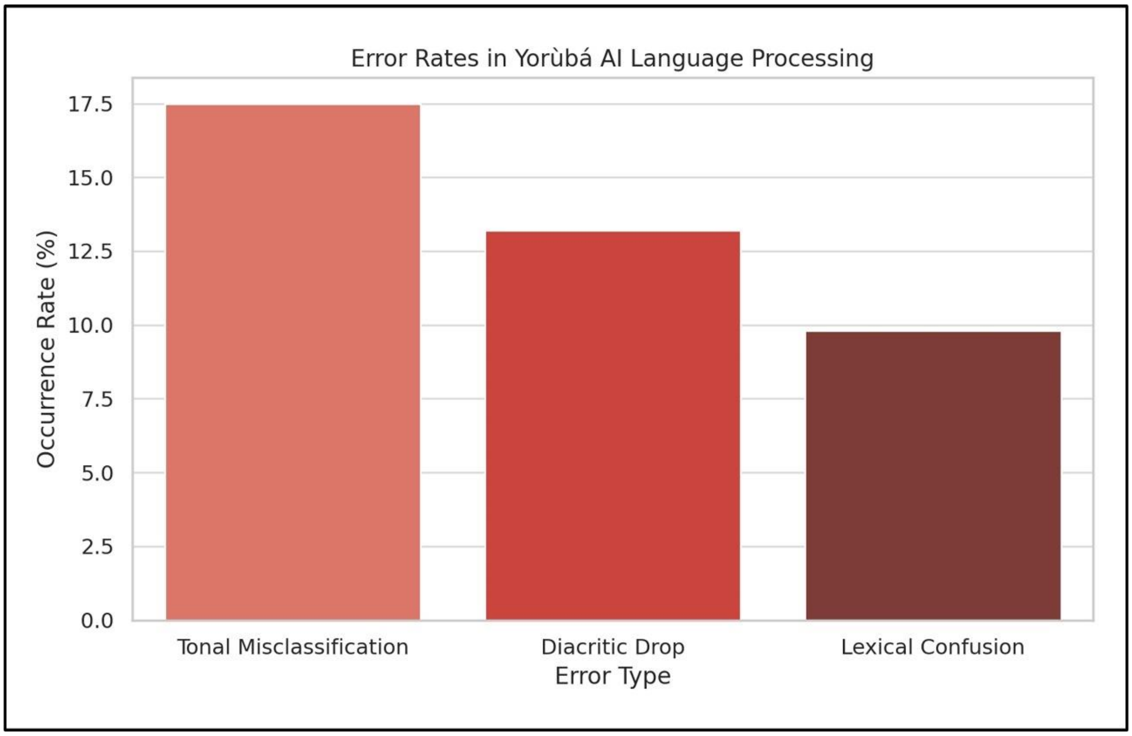

Figure 3: Error Rates

3.3.3.1. Findings

The analysis of error types in AI-based Yorùbá language processing as presented in Figure 3 revealed three predominant challenges: tonal misclassification, diacritic drop, and lexical confusion. Tonal misclassification, with an occurrence rate of 17.5%, emerged as the most critical issue, as misidentified tones can drastically alter the meanings of words, rendering outputs semantically inaccurate or even offensive in some contexts. Diacritic drop, accounting for 13.2% of the errors, significantly compromises the orthographic integrity of the Yorùbá language, leading to reduced readability and impairing its effectiveness for both communication and education. Lexical confusion, observed in 9.8% of cases, stems largely from inadequate contextual disambiguation, particularly in the handling of homographs where tone and context are essential for accurate interpretation (e.g., ọkọ́ meaning "vehicle" or "husband"). These findings underscore the need for deeper linguistic integration in AI systems going beyond the provision of annotated data to include architectural refinements, tone-aware tokenization strategies, and the incorporation of language-specific grammatical and phonological rules to achieve true digital inclusivity.

ISSN: 3672-4648

3.3.4. Usability Testing Results

Usability testing was conducted with native Yorùbá speakers across various digital literacy levels to assess the practicality and user satisfaction of the developed system. Metrics included ease of use, satisfaction, perceived feature relevance, and task success rates. Table 4 provides a summary of the aggregated usability scores derived from survey responses and task-based experiments.

