

African Multidisciplinary Journal of Development (AMJD)

Page: 18 – 27

<https://amjd.kiu.ac.ug>

ARTIFICIAL INTELLIGENCE AS A CATALYST FOR THE PRESERVATION OF YORUBA DIALECTAL DIVERSITY: PEDAGOGICAL, LINGUISTIC, AND TECHNOLOGICAL PERSPECTIVES

¹ Folahan Olaniyi Olukayode & ² Abiola Olaide Saheed

¹ Department of Educational Technology Adeyemi Federal University of Education, Ondo

² Department of Yorùbá Adeyemi Federal University of Education, Ondo

Abstract

Yorùbá dialects embody layers of historical memory, cultural practice, and linguistic knowledge that remain largely underrepresented in formal education and digital environments. Although the Yorùbá language continues to be widely spoken, many of its dialectal forms face gradual erosion as schooling, media, and technology favor standardized varieties. This article presents a critical review of how artificial intelligence may contribute to sustaining Yorùbá dialectal diversity when guided by linguistic insight and pedagogical purpose. Drawing on scholarship from Yorùbá linguistics, educational technology, and language documentation, the paper examines the cultural foundations of dialectal variation, the challenges associated with traditional preservation methods, and the emerging role of intelligent digital tools. Particular attention is given to the pedagogical implications of AI-supported documentation, especially its potential to support dialect-inclusive teaching, teacher education, and community-based learning. The review also reflects on ethical and policy considerations, emphasizing the need for culturally grounded approaches that respect linguistic ownership and local knowledge systems. By bringing together educational, linguistic, and technological perspectives, the article argues that artificial intelligence, when thoughtfully applied, can support the continuity of Yorùbá dialects without diminishing their cultural specificity.

Keywords: Yorùbá dialects, Artificial intelligence, Language documentation, Indigenous language education, Educational technology

Introduction

The Yorùbá language occupies a central place in the cultural, intellectual, and historical life of a large population in southwestern Nigeria and the wider Atlantic world (Eichler, 2021). Beyond its role as a medium of everyday communication, Yorùbá serves as a carrier of worldview, ethics, artistic expression, and indigenous knowledge systems. Its dialects such as Ìkálẹ̀, Ifẹ̀, Oòndó, Èkìtì, Àkúrẹ̀ are shaped by geography, history, and patterns of social interaction, represent rich repositories of linguistic and cultural meaning. Yet, in recent decades, many of these dialects have experienced reduced transmission, especially among younger speakers, as formal education, urbanization, and digital communication increasingly privilege standardized varieties and global languages (Idowu & Abbass, 2025).

From a linguistic perspective, dialectal diversity is not a peripheral feature of Yorùbá but one of its defining strengths. Variations in tone realization, lexical choice, and phonological pattern across regions reflect long-standing socio-historical processes and local identities (Ahia et al., 2024). However, these same features that make Yorùbá dialects linguistically rich also make them vulnerable in contemporary settings where educational materials, media content, and digital tools

rarely accommodate such variation. As a result, dialects that are less represented in print or broadcast domains face gradual marginalization, even within communities where Yorùbá remains widely spoken (Foley, 2019).

Educational technology scholarship has increasingly drawn attention to the role of digital tools in supporting language learning, documentation, and cultural continuity. Advances in artificial intelligence, particularly in speech processing and language data analysis, have opened new possibilities for engaging with under-resourced and tone-sensitive languages. Studies outside Africa have shown that intelligent systems, when carefully designed, can assist in capturing spoken language, organizing large bodies of linguistic data, and supporting pedagogical innovation for minority languages (Ghaith, 2024). These developments invite critical reflection on how similar approaches might be adapted to the Yorùbá context without erasing its internal diversity.

