



## Nigerian Indigenous Languages Readiness for Inclusivity and Online Accessibility: The Challenges and the Way Forward

Abdullahi Bn Umar<sup>1</sup>, Zainab Mansour Sarki<sup>2</sup>, Godspower Osaretin Ekuobase<sup>3,4</sup>

<sup>1</sup>Department of Computer Science, YMS Federal University of Education, Kano, Kano State  
<https://orcid.org/0009-0001-5137-7831>

<sup>2</sup>Department of Computer Science, YMS Federal University of Education, Kano, Kano State

<sup>3</sup>Department of Computer Science, University of Benin, Benin City, Edo State  
<https://orcid.org/0000-0003-3037-2529>

<sup>4</sup>Department of Computer Science, National Open University of Nigeria, Jabi, Abuja

**ABSTRACT:** Language inclusivity and online accessibility are crucial for language development and survival in today's digital landscape. However, Nigeria's rich linguistic diversity, with over 510 languages, is starkly contrasted by the fact that only three maintain a significant online presence, presenting a concerning scenario. The objective of this study is to investigate Nigeria's efforts towards inclusivity and development of indigenous languages for online visibility and accessibility. To achieve this objective, a content analysis approach was employed to examine online textual literature from Google Scholar and Scopus databases. The literature was coded and categorized based on the study's conceptual framework. The findings revealed that, although some literature reports linguistic development in certain languages, little effort and limited community and government commitment were evident.

**KEYWORDS:** Inclusivity, online content, indigenous language, visibility, accessibility, language development.

### 1. INTRODUCTION

It is not an overstatement that English language has continued to dominate world languages online, represent the global language of globalization and the internet, with over 80% of online content visibility and web accessibility (Abdullahi and Ekuobase, 2024; Akporokah and Okagbare, 2025). This increasing socioeconomic dominance of English language over diverse world's indigenous languages in this digital technology ecosystem, not only limits their online visibility and digital engagement but, more importantly, threatens their socioeconomic survival and cultural heritage, particularly in developing African economies like Nigeria – a country estimated to have over 510 diverse languages, endowed with rich cultural heritage and diversity, but only three of these languages – Hausa in the North, Yoruba in the Southwest, and Igbo in the Southeast – being officially recognized as regional and prominent Nigerian indigenous languages by the government since the late 1960s. Although these languages were designated as regional languages, they are spoken by approximately 50% of the population in this regions. Since then, these languages have continued to grow linguistically, economically, and academically, being part of the Nigerian education curriculum and taught as subjects at both secondary and higher education levels. Furthermore, these languages are making strides in formalization, digitization, translation resources, and online representation, prerequisite for online linguistic content visibility. Language accessibility depends on linguistic resource availability, technology support and other crucial factors. Accessibility encompasses internet connectivity, literacy skills and effective technology utilization ultimately facilitating language development, access, and preservation. Technologies represent the practical application of knowledge in real-life engagement, serving as tools that simplify and facilitate tasks, making their deployment crucial in areas such as language development and access.

Nigeria's linguistic landscape is complex, with over 500 indigenous languages, yet only three – Hausa, Yoruba, and Igbo – have been officially recognized. This disparity underscores the intricate dynamics of language in Nigeria, necessitating an understanding of the complexities arising from regional disparities in language use with multiple dialects, morphological variations, and language shift. Aligning with the efforts by (Akporokah and Okagbare, 2025), this study investigates the preparedness of diverse indigenous Nigerian languages for digitization, online visibility and accessibility, guided by the following research questions:

**RQ1:** What are the challenges in making indigenous languages accessible online?

**RQ1:** How can accessibility be improved for indigenous languages?

**RQ1:** What strategies can enhance inclusivity for indigenous languages in the digital landscape?

Answering these research questions will offer valuable insights into the current state of indigenous language development efforts and the challenges that lied ahead.

### 1.1 THEORETICAL FRAMEWORK

Three theoretical framework – cultural responsiveness theory, inclusive design and sustainable development goals provides a guide to this research study. While cultural responsiveness theory stresses respect for cultural and linguistic diversity in digitization and development of languages, inclusive design encourages digital technology tools and services that meet the needs of diverse users including the under privileged in the society. Sustainable Development Goals on the other hand, supports efforts in aligning digitalization and language development targeted towards equitable and qualitative access opportunities for all without boundaries. These theories provides for assessment on the readiness of Nigerian indigenous languages inclusivity and online accessibility.

### 2. REVIEW PROCEDURE

In this study, a content analysis approach was applied to examine online literature retrieved from Google Scholar and Scopus databases. This method was used not only because it allows for a objective and systematic strategy for data analysis, enhancing reliability and reducing bias, but also for its suitability for analyzing textual-based data with a view to identifying trends in language development particularly in Nigeria. The related literature to the chosen databases – Google Scholar and Scopus used in this study, was identified and collected using database-specific search terms presