Drivers of Suba Language Endangerment in Kenya and the Role of Artificial Intelligence (AI) in Mitigation

Hesborn Ondiba, Tokyo University of Science, Japan

The Southeast Asian Conference on Education 2025
Official Conference Proceedings

Abstract
This study investigates the critical factors contributing to the endangerment of the Suba language in Kenya. Primarily spoken by older members of the Suba community, the language faces significant challenges driven by several interconnected factors. A literature review and ethnographic research identified four primary drivers of Suba language decline: the influence of dominant languages, cultural assimilation through integration with Luo communities, geographical dispersion, and economic migration driven by shifts in livelihood opportunities. Proposed AI-driven solutions include digital ecosystems for Suba language learning, virtual cultural repositories, and AI-enabled economic platforms. These tools aim to document Suba's linguistic heritage, foster intergenerational knowledge transfer, and empower the community economically. Additionally, AI-powered community networking platforms and localized broadcasting are proposed to bridge the geographical and cultural gaps among dispersed Suba speakers. The research suggests a strategic framework for using AI to counter these drivers of language decline, prioritizing community collaboration and cultural sensitivity. This approach aims to preserve the Suba language and offers a scalable model for safeguarding endangered languages in Kenya and beyond.

Keywords: Suba, AI, endangerment, revitalization

Introduction

Endangered languages in Kenya, including Suba, face various threats that align with global trends. The dominance of widely spoken languages, economic migration, and cultural assimilation significantly contribute to the erosion of minority languages (Obiero, 2008). Limited resources for documentation and preservation further worsen these issues, leaving many Indigenous languages at risk of extinction. The loss of these languages goes beyond linguistics; it represents the disappearance of unique worldviews, traditional knowledge, and cultural heritage (Rovira, 2008).

The Suba language, primarily spoken by the Suba community in Kenya, is a notable example of this challenge. Representing a significant piece of Kenya's linguistic and cultural heritage, the Suba language is rooted in the history of the Suba people. The Suba community is a heterogeneous Bantu group and ties to the Ganda, Luhya, and Soga tribes (Ndenda, 2019). Historically, the Suba migrated to Kenya, settling around Lake Victoria, particularly on Rusinga and Mfangano islands, where they developed a rich oral tradition that preserved their cultural knowledge, folklore, and identity (Ndenda, 2019). However, the Suba language has experienced a steep decline over time. With an estimated 157,787 speakers, many of whom are older, the language is increasingly at risk as younger generations shift to dominant languages such as Luo and Swahili (KNBS, 2019). This generational language shift raises concerns about losing a vital component of Kenya's cultural identity and diversity.

This study aims to investigate the critical factors driving the endangerment of the Suba language and explore how Artificial Intelligence (AI) can play a transformative role in addressing these challenges. By examining the interplay of linguistic, cultural, and socio-economic factors, the research seeks to identify actionable strategies for preserving the Suba language. In doing so, the study highlights the potential of AI-driven tools to support documentation, analysis, and revitalization efforts, fostering a sustainable approach to safeguarding endangered languages in Kenya and beyond.

Materials and Methods

This study adopts a qualitative approach using secondary data to investigate the drivers of Suba language endangerment and the potential role of Artificial Intelligence (AI) in its preservation. The literature review focused on peer-reviewed articles, books, and reports on language endangerment, particularly on studies on the Suba language and other minority languages in Kenya. Key sources included census data, historical accounts, and government policy documents, providing a comprehensive understanding of the historical, cultural, and socio-political contexts impacting the Suba language. Additionally, the study reviewed existing AI-based tools and frameworks for linguistic documentation, analysis, and revitalization to assess their applicability in addressing the challenges specific to the Suba language.

The collected secondary data were analyzed using a thematic analysis approach. Key themes were identified to categorize the drivers of language decline. These themes were used to establish patterns and connections, highlighting the interplay of these factors and their cumulative effect on the decline of the Suba language. This systematic analysis provided critical insights into the Suba language's challenges and the opportunities for leveraging AI in its preservation.