Within Nigeria, scholars have begun to emphasize the urgency of aligning technological innovation with indigenous language realities. Tone and prosody, which are central to Yorùbá meaning, present particular challenges for digital representation, yet they also offer opportunities for more nuanced computational modeling. Recent discussions by Yorùbá scholars further stress that any technological intervention must be grounded in linguistic expertise and community knowledge if it is to support, rather than distort, existing language practices (Idowu & Abbass, 2025).

Artificial intelligence, when approached as a supportive rather than substitutive tool, holds promise as a catalyst for the preservation of Yorùbá dialectal diversity. Its relevance lies not only in technical capacity but also in its potential pedagogical value (Ajuzieogu, 2022). AI-supported resources can contribute to dialect-inclusive teaching, enable learners to encounter linguistic variation in structured ways, and support educators in developing context-sensitive instructional materials (Adedeji, 2017). From this standpoint, the intersection of AI, language documentation, and education becomes a site for rethinking how indigenous languages are represented and taught in both formal and informal learning spaces.

Conceptual Clarifications and Scope

A discussion that brings together artificial intelligence, Yorùbá dialects, and educational practice requires careful clarification of key concepts in order to avoid ambiguity and conceptual overlap. From both linguistic and educational technology perspectives, the meanings attached to terms such as dialect, artificial intelligence, documentation, preservation, and revitalization shape not only the direction of inquiry but also the ethical and practical implications of proposed interventions.

Within Yorùbá studies, the notion of dialect goes beyond minor phonetic differences. Yorùbá dialects are structured varieties with systematic distinctions in pronunciation, tone realization, vocabulary, and, in some cases, morphosyntactic patterns. These variations are closely tied to historical settlement patterns, social organization, and cultural practices. Conceptualizing Yorùbá dialects as legitimate linguistic systems, rather than deviations from a standard form, is essential for any meaningful documentation or pedagogical engagement. This position aligns with contemporary sociolinguistic thought, which emphasizes linguistic plurality and resists hierarchies that privilege standardized forms over community-based ones (Adeniyi, 2021).

Artificial intelligence, as used in this article, does not refer to a single technology but to a broad family of computational approaches that enable machines to process, recognize, and generate human language. In educational technology research, AI is often framed as an enabling infrastructure that supports learning, assessment, and knowledge organization, rather than as an autonomous replacement for human expertise (Luckin, 2018). Applied to language contexts, AI includes tools for

speech recognition, audio annotation, pattern analysis, and the organization of large linguistic datasets. For tone languages such as Yorùbá, these tools raise specific questions about accuracy, cultural interpretation, and the limits of automated analysis (Oyewusi, 2024).

The concepts of documentation, preservation, and revitalization, though related, serve different purposes in language-focused scholarship. Documentation refers primarily to the systematic recording and description of language use in its natural settings, often with an emphasis on creating lasting records for research and community reference (Akujobi & Samuel-Enogwe, 2024). Preservation extends this effort by seeking to maintain access to these records over time, particularly in formats that remain usable as technologies evolve. Revitalization, by contrast, is inherently social and pedagogical, involving deliberate efforts to increase language use, visibility, and transmission, especially among younger generations. In the Yorùbá context, revitalization does not necessarily imply language endangerment in the strict sense, but rather addresses the diminishing visibility of dialectal forms within formal education and digital spaces (Zahrasadat et al., 2024).

From an educational technology standpoint, scope is a critical consideration. This review does not attempt to provide technical specifications of AI systems or detailed computational models. Instead, it focuses on conceptual and applied dimensions that connect linguistic knowledge with educational practice. The emphasis is on how AI-supported approaches can assist in capturing dialectal variation, organizing linguistic resources, and supporting dialect-inclusive pedagogy. This orientation reflects concerns raised by Nigerian scholars who argue that technology-driven language initiatives must remain attentive to local educational realities, teacher preparedness, and community participation (Hambardzumyan, 2026).