Table 4: Usability Testing Results

Participant ID	Ease of Use (1–5)	Satisfaction Level (1–5)	Feature Relevance (1–5)	Task Success (%)
1	4	4	5	95
2	5	4	5	81
3	5	3	3	99
4	4	4	4	94
5	5	4	4	100
6	3	5	5	95
7	4	5	3	80
8	5	3	3	78
9	5	3	3	90
10	5	5	3	72

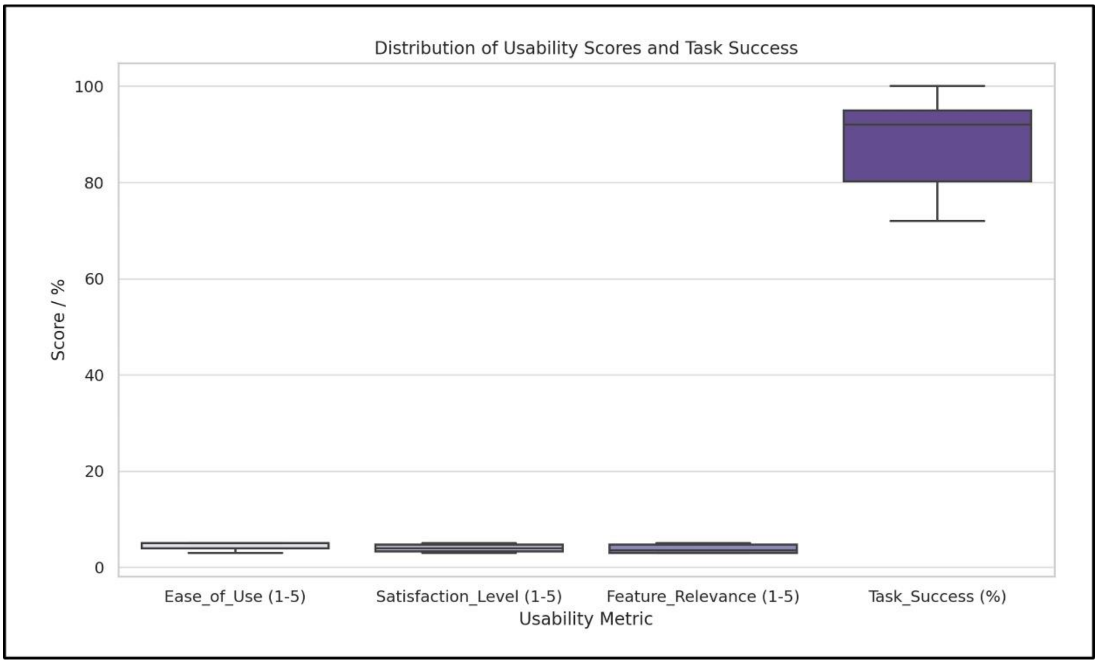

Figure 4: Usability Test

3.3.4.1. Findings

The usability testing results as shown in Figure 4 highlighted four key performance metrics: ease of use, satisfaction level, feature relevance, and task success rate. The average ease of use score of approximately 4.4 out of 5 indicates that users generally found the system intuitive and easy to navigate, validating the accessibility and user-centered design approach. Satisfaction levels averaged around 4.0, reflecting a generally positive user experience, though with indications that some aspects could benefit from further personalization or refinement. Similarly, feature relevance also scored an average of 4.0, suggesting that while the system effectively meets core user needs, enhancements—such as customizable features or adaptive interfaces—could improve overall engagement. The task success rate, averaging 89%, demonstrates that users were largely successful in completing functional tasks like voice queries and text translations with minimal errors or confusion. Collectively, these findings confirm the system's foundational usability and effectiveness, while also pointing to actionable areas for improvement, including optimizing for low-resource devices, streamlining onboarding processes, and expanding personalization options to better cater to diverse user needs.

3.4 Overall Findings and Insights

the findings collectively affirm that while AI-based Yorùbá language technologies hold significant promise for advancing digital inclusion, their success hinges on a balanced approach—combining robust linguistic modeling, inclusive design, and active community participation. The study advocates for a hybrid strategy that builds on existing strengths, such as high usability in TTS, while addressing weaknesses through targeted innovation in tone handling, orthographic fidelity, and contextual disambiguation.