Key Drivers of Suba Language Endangerment

Table 1 presents the primary factors contributing to the decline of the Suba language, alongside proposed AI-driven strategies to mitigate each issue. The identified drivers include the influence of dominant languages, cultural assimilation, economic migration, and geographical dispersion. Corresponding AI interventions encompass digital language learning platforms, virtual cultural repositories, AI-enabled economic initiatives, and community networking tools. This alignment underscores the potential of AI technologies in addressing the multifaceted challenges facing the preservation of the Suba language.

Table 1: Key Drivers of Suba Language Endangerment and Corresponding AI Mitigation Strategies

Driver	**AI Mitigation**
Influence of Dominant Languages	AI-powered digital ecosystems for Suba language learning (e.g., chatbots, bilingual dictionaries, language learning apps).
Cultural Assimilation	AI-based cultural repositories, transcription tools, and virtual reality (VR) for documenting and showcasing Suba traditions.
Economic Migration and Livelihoods	AI-enabled digital marketplaces, recommendation systems, and livelihood training programs tailored for Suba speakers.
Geographical Dispersion	AI-driven community networking platforms, virtual storytelling sessions, and geotargeted broadcasting in the Suba language.

Influence of Dominant Languages

Luo, Swahili, and English dominance has significantly marginalized the Suba language. During the colonial era, missionary activities were a pivotal factor in this marginalization, as they often relied on Luo-speaking intermediaries, entrenching Luo in religious and educational domains. The perception of Luo as a language of prestige and opportunity compels many Suba families to encourage their children to adopt it, especially in environments dominated by Luo speakers, such as market transactions and public meetings. In regions like Rusinga and Gembe, Luo has almost entirely replaced Suba as the language of everyday communication. Suba is now spoken primarily by older generations and is often reserved for private or symbolic uses, such as rituals and storytelling (Obiero, 2010).

Educational policies that promote English and Swahili as the primary languages of instruction in Kenyan schools further diminish the functional value of Suba. The elevation of English as the language of science, technology, and socio-economic mobility has accelerated Suba's decline by displacing it from educational and formal domains. Consequently, children are introduced to English and Swahili early in their education, leaving Suba largely absent from formal learning environments. Furthermore, the lack of Suba literacy materials and codified instructional resources reinforces its invisibility within institutional contexts. This institutional neglect alienates young speakers, as schools prioritize English and Swahili for formal instruction and Luo for informal socialization (Barasa, 2023).

The prevalence of these dominant languages in education, governance, and public life continues to erode Suba's functional utility, rendering it less relevant in daily interactions and

institutional contexts. Wamalwa and Oluoch (2013) highlight that the shift to dominant languages represents a strategic adaptation by families seeking socio-economic mobility, often at the expense of their linguistic heritage. For many families, the tangible benefits of mastering English or Swahili frequently outweigh the perceived cultural value of preserving Suba.

Cultural Assimilation Through Integration With Luo Communities. The historical process of assimilation, beginning as early as the 19th century through intermarriage, migration, and trade, has brought Suba culture and traditions to extinction. Key elements of Luo culture, including naming systems and marriage customs, have been widely adopted by the Suba, significantly diluting their unique identity (Obiero, 2008). This cultural convergence was further driven by Suba's active participation in Luo-dominated economic and social networks, where trade and intermarriage facilitated the integration of Luo norms and practices (Ndenda, 2019).

In many cases, the assimilation has been so extensive that Suba speakers no longer identify with their linguistic heritage, fully aligning themselves with Luo culture and language. Luo has become the first language of younger generations, marking a profound cultural and linguistic allegiance shift. Some members of the Suba community now self-identify as Luo, with Suba culture surviving only as a relic preserved by older people or showcased during specific events such as funerals or festivals (Omollo & Kingwara, 2024).

The emergence of the term "Luo Abasuba" underscores the depth of this assimilation, describing Suba individuals who have fully integrated into Luo society. This linguistic and cultural absorption is also reflected in historical land migrations and disputes between the Suba and Luo, where geographic proximity necessitated compromise and adaptation to Luo dominance (Ndenda, 2019). As a result, Suba's distinct cultural identity has been gradually overshadowed by the pervasive influence of Luo traditions and practices.