Geographically and linguistically, the scope of the article centers on Yorùbá dialects spoken within Nigeria, while acknowledging the wider Yorùbá-speaking diaspora. The review prioritizes studies published from 2017 onward in order to reflect current debates in artificial intelligence, language documentation, and educational technology. Both international and Nigerian scholarship are drawn upon, with particular attention to contributions by Yorùbá linguists whose work provides culturally grounded insights into dialectal structure and use.

By clarifying these concepts and boundaries, the article establishes a shared understanding that guides subsequent discussions. This approach ensures that artificial intelligence is examined not as an abstract technological trend, but as a situated tool whose value depends on linguistic sensitivity, pedagogical purpose, and respect for Yorùbá cultural contexts.

Yorùbá Dialectal Diversity: Linguistic and Cultural Foundations

Yorùbá dialectal diversity is best understood as the outcome of long-standing historical, social, and cultural processes rather than as a deviation from an assumed norm. Across Yorùbá-speaking regions, varieties have evolved through patterns of migration, political organization, intergroup contact, and ecological setting. These processes have produced dialects that share a common grammatical core while exhibiting systematic differences in sound patterns, tonal realization, lexical choice, and pragmatic usage. Such variation is not incidental; it reflects localized identities and ways of interpreting social experience (Atolagbe, 2020).

From a linguistic standpoint, tone remains one of the most salient features distinguishing Yorùbá dialects. While the three-level tonal system is shared across varieties, the phonetic realization of tones, tonal sound patterns, and tone–segment interaction vary from one dialect area to another. These differences affect meaning construction in subtle yet significant ways, particularly in oral genres such as praise poetry, chants, and ritual speech. Nigerian linguists have emphasized that tone

should be treated as a central organizing principle in any serious account of dialectal structure, especially in documentation efforts that aim to preserve communicative nuance (Ahia et al., 2024). Lexical variation further illustrates the depth of dialectal diversity within Yorùbá. Words associated with agriculture, kinship, cuisine, and religious practice often differ across regions, reflecting localized knowledge systems and cultural priorities. These lexical distinctions are frequently embedded in idiomatic expressions and proverbs, making them resistant to direct translation into standardized forms without loss of meaning. Jimoh et al. (2025) observes, such expressions serve not only communicative functions but also act as vehicles for moral instruction and communal memory.

Cultural practices provide an essential context for understanding why dialects persist and how they function. Festivals, oral performances, naming practices, and indigenous education systems rely heavily on dialect-specific forms of expression. In many communities, elders transmit historical narratives and ethical values through speech genres that are tightly bound to local varieties. When these genres are removed from their linguistic environment or rendered in generalized forms, their cultural force is diminished (Fagbe, 2024). This observation has direct implications for educational technology, where digital learning materials often prioritize uniformity over contextual richness. The interaction between language and identity is particularly evident in the attitudes speakers hold toward their dialects. While Standard Yorùbá enjoys greater visibility in formal education and media, local varieties continue to signal social affiliation within communities. Recent studies note that younger speakers may understand dialects passively but lack confidence in active use, a pattern linked to schooling practices that marginalize non-standard forms (Akomoledede et al., 2025). This shift underscores the need for approaches that recognize dialects as pedagogical resources rather than obstacles to literacy and academic success.

For educational technologists, these linguistic and cultural foundations highlight the importance of context-aware design. Digital tools that fail to account for dialectal variation risk reinforcing linguistic hierarchies and overlooking community knowledge. Conversely, technologies that are informed by linguistic scholarship and cultural practice can support more inclusive forms of learning and documentation. By grounding technological innovation in the lived realities of Yorùbá speakers, dialectal diversity becomes an asset for education and cultural continuity rather than a problem to be simplified (Ahia et al., 2024).

Challenges in the Documentation and Preservation of Yorùbá Dialects

Efforts to document and preserve Yorùbá dialects face a range of interrelated linguistic, educational, and technological challenges. One of the most persistent difficulties lies in the dominance of standardized Yorùbá in formal domains. School curricula, print media, and broadcast content largely privilege a normalized variety, often treating dialectal forms as informal or unsuitable for academic use. This practice has narrowed the spaces in which dialects are used, recorded, and valued, leading to limited institutional support for their systematic documentation (Ahia et al., 2024).