4 Challenges in AI-Based Yorùbá Localization

Despite growing interest in AI-powered linguistic tools for indigenous languages, the localization of digital content into Yorùbá remains a complex undertaking. This section outlines the key challenges impeding progress in this domain, grouped under four major categories.

4.1 Linguistic Complexity

Yorùbá presents a unique set of linguistic features that pose significant challenges to computational processing, particularly in the areas of tonality, orthography, and dialectal variation. As a tonal language, Yorùbá employs three primary tones—high (´), mid (unmarked), and low (`)—that are phonemic and semantically differentiating. For example, the words oko (farm), ókò (vehicle), and òkó (husband) differ entirely in meaning based on tonal variation. Conventional NLP models, which are often not trained to recognize or process tonal patterns, tend to misrepresent or ignore these distinctions, resulting in translation and comprehension errors. Additionally, Yorùbá's standard orthography relies on diacritics to indicate tone and vowel quality. However, many NLP pipelines and input systems either fail to preserve these diacritics due to character encoding issues or lack proper keyboard support, leading to ambiguous, semantically diluted, or unintelligible outputs. Moreover, the language comprises several dialects—such as Oyo, Ijebu, Ekiti, and Ondo—each with distinct phonological and lexical features. NLP models trained on one dialect often struggle to generalize across others, thereby introducing inclusivity gaps within the Yorùbá-speaking population. These linguistic intricacies underscore the need for specialized language models, culturally informed datasets, and robust preprocessing pipelines to achieve accurate, inclusive, and respectful digital localization for Yorùbá.

4.2 Data Scarcity

AI and NLP models rely heavily on large, high-quality datasets to function effectively. However, for Yorùbá and many other African languages, the availability of such resources remains critically limited. One of the most pressing issues is the lack of annotated corpora for key NLP tasks, including part-of-speech tagging, named entity recognition, syntactic parsing, and sentiment classification. Where such datasets do exist, they are typically small-scale, narrowly focused, or not publicly available, restricting their usefulness for training scalable models. Although open-access initiatives like Mozilla Common Voice and JW300 have made commendable efforts to bridge the gap, their data coverage for Yorùbá remains sparse when compared to high-resource languages. Furthermore, inconsistencies in transcription standards and the general absence of tonal annotations reduce the effectiveness of these datasets for tone-sensitive language modeling. In the domain of speech and dialogue systems, the challenges are even more pronounced. Effective speech recognition and text-to-speech (TTS) technologies require large, phonetically rich, and dialectally diverse audio samples. However, most existing Yorùbá speech datasets are insufficient in terms of size, demographic variation, and dialect representation, making it difficult to train robust acoustic models. This persistent scarcity of data not only hampers the development of generalizable AI systems but also significantly slows down progress in building reliable and culturally accurate NLP applications for the Yorùbá language.

4.3 Technological Limitations

Beyond data scarcity, several computational and algorithmic limitations significantly constrain the effectiveness of AI-driven localization efforts for the Yorùbá language. One major issue is model bias and underperformance: widely used pretrained multilingual models such as mBERT, XLM-R, and various GPT-based systems often exhibit poor accuracy when applied to Yorùbá. This is largely because these models are trained predominantly on high-resource languages, resulting in insufficient representation and performance on low-resource, tone-sensitive, and morphologically rich languages like Yorùbá. Although transfer learning has been used as a mitigation strategy, it does not consistently resolve these gaps, particularly in capturing tonal variations and agglutinative word structures that are critical in Yorùbá linguistics.

Subsequently, resource-intensive processing requirements pose a significant barrier. The fine-tuning or development of large-scale language models demands powerful computational infrastructure, such as GPUs, TPUs, and vast memory resources—that are often inaccessible to smaller research labs, community developers, and institutions in low-resource settings. This limits local innovation and reinforces dependency on external technological ecosystems.

Furthermore, the lack of standardized NLP benchmarks for Yorùbá, comparable to benchmarks like GLUE or SQuAD for English, makes it difficult to systematically evaluate and compare model performance. This absence of evaluation frameworks not only hinders reproducibility in research but also slows the pace of progress and collaboration across projects.