Economic Migration and Changing Livelihood Opportunities. The Suba community traditionally relied on fishing, farming, and boatbuilding for their livelihoods. However, these activities have significantly declined due to economic and environmental changes. Environmental pressures, such as natural resource degradation and overfishing in Lake Victoria, have rendered fishing less viable as an occupation (Ndenda, 2019). Expanding commercial fishing enterprises dominated by non-Suba groups have marginalized traditional fishing practices, forcing many Suba to adopt alternative livelihoods (Obiero, 2008). Conservation initiatives, including fishing restrictions in certain areas to protect biodiversity, have inadvertently displaced Suba fishermen, compelling them to seek other income sources (Odundo, 2016). These economic and environmental shifts have driven the Suba to integrate into Luo-dominated markets, where Luo language and norms prevail, further weakening the use of Suba within households.

Urban migration has emerged as another significant livelihood strategy for the Suba community in response to these rising economic and environmental challenges. Many younger Suba have moved to urban centres, such as Nairobi, due to limited employment opportunities in rural areas, particularly in traditional Suba industries like fishing and subsistence farming (Obiero, 2008). This migration has created a generational divide, with older Suba speakers remaining in rural areas while younger generations lose connection to their linguistic heritage in urban settings (Barasa, 2023).

Children raised in urban environments, where school systems prioritize English and Swahili, often have little to no exposure to the Suba language or culture. This disconnection may lead

to identity crises as they struggle to reconcile their ancestral heritage with their socio-educational environment. Consequently, urban migration has disrupted traditional livelihoods and accelerated the erosion of Suba's cultural and linguistic identity. Ndenda (2019) emphasizes that economic migration has fragmented rural Suba-speaking communities, making it increasingly difficult to maintain a critical mass of active speakers who can preserve and transmit the language.

Geographical Dispersion

The Suba community faces significant challenges due to their geographical spread across areas such as Mfangano, Rusinga, Takawiri Islands, Gwasi Hills, and Migori (Ndenda, 2019). The isolation of these locations, particularly the islands and remote mainland regions, has made communication among Suba groups difficult. This separation has forced them to rely heavily on Luo-speaking neighbours for trade and daily activities, further reducing the use of the Suba language. On Rusinga Island and parts of Gwasi, many Suba people now predominantly use Dholuo, which has become the primary language for governance and social interactions. In areas like Migori, where Suba communities are scattered among Luo-dominated populations, younger generations have ultimately adopted Luo, severing their connection to the Suba language (Omollo & Kingwara, 2024).

This geographical dispersion has also weakened Suba subgroups' cultural and linguistic ties. Physical separation hinders the maintenance of shared traditions and limits opportunities to pass on the language to future generations (Obiero, 2008). As a result, the uneven use of the Suba language across different regions isolates communities. It accelerates its decline by reducing the opportunities for younger generations to learn, practice, and sustain it.

AI Mitigation Strategies

Digital Ecosystems for Suba-Language Learning and Use

AI-powered digital ecosystems offer a promising solution to counter the decline of the Suba language amidst the dominance of Luo, Swahili, and English (Ondiba, 2025a). These ecosystems combine technological tools, cultural preservation strategies, and user-centric design to support language learning and revitalization (Low et al., 2022). Beyond traditional educational tools, they integrate cultural identity to address barriers to language preservation. The revitalization of te reo Māori through AI tools, educational initiatives, and community engagement illustrates how digital ecosystems can empower endangered languages to reclaim their relevance in modern sociolinguistic contexts (Raj, 2024).

AI chatbots can facilitate conversational practice in Suba as virtual language tutors, enabling contextual language learning while incorporating dominant languages like Luo, Swahili, and English (Low et al., 2022). Digital Language Ecosystem (DLE) features like quizzes, forums, and interactive workshops could enhance language learning while promoting problem-solving and creative skills in Suba-specific contexts (Pinto-Llorente & Izquierdo-Álvarez, 2024). These tools help users maintain a connection to Suba even in environments where other languages dominate, offering interactive and adaptive learning experiences.