From a linguistic perspective, the structural complexity of Yorùbá dialects poses additional challenges. Variations in tonal patterns, phonetic realization, and lexical usage require careful, context-sensitive analysis. Traditional documentation methods, which often rely on written transcription alone, struggle to capture tonal nuance and performance-based meanings that are central to many dialects. Nigerian linguists have noted that when tone and prosody are inadequately represented, recorded data may flatten meaningful distinctions and misrepresent actual language use (Atolagbe, 2020).

Resource constraints further complicate documentation efforts. Many dialect-speaking communities lack access to sustained research funding, trained field linguists, or appropriate recording equipment. As a result, documentation projects are often small in scale and short in duration, limiting their long-term value. Even when recordings are made, issues of storage, accessibility, and maintenance arise, particularly in contexts where digital infrastructure is uneven (Jimoh et al., 2025). These constraints underscore the gap between the richness of spoken dialects and the fragility of their recorded forms.

Educational practices also contribute to preservation challenges. In many classrooms, learners are discouraged from using dialectal forms, reinforcing the perception that such varieties lack educational worth. This discouragement affects intergenerational transmission, as younger speakers may understand dialects but avoid active use. Scholars of Yorùbá descent have argued that this pattern reflects broader language ideologies rather than linguistic inadequacy, and that education systems play a key role in shaping attitudes toward dialects (Fagbe, 2024).

Technological interventions, while promising, introduce their own difficulties. Most existing digital tools are designed for global or high-resource languages and do not readily accommodate tonal complexity or dialectal variation. International studies highlight the risk of technological bias, where under-resourced languages are forced into frameworks that do not reflect their structure. For the Yorùbá language, this can result in partial or distorted representations if local linguistic expertise is not integrated into tool development for its dialects (Ahia et al., 2024).

Artificial Intelligence and Indigenous Language Documentation

The documentation of indigenous languages has historically depended on manual recording, transcription, and archiving practices that demand significant time, expertise, and financial resources. While these methods have produced valuable records, they often struggle to keep pace with the rapid social changes affecting language use. Recent developments in artificial intelligence have introduced alternative ways of approaching documentation, particularly for languages that rely heavily on oral transmission. From the combined perspectives of educational technology and Yorùbá linguistics, AI may be understood as an assistive framework that extends human capacity rather than a substitute for linguistic knowledge or community engagement.

In the context of indigenous languages, AI-driven tools have been explored for tasks such as speech-to-text processing, audio segmentation, pattern recognition, and the organization of large collections of spoken data. International studies suggest that these tools can support more efficient handling of linguistic materials, especially where large volumes of audio recordings are involved (Jimoh et al., 2025). However, scholars caution that such tools must be adapted carefully to the structural features of each language. For tone languages like Yorùbá, the accurate representation of pitch movement, rhythm, and contextual meaning remains a major concern (Adeniyi, 2021).

Nigerian researchers have emphasized that indigenous language documentation cannot be reduced to technical capture alone. Idowu & Abbass (2025) argues that without attention to cultural context, AI-supported records risk stripping speech of its social meaning. Proverbs, chants, and ritual utterances common in Yorùbá dialects are not merely strings of words but performances shaped by setting, audience, and shared knowledge. Effective documentation therefore requires collaboration between technologists, linguists, and community members to guide data selection, interpretation, and annotation.

From an educational technology standpoint, AI offers possibilities for transforming documented materials into usable learning resources. Recorded dialectal data can be organized into searchable archives, annotated for pedagogical use, and integrated into teacher education and community

learning initiatives. Zahrasadat et al. (2024) note that when digital language resources are designed with educational goals in mind, they are more likely to support sustained use rather than remain static archives. This perspective aligns with calls by Yorùbá scholars for documentation projects that serve both academic inquiry and community learning needs (Akujobi & Samuel-Enogwe, 2024).