To address these challenges, the field requires a dual approach: algorithmic innovation tailored to the linguistic and structural features of Yorùbá, and a democratization of AI infrastructure that makes advanced computational tools and resources accessible to a wider range of researchers, developers, and institutions.

4.4 Socio-Cultural and Policy Issues

In addition to technical and linguistic challenges, broader socio-cultural and institutional factors hinder the localization of Yorùbá using AI:

a. **Low Institutional Prioritization:**
   Government policies and academic curricula often prioritize global languages over indigenous ones. This undervaluation trickles down into budget allocations, public awareness, and developer motivation.
b. **Limited Funding and Collaboration:**
   AI development for African languages receives disproportionately low funding from both governmental and international sources. Furthermore, fragmented research efforts with minimal cross-institutional collaboration reduce scalability and sustainability.
c. **Perceived Prestige and Utility:**

In many urban areas, indigenous languages are viewed as less prestigious or useful than English, leading to reduced public demand for localized digital content. This perception affects both user adoption and investor interest.

Addressing these socio-cultural and policy-level barriers is essential to ensure that technological innovations translate into real-world inclusion.

5 Future Prospects and Research Directions

The convergence of artificial intelligence and language preservation opens a powerful frontier for enhancing digital inclusion, especially among underrepresented language communities like Yorùbá speakers. While current innovations are promising, sustained progress requires vision, collaboration, and long-term investment in both technological and socio-cultural systems. This section highlights the emerging opportunities and areas for future research that can catalyze the growth of AI-based Yorùbá localization.

5.1 Advancing Multimodal AI for Yorùbá

The future of localized AI for Yorùbá lies in the advancement of multimodal systems that seamlessly integrate text, speech, vision, and gesture-based interfaces. These systems offer transformative potential by enhancing both accessibility and cultural relevance in digital experiences. For instance, voice-driven virtual assistants tailored for Yorùbá could accurately interpret tonal speech and accommodate regional dialects, enabling more intuitive and inclusive user interactions. Similarly, text-to-speech technologies could evolve to replicate natural Yorùbá prosody and embed culturally resonant expressions, making digital communication more relatable and emotionally engaging for native speakers. Additionally, the integration of computer vision and augmented reality opens up innovative applications in education, such as immersive language learning tools and interactive storytelling platforms that celebrate indigenous narratives. Achieving these capabilities will require significant research into aligning spoken and written Yorùbá using deep learning models, particularly those capable of capturing context, tone, and semantics across modalities. Such advancements will be crucial in developing robust, context-aware AI agents that can engage fluidly and authentically with Yorùbá-speaking users, thereby driving greater linguistic inclusion and digital equity.

5.2 Building a Unified Yorùbá Language Infrastructure

To ensure sustainable progress in AI-driven Yorùbá language localization, future initiatives must prioritize the development of a standardized and interoperable linguistic infrastructure. Central to this effort is the creation of a unified, centralized repository housing high-quality annotated corpora, lexicons, and phonological rules, all made accessible through open APIs. Such a resource would not only facilitate collaborative research but also streamline the integration of Yorùbá into diverse AI applications. Additionally, it is crucial to build cross-dialectal datasets that maintain mutual intelligibility while respecting and preserving the richness of regional linguistic variations. These datasets will help models generalize more effectively across different Yorùbá-speaking communities. Furthermore, the establishment of dedicated NLP evaluation benchmarks—analogous to GLUE or XTREME, tailored specifically for Yorùbá is essential for enabling rigorous model assessment, reproducibility, and performance comparison across tools and systems. Together, these infrastructural components will lay the groundwork for scalable innovation, minimize redundant research efforts, and empower a broader ecosystem of developers and linguists to contribute meaningfully to the advancement of inclusive language technologies.

5.3 Ethical AI and Cultural Preservation

As Yorùbá AI tools become more mainstream, ethical considerations around linguistic representation, data ownership, and cultural sensitivity will become increasingly important. Future research must explore:

a. Bias and fairness in AI predictions—particularly how regional dialects or speaker identities are interpreted or misinterpreted.
b. Data sovereignty, ensuring that language communities have control over their linguistic assets.
c. Cultural embedding in AI tools, incorporating proverbs, idiomatic expressions, and oral traditions to enhance relevance and relatability.