Translation AI is another critical component of this ecosystem. AI-powered translation tools can bridge linguistic gaps between Suba and dominant languages by creating bilingual dictionaries, subtitling traditional media, and enabling real-time translation during cultural

events or community gatherings (Jafari, 2023). For instance, Microsoft's collaboration with the Nunavut government to develop AI models for Inuktut languages demonstrates the transformative potential of such tools in fostering intergenerational communication and creating multilingual archives (Low et al., 2022). For Suba, these tools could document and preserve linguistic heritage while integrating the language into modern educational and cultural contexts. Additionally, AI-driven translation systems can provide immediate, context-specific feedback to learners, enhancing comprehension and communication skills in multilingual settings (Pinto-Llorente & Izquierdo-Álvarez, 2024). These efforts promote Suba's practical utility and elevate its status, ensuring its presence in educational, cultural, and public domains.

Cultural Repositories and Virtual Reality

One way to address cultural assimilation challenges in the Suba community is by creating AI-powered digital archives. These archives can preserve Suba cultural artefacts, oral histories, and traditional practices, safeguarding them for future generations. As Ondiba (2025b) emphasizes, integrating AI technologies can ensure accurate transcription and annotation of Suba oral traditions and the ethical and secure handling of sensitive Indigenous knowledge. This provides younger generations with engaging and educational access to their cultural roots. Bekele et al. (2018) also highlight how virtual reality (VR) and augmented reality (AR) technologies can enhance cultural preservation by transforming digitized artefacts and data into interactive formats. For example, the Digital Periegesis project used VR to document Greek cultural heritage (Bekele et al., 2018). A similar approach could be adapted for Suba oral histories and practices, ensuring accessibility and interactivity.

Smartphone applications using gamified learning techniques can also play a key role in preserving Suba language and culture. These apps could include interactive features for vocabulary building, storytelling, and participation in cultural activities (Gray, 2023). For instance, gamified scenarios based on traditional practices, like crafting fishing gear or narrating folk tales, could engage users while transmitting cultural knowledge. Inspired by Māori examples, community-driven storytelling apps and role-playing tasks rooted in cultural practices could sustain engagement and encourage intergenerational knowledge transfer (Low et al., 2022).

Gamification strategies, such as rewards for language milestones or role-playing based on Suba traditions, have improved language and cultural retention. Studies indicate that gamified learning creates immersive and enjoyable experiences, which are highly effective for sustaining engagement and promoting long-term cultural preservation (Hamari et al., 2014).

Economic Empowerment Through Language Revitalization

AI-driven strategies inspired by successful rural tourism models can empower Suba speakers economically while preserving their linguistic heritage. One approach is the creation of Suba-centric digital marketplaces, leveraging AI to help local entrepreneurs sell traditional goods such as fishing crafts, textiles, and artisanal products (Semwal et al., 2024). These platforms, modelled after AI-enabled virtual marketplaces that connect tourists with local artisans, could incorporate the Suba language in branding and product descriptions, promoting Suba culture while creating economic incentives for its preservation.

AI-powered recommendation systems could match sellers with global buyers, translating Suba-language content for international audiences and increasing the visibility of Suba's identity.

Research shows that integrating cultural elements into digital marketplaces boosts community participation and economic benefits (Siddiqui et al., 2022). Additionally, participatory decision-making frameworks could involve Suba stakeholders in shaping these platforms to reflect their cultural and financial priorities (Semwal et al., 2024).

AI-driven livelihood training programs can equip the Suba community with modern sustainable fishing and farming skills while incorporating Suba terminologies and traditional ecological knowledge. These programs could use AI tools, such as virtual simulations and predictive analytics, to teach best practices tailored to the community's cultural context (Jayadatta, 2024). For instance, an AI application could guide fishermen in identifying optimal fishing seasons while embedding Suba terminology for species and techniques, seamlessly integrating the language into daily economic activities. Similar initiatives have successfully preserved indigenous knowledge in other communities facing modernization and environmental challenges (Low et al., 2022).