At the same time, concerns about data ownership, consent, and representation remain central. International debates on language technology highlight the risk of extracting linguistic data without adequate benefit to speaker communities. Nigerian scholars echo this concern, stressing that indigenous language documentation should empower speakers and respect local authority over linguistic heritage (Eichler, 2021). In this regard, artificial intelligence should be framed as a participatory tool, guided by ethical principles and cultural accountability.

Overall, the relevance of AI to indigenous language documentation lies not in automation alone, but in its potential to support careful, collaborative, and context-sensitive practices. For Yorùbá dialects, such an approach offers a pathway for combining technological innovation with linguistic depth and educational purpose.

Pedagogical Implications of AI-Supported Dialect Preservation

The integration of artificial intelligence into the documentation and preservation of Yorùbá dialects carries significant pedagogical potential. Beyond mere archival purposes, AI-enabled tools can transform how dialectal knowledge is transmitted, learned, and valued within formal and informal educational contexts. By leveraging computational capabilities such as speech recognition, corpus analysis, and adaptive learning systems, educators can create learning environments that accommodate dialectal variation while supporting literacy, comprehension, and cultural awareness (Luckin, 2018).

One important implication is the possibility of dialect-inclusive teaching. Traditional classroom settings often privilege standardized Yorùbá, inadvertently marginalizing regional varieties. AI-supported resources can provide students with exposure to multiple dialectal forms through interactive audio archives, automated pronunciation feedback, and context-rich digital texts. Such exposure allows learners to recognize the legitimacy of dialects, improving both comprehension and engagement while fostering positive attitudes toward linguistic diversity (Ahia et al., 2024).

Teacher education is another critical area where AI can play a role. Training programs that incorporate AI-driven dialectal corpora enable educators to develop pedagogical strategies that respond to dialectal variation rather than suppress it. This approach aligns with Nigerian scholarship emphasizing the need for teacher preparedness to navigate dialectal complexity in classrooms (Akomoledé et al., 2025).

Moreover, AI tools can facilitate community-based learning initiatives by allowing local speakers to contribute recordings, validate transcriptions, and create culturally grounded educational content. This participatory dimension strengthens intergenerational transmission, ensuring that dialects remain living languages rather than static records (Ajuzieogu, 2022).

Despite these opportunities, caution is necessary. Educational deployment of AI must consider access, infrastructural limitations, and the ethical use of community data. When these concerns are addressed, AI-supported dialect preservation can enrich pedagogy, enhance cultural literacy, and support sustainable language revitalization in ways that are both technologically informed and culturally sensitive.

Technological Perspectives on AI Deployment for Yorùbá Dialects

The deployment of artificial intelligence for Yorùbá dialect preservation requires careful consideration of both technological possibilities and the linguistic intricacies inherent in a tone-

sensitive language. AI applications in this context are not limited to automated transcription or corpus management; they encompass the creation of intelligent systems capable of recognizing, analyzing, and generating dialectal forms while respecting their phonological, morphological, and prosodic characteristics (Adeniyi, 2021).

Digital corpora are central to technological interventions. By collecting audio, video, and textual data from speakers across dialect regions, researchers and developers can construct comprehensive, searchable archives that serve both linguistic research and educational purposes. Annotated corpora that capture tonal variation, lexical specificity, and pragmatic usage allow AI systems to provide more accurate outputs, enhancing both documentation quality and pedagogical applications (Ahia et al., 2024). Importantly, such corpora must be continuously validated by community members to ensure cultural authenticity and prevent the erosion of local knowledge systems (Zahrasadat et al., 2024).