Participatory design approaches that empower local speakers, elders, and cultural custodians will ensure that technology uplifts rather than erodes indigenous identity.

5.4 Cross-Language Transfer and African NLP Collaboration

The development of AI tools for Yorùbá provides valuable, transferable insights that can significantly benefit other African languages with similar linguistic characteristics, such as Igbo, Ewe, and Fon. These languages, which often share tonal, morphological, and syntactic features within the Niger-Congo and Bantu language families, can leverage advancements in Yorùbá NLP as a foundation for broader cross-lingual innovation. Future prospects in this space include cross-lingual model training that exploits these shared linguistic traits to improve performance and transfer learning efficiency across multiple African languages. Additionally, the formation of pan-African NLP research coalitions will be instrumental in driving progress. These coalitions could facilitate the sharing of datasets, foster technical mentorship, and encourage the development of open-source tools tailored to African language contexts. Beyond collaboration, there is a growing need to develop multilingual language models specifically designed for Africa's linguistic ecology—models capable of zero-shot or few-shot learning that can adapt to lesser-resourced languages with minimal labeled data. Such collaborative and context-aware approaches are key to accelerating the creation of equitable, inclusive, and culturally grounded language technologies that truly reflect the diversity and richness of Africa's linguistic landscape.

5.5 Educational Integration and Capacity Building

a. To sustain innovation in AI for Yorùbá localization, educational systems must be actively involved. Research should focus on:
b. Curriculum development that introduces students to AI, computational linguistics, and Yorùbá language processing.
c. Teacher training programs that utilize localized AI tools for bilingual or multilingual instruction.
d. AI research hubs in universities dedicated to indigenous language technologies, fostering cross-disciplinary collaboration among computer scientists, linguists, and sociologists.

This human-centered development model will help build a pipeline of skilled researchers and practitioners from within the language community. The future of AI-based Yorùbá localization is bright—if pursued with a balance of technical innovation, ethical responsibility, and cultural sensitivity. Sustained progress will depend on collaborative partnerships, policy support, and a shared commitment to digital equity and linguistic inclusion.

6 Conclusion

This study demonstrates that Artificial Intelligence, when thoughtfully applied, can play a transformative role in enhancing digital inclusion for Yorùbá speakers. By localizing AI systems to accommodate the linguistic nuances of the Yorùbá language including tone sensitivity, orthographic integrity, and cultural relevance this research affirms the potential for inclusive technological experiences that empower indigenous communities. The analysis revealed persistent challenges such as data scarcity, tonal misclassification, and limited institutional support, but also highlighted promising outcomes through user satisfaction and task success rates in speech and text interfaces. To address these gaps, the study proposed a suite of solutions encompassing AI model customization, community-driven data annotation, inclusive design principles, and policy advocacy. These strategies not only support the advancement of Yorùbá NLP but also provide a replicable blueprint for other under-resourced African languages. Future work will involve expanding the annotated datasets to include more dialects, developing a centralized open-access linguistic resource hub, and creating benchmark evaluation tasks specifically for Yorùbá NLP. Further exploration into multimodal AI systems integrating speech, text, and vision will also be critical for educational tools, voice assistants, and culturally grounded AI applications across the African continent.

Competing Interests

The authors declare that there are no competing financial, professional, or personal interests that could have influenced the conduct, analysis, or presentation of this research. All interpretations and conclusions presented in this paper are solely the authors' and are based on empirical data and scholarly analysis.

Ethical Approval

This study was conducted in accordance with established ethical guidelines for research involving human participants. Informed consent was obtained from all participants prior to data collection, and all responses were anonymized to ensure confidentiality and privacy. Participation was voluntary, and respondents were free to withdraw from the study at any point without penalty.