AI innovations in skill development can also help address job displacement caused by modernization. For example, platforms like M-Shamba in Kenya provide personalized agronomic advice and real-time market data, empowering smallholder farmers with actionable insights (Olagunju, 2024). Tailored versions of such platforms could be developed for Suba fishermen and farmers, integrating their linguistic and ecological knowledge into AI-based recommendations. These tools would enhance economic outcomes and ensure the active use and preservation of the Suba language.

Connectivity and Networking

To address the challenges of geographical dispersion, AI-driven community networking platforms can provide dispersed Suba populations with virtual spaces for interaction and collaboration. These platforms could include regional forums, group messaging systems, and event planning tools that promote the use of the Suba language. For instance, AI-powered platforms for revitalizing the Māori language have successfully hosted virtual cultural events, storytelling sessions, and gamified language learning experiences (Sharofova, 2023). A similar platform for the Suba community could host virtual storytelling sessions, allowing members from different regions to share oral traditions in their native language. This preserves Suba cultural heritage and fosters linguistic ties across geographically separated groups. As demonstrated in UNESCO-supported projects, interactive forums within language learning apps can further enhance collaboration between learners and native speakers, fostering a sense of community (Sharofova, 2023).

AI-enabled language broadcasting offers another impactful strategy for connecting dispersed Suba speakers. By leveraging AI tools, localized radio programs and online streaming content can be produced entirely in the Suba language, featuring news, cultural stories, and educational material. Geotargeted broadcasts tailored to areas with high Suba populations ensure content relevance and accessibility. For example, transcription tools like those used in the Cherokee language project can convert oral narratives into scripts, enabling high-quality radio segments and podcasts that feature Suba folklore (Sharofova, 2023). Research shows that broadcasting in indigenous languages enhances language retention and fosters cultural pride in diverse communities (Raj, 2024).

Conclusion

This study explored the critical factors contributing to the endangerment of the Suba language in Kenya and demonstrated how AI-driven solutions can address these challenges. The findings identified four primary drivers of language decline: the dominance of Luo, Swahili, and English; cultural assimilation through integration with Luo communities; economic migration; and geographical dispersion. These interconnected factors have significantly eroded the use and transmission of Suba, placing it at risk of extinction. By leveraging AI technologies, such as digital ecosystems for language learning, virtual cultural repositories, AI-driven economic platforms, and community networking tools, this study highlighted practical strategies for preserving the Suba language. These tools offer scalable solutions to document and promote Suba culture while fostering economic empowerment and connectivity among dispersed populations.

While the proposed solutions show promise, the study acknowledges limitations, such as the need for community buy-in and access to digital infrastructure, which are critical for successful implementation. Future research should focus on refining these AI frameworks and assessing their long-term impact on language preservation and community engagement. Ultimately, this research emphasizes the transformative potential of AI in safeguarding endangered languages. For the Suba community, adopting these strategies supports linguistic and cultural preservation and strengthens their collective identity and resilience in a rapidly evolving world.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP 24K22472.

References

Barasa, D. (2023). Language ideologies, policies and practices within the multilingual Kenyan context. *Journal of Linguistics, Literary and Communication Studies*, *2*(1), 55–62.

Bekele, M. K., Pierdicca, R., Frontoni, E., Malinverni, E. S., & Gain, J. (2018). A survey of augmented, virtual, and mixed reality for cultural heritage. *Journal on Computing and Cultural Heritage (JOCCH), 11*(2), 1–36.

Gray, E. (2023). AI and indigenous language preservation. *Medium.com.* https://medium.com/aimonks/ai-and-indigenous-language-preservation-3005d3567ab0

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE, Hawaii, USA.

Jafari, Z. (2023). The role of AI in supporting indigenous languages. *AI and Tech in Behavioral and Social Sciences*, *2*(4), 11–18. https://doi.org/10.61838/kman.aitech.2.4.2

Jayadatta, S. (2024). A Study on AI-Driven Agricultural Innovations for Rural and Industrial Development in Indian Context. *Journal of Rural and Industrial Development*, *12*(1), 1–16.