AI-driven tools also facilitate the development of interactive educational technologies. Mobile applications, digital dictionaries, and speech-based learning platforms can integrate dialectal data, offering learners opportunities for active engagement with regional varieties. This approach addresses previous technological limitations that favored high-resource languages, promoting inclusivity and cultural continuity (Akomoledé et al., 2025).

Sustainability and scalability are additional technological considerations. For Yorùbá dialects, infrastructure limitations such as inconsistent electricity, limited internet access, and device scarcity can constrain AI deployment. Cloud-based systems, offline-capable applications, and lightweight processing models offer practical solutions, ensuring that interventions reach underserved communities while maintaining usability and reliability (Oyewusi, 2024).

Finally, ethical and participatory frameworks are integral to technological design. Yoruba-speaking communities should guide decisions regarding data collection, storage, and dissemination. Ensuring informed consent and equitable access fosters trust and aligns technological innovation with the social and cultural objectives of dialect preservation (Ajuzieogu, 2022). When these factors are integrated, AI deployment can move beyond experimental applications, offering scalable, culturally sensitive, and sustainable solutions for preserving Yorùbá dialectal diversity.

Ethical, Cultural, and Policy Considerations

The preservation and documentation of Yorùbá dialects through artificial intelligence are not purely technical endeavors; they are deeply embedded in ethical, cultural, and policy dimensions that shape both process and outcome. Ethical considerations begin with the principle of informed consent and the recognition of community ownership over linguistic data. Yorùbá-speaking communities are not passive subjects but custodians of their dialects, and their participation in data collection, annotation, and validation is essential for culturally sensitive documentation (Eichler, 2021). Without meaningful engagement, AI-driven interventions risk reinforcing extractive practices in which linguistic knowledge is removed from its social context and appropriated without benefit to the speakers themselves.

Cultural integrity represents another central concern. Yorùbá dialects are inextricably linked to oral performance, ritual practice, and social identity. Automated transcription, tone analysis, or digital translation must not flatten or distort these nuances. Scholars have highlighted that proverbs, songs, and ritual utterances often convey meanings that extend beyond lexical content, and these dimensions are difficult to capture using AI without careful human supervision (Fagbe et al., 2024). Preserving dialectal diversity thus entails designing AI systems that respect performative and

pragmatic dimensions, while allowing local communities to guide decisions regarding what, how, and for whom linguistic data are made accessible.

Policy considerations are equally vital. National and regional language policies in Nigeria have historically emphasized standardized forms, limiting formal recognition and resource allocation for dialectal varieties. Integrating AI into language preservation requires policy frameworks that incentivize dialect documentation, support teacher training, and ensure equitable access to digital resources (Akujobi & Samuel-Enogwe, 2024). Policies must also address technological infrastructure, data protection, and long-term sustainability to prevent digital resources from becoming obsolete or inaccessible.

Internationally, debates on AI ethics underscore the risk of linguistic homogenization and algorithmic bias. Low-resource languages such as Yorùbá dialects may be marginalized if AI systems are designed without attention to tonal complexity, cultural context, and community priorities (Jimoh, De Wille, & Nikolov, 2025). Nigerian scholars emphasize that these concerns require policies that embed cultural accountability, participatory design, and local capacity building as central pillars of AI-supported language initiatives (Gaith, 2024).

In sum, ethical, cultural, and policy considerations are inseparable from technological and pedagogical objectives. Successful AI-mediated preservation of Yorùbá dialects depends on frameworks that honor community authority, safeguard cultural meaning, and align technological interventions with supportive policy environments.

Toward an Integrated Framework for AI-Driven Dialect Preservation

Building upon the linguistic, pedagogical, technological, and ethical considerations discussed in this article, it becomes clear that effective preservation of Yorùbá dialects through artificial intelligence requires an integrated, multidimensional framework. Such a framework must align the expertise of linguists, educators, technologists, and community stakeholders, ensuring that AI interventions are both culturally grounded and pedagogically relevant (Ajuzieogu, 2022).