References

[1] A. Shaban, "Digitalization and exclusion—digital divides and development," in *Digital Geographies—Theory, Space, and Communities: A Machine-Generated Literature Review*, Springer, 2025, pp. 255–496.
[2] M. Pérez-Escolar and F. Canet, "Research on vulnerable people and digital inclusion: toward a consolidated taxonomical framework," *Univers. access Inf. Soc.*, vol. 22, no. 3, pp. 1059–1072, 2023.
[3] A. P. Wu, S.-L. R. Gonsalves, and D. A. Wagner, "The Impact of Technology on a Multilingual World: Problems and Opportunities," in *ICT and International Learning Ecologies*, Routledge, 2021, pp. 42–60.
[4] T. Onadipe-Shalom, "The Social Media and the Revitalisation of Yorùbá Language and Culture," *NIU J. Humanit.*, vol. 9, no. 4, pp. 69–77, 2024.
[5] S. Karunya, M. Jalakandeshwaran, T. Babu, and R. Uma, "AI-Powered Real-Time Speech-to-Speech Translation for Virtual Meetings Using Machine Learning Models," in *2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS)*, IEEE, 2023, pp. 1–6.
[6] M. R. Costa-Jussà *et al.*, "No language left behind: Scaling human-centered machine translation," *arXiv Prepr. arXiv2207.04672*, 2022.
[7] L. E. Bailey and N. Nyabola, "Digital equity as an enabling platform for equality and inclusion," *Pathfind. Peaceful, Just, Incl. Soc. Cent. Int. Coop. June 2021, https//cic. nyu. edu/resources/digital-equityas-an-enabling-platform-for-equality-and-inclusion*, 2021.
[8] A. Tella and J. Ngoaketsi, "Global Initiatives for Digital Preservation of Indigenous Languages in the Fourth Industrial Revolution," *Digit. Media Preserv. Indig. Lang. Africa Towar. a Digit. Sustain. Soc.*, vol. 113, 2024.
[9] J. Taylor and T. Kochem, "Access and empowerment in digital language learning, maintenance, and revival: a critical literature review," *Diaspora, Indig. Minor. Educ.*, vol. 16, no. 4, pp. 234–245, 2022.
[10] P. Thotad, S. Kallur, and S. Amminabhavi, "Automatic question generator using natural language processing," *J. Pharm. Negat. Results*, vol. 13, pp. 2759–2764, 2022.
[11] B. Manka, T. N. Losenje, and S. G. Mbotake, "The Role of Machine Translation (MT) in the Promotion of National African languages," 2025.
[12] A. K. Kandji, C. Ba, and S. Ndiaye, "State-of-the-Art Review on Recent Trends in Automatic Speech Recognition," in *International Conference on Emerging Technologies for Developing Countries*, Springer, 2023, pp. 185–203.
[13] V. M. Reddy, T. Vaishnavi, and K. P. Kumar, "Speech-to-text and text-to-speech recognition using deep learning," in *2023 2nd international conference on edge computing and applications (ICECAA)*, IEEE, 2023, pp. 657–666.
[14] A. T. Helen, O. I. Uzougbo, T. Rashidat, and T. F. P. Ado-ekiti, "From the Ground Up : Leveraging AI to Transform Education at the School Level in Nigeria," *Tech-sph. Multidiscip. Int. J.*, vol. 2, no. 1, pp. 1–16, 2025, doi: https://doi.org/10.5281/zenodo.15162499.
[15] S. A. Oyelere, "Lexical prosody in Naija." Université de Nanterre-Paris X; University of Ibadan, 2021.
[16] J. Nakatumba-Nabende *et al.*, "Building Text and Speech Benchmark Datasets and Models for Low-Resourced East African Languages: Experiences and Lessons," *Appl. AI Lett.*, vol. 5, no. 2, p. e92, 2024.

[17] A. O. Olowojebutu, I. U. Onwuegbuzie, and K. K. Akomolede, "The Role of AI in Promoting Linguistic and Financial Inclusion: The Lédèe Yorùbá API," *Tech-sph. J. Pure Appl. Sci.*, vol. 2, no. 1 SE-Articles, pp. 1–17, Feb. 2025, doi: 10.5281/zenodo.15258443.

[18] J. Koury-Gaioso, "Navigating Digitalization: Intersectional Perspectives on the Digital Divide and the Use of Digital Tools in the Introduction Programme." University of Agder, 2024.

[19] J. Ding, "Corpus-based Translation Studies: Examining Media Language through a Linguistic Lens," in *SHS Web of Conferences*, EDP Sciences, 2024, p. 1012.