KNBS. (2019). Kenya National Bureau of Statistics; *the 2019 Kenya population and housing census.* https://www.knbs.or.ke/wp-content/uploads/2023/09/2019-Kenya-population-and-Housing-Census-Volume-1-Population-By-County-And-Sub-County.pdf

Low, D. S., Mcneill, I., & Day, M. J. (2022). Endangered languages: A sociocognitive approach to language death, identity loss, and preservation in the age of artificial intelligence. *Sustainable Multilingualism*, *21*(1), 1–25.

Ndenda, M. A. (2019). Population movement, settlement, and the construction of society to the east of Lake Victoria in precolonial times: the western Kenyan case. *The East African Review, 52*(1), 83–108.

Obiero, O. J. (2008). Evaluating language revitalization in Kenya: the contradictory face and place of the local community factor. *Nordic Journal of African Studies*, *17*(4), 22–22.

Obiero, O. J. (2010). A case of a mother tongue and another mother tongue in school: efforts at revitalization of Olusuba language of Kenya. *Journal of Third World Studies, 27*(2), 267-291.

Odundo, S. (2016). *Relationship between environmental conservation initiatives and household's livelihood: A case of a community in Suba Sub-county, Homa Bay County* (Doctoral dissertation, University of Nairobi, Kenya). University of Nairobi Repository. http://erepository.uonbi.ac.ke/bitstream/handle/11295/97820/

Olagunju, O. O. (2024). Harnessing Artificial Intelligence for Youth Engagement in Agriculture: Lessons from Global Practices and Prospects for Nigeria. *International Journal of Advance Social Sciences and Education (IJASSE), 2*(2), 83–94. https://doi.org/10.59890/ijasse.v2i2.1490

Omollo, K., & Kingwara, C. (2024). How Abasuba accepted Luo's love and lost their culture. *The Saturday Standard.* The Standard Group PLC. https://www.standardmedia.co.ke/the-standard-insider/article/2001386038/how-abasuba-accepted-luo-love-and-lost-their-culture

Ondiba, H. (2025a). Challenges and Opportunities of AI in Revitalizing and Preserving Endangered Languages in Kenya. *Proceedings of the 2024 16th Asian Conference on Education.* (pp. 1–12).

Ondiba, H. (2025b). Proactive AI-driven cybersecurity for endangered language preservation: Safeguarding the Suba linguistic corpus. *Proceedings of the 4th IEEE International Conference on AI in Cybersecurity (ICAIC)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICAIC63015.2025.10848675

Pinto-Llorente, A. M., & Izquierdo-Álvarez, V. (2024). Digital Learning Ecosystem to Enhance Formative Assessment in Second Language Acquisition in Higher Education. *Sustainability*, *16*(11), 4687–4704.

Raj, A. (2024). Preserving indigenous languages with AI. Tech Wire Asia. https://techwireasia.com/tag/artificial-intelligence/

Rovira, L. C. (2008). The relationship between language and identity. The use of the home language is a human right of the immigrant. *REMHU-Revista Interdisciplinar da Mobilidade Humana*, *16*(31), 63–81.

Semwal, R., Tyagi, N. T. P. K., Pandey, U. K., Dafouti, B. S., & Arya, V. K. (2024). Revitalizing rural tourism in India: A comprehensive framework for AI integration. *Proceedings of the International Conference on Innovation and Regenerative Trends in Tourism and Hospitality Industry (IRTTHI 2024)* (pp. 249–267). Atlantis Press.

Sharofova, S. (2023). The Impact of AI on Endangered Languages: Can Technology Save or Kill? *Texas Journal of Philology, Culture and History*, *25*(1), 52–59.

Siddiqui, M. S., Syed, T. A., Nadeem, A., Nawaz, W., & Alkhodre, A. (2022). Virtual tourism and digital heritage: an analysis of VR/AR technologies and applications. *International Journal of Advanced Computer Science and Applications, 13*(7). 303–315.

Wamalwa, E. W., & Oluoch, S. (2013). Language endangerment and language maintenance: Can endangered indigenous languages of Kenya be electronically preserved? *International Journal of Humanities and Social Science*, *7*(3), 258–266.