At its core, the framework should begin with community-centered data collection. Indigenous speakers must participate in the selection, recording, and annotation of dialectal material to ensure authenticity and prevent misrepresentation. This participatory approach safeguards cultural knowledge and strengthens intergenerational transmission, while simultaneously generating high-quality corpora suitable for AI applications (Eichler, 2021; Fagbe et al., 2024).

The second component involves linguistic and computational alignment. AI systems must be designed to capture the tonal, prosodic, and lexical nuances of Yorùbá dialects. Unlike standard languages, dialectal forms exhibit subtle variation in pitch, stress, and morphosyntactic structures. Integrating these features into speech recognition, text-to-speech, and analytical tools enhances both documentation fidelity and pedagogical utility (Adeniyi, 2021; Ahia et al., 2024).

A third dimension addresses pedagogical integration. Digital resources derived from AI-enabled documentation should be adaptable for classroom and community learning, supporting dialect-inclusive curricula and teacher training programs. Such resources can reinforce learner engagement, promote positive attitudes toward linguistic diversity, and facilitate practical language use (Zahrasadat et al., 2024; Luckin, 2018).

Finally, the framework emphasizes ethical, policy, and sustainability principles. Guidelines for informed consent, data ownership, and equitable access are necessary to protect community interests. Policy support at local and national levels can ensure infrastructure provision, resource allocation, and long-term maintenance of digital archives (Gaith, 2024; Jimoh et al., 2025).

In combining these dimensions, the framework positions AI as an enabling tool rather than a prescriptive solution. It encourages co-creation, cultural accountability, and pedagogical innovation, offering a holistic model for sustaining Yorùbá dialectal diversity in a rapidly changing technological landscape.

Conclusion

The preservation and revitalization of Yorùbá dialects through artificial intelligence represents a convergence of linguistic scholarship, educational innovation, and technological development. This review has demonstrated that dialectal diversity is both a linguistic asset and a cultural resource, reflecting historical continuity, social identity, and localized knowledge systems. Yet, traditional methods of documentation and preservation face limitations related to resource constraints, accessibility, and the inadequacy of standardized approaches to capture tonal and lexical variation. Artificial intelligence offers promising solutions by enabling efficient speech recognition, corpus management, and interactive learning environments that are sensitive to dialectal nuance. When integrated thoughtfully, AI can support pedagogical innovation, enhance teacher education, and foster community participation, thereby reinforcing both language transmission and cultural continuity. At the same time, technological interventions must be guided by ethical frameworks that respect community authority, ensure informed consent, and protect data ownership.

Ultimately, sustaining Yorùbá dialectal diversity requires a holistic approach that balances linguistic accuracy, technological sophistication, pedagogical relevance, and cultural accountability. An integrated framework rooted in participatory design, policy support, and continuous community engagement can position artificial intelligence not merely as a tool, but as a catalyst for the preservation and revitalization of Yorùbá dialects in the digital age. Such an approach offers a model for other indigenous languages facing similar pressures, demonstrating how modern technology can honor and sustain linguistic heritage without compromising its cultural integrity.

Recommendations

Language documentation projects should prioritize active involvement of Yorùbá-speaking communities in the collection, annotation, and validation of dialectal data. Engaging elders, local educators, and cultural custodians ensures authenticity and enhances intergenerational transmission of dialectal knowledge.

Educational institutions should leverage AI-enabled resources to support dialect-inclusive curricula. Tools such as interactive speech recognition platforms, digital dictionaries, and adaptive learning modules can expose learners to diverse Yorùbá dialects, fostering positive attitudes toward linguistic variation while improving comprehension and oral proficiency.

Developers of AI tools for Yorùbá language should incorporate the language's tonal, prosodic, and lexical complexities into system design to accommodate the peculiarities of its dialects. Collaboration with linguists specializing in Yorùbá ensures that AI outputs maintain semantic integrity, preventing distortion or oversimplification of dialectal forms.

All AI-driven documentation efforts should adhere to ethical standards that prioritize informed consent, community ownership, and equitable access to linguistic resources. Policies must safeguard speakers' rights and ensure that digital materials benefit the communities from which they originate. Government agencies, educational authorities, and research institutions should establish policies and provide infrastructure that sustain AI-assisted dialect preservation. Investments in electricity, internet connectivity, and cloud-based storage systems are essential for the long-term accessibility and maintenance of digital archives.

References

- 1) Adeniyi, K. (2021). A tonal identification of Yoruba dialects. *Dialectologia: revista electrònica*, 28, 1-31.
- 2) Ahia, O., Aremu, A., Abagyan, D., Gonen, H., Adelani, D., Abolade, D., Smith, N.A., Tsvetkov, Y. (2024). Voices unheard: NLP resources and models for Yorùbá regional dialects. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 4392–4409
- 3) Ajuzieogu, U.C (2022). Multimodal Generative AI for African language preservation: A framework for language documentation and revitalization. *Research (RPWCS2024)*.
- 4) Akomolede, K. K., Onwuegbuzie, I. U., & Olowojebutu A. O. (2025). Enhancing digital inclusion through ai-based Yorùbá language localization: challenges, solutions, and future prospects. *Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)*, 2(1), 1-16.
- 5) Akujobi S. O. & Samuel-Enogwe E. C. (2024). The Role of codification in language revitalization: a study of endangered languages in Nigeria. *Journal of Language Studies and Arts (JOLSA)*, 2(3), 52-61.
- 6) Eichler, J. (2021). 'ICH, inequalities and participation: who decides on heritage?'. *International Journal of Human Rights* 25(5) 793–814.
- 7) Fagbe, A., Josiah, G. U. & Ufomba, E. E. (2024). Creation, Preservation and Documentation of Oral Tradition in Ancient Yorùbá and Hebraic Cultures. *Yoruba Studies Review*. 9(1), 127-142.
- 8) Foley, J. Miles (2019, January 6). oral tradition. *Encyclopedia Britannica*. <https://www.britannica.com/topic/oral-tradition>
- 9) Gaith, K. (2024). AI integration in cultural heritage conservation – Ethical considerations and the human imperative. *International Journal of emerging and Disruptive Innovation in Education: VISIONARIUM*, 2(1).
- 10) Hambardzumyan, G. (2026). AI driven vocabulary acquisition for ESL learners: a comparative study of human-duided Vs AI-assisted learning. *European Journal of English Language and Literature Studies*, 14 (1),1-22.
- 11) Idowu & Abbass (2025). Reimagining Yorùbá oral tradition: an exploratory study of AI-Generated storytelling. *Awka Journal of English Language and Literary Studies (AJELLS)*. 12(2), 102-131
- 12) Jimoh, A. T., De Wille, T. & Nikolov, S. N. (2025). Bridging Gaps in Natural Language Processing for Yorùbá: A Systematic Review of a Decade of Progress and Prospects. *Natural Language Processing Journal*, 2(13), 1-43.
- 13) Luckin, R. (2018). *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*. UCL Press
- 14) Oyewusi, W. (2024). AI Literacy in Low-Resource Languages: Insights from creating AI in Yoruba videos. *arXiv*. <https://arxiv.org/abs/2403.04799>
- 15) Tolagbe, E. A. (2020). "A Variationist Study and Mutual Intelligibility Testing of Yoruba and Variants of Ìgbómìnà Dialectsin Kwara State, Nigeria. *Issues in Language & Literary Studies*, 6(1), 137-147.
- 16) Zahrasadat, H., Ubhaya, P., Rezaie Fatima, R., Oluwafikayo, A., Tilanka, C., Aditya, J., Kasun, G., & Mallikarachchige, W. (2024). Preserving dying languages to sustain culture and knowledge: An experiment in utilizing augmented reality to revive the Yoruba language in Nigeria. 1. 1-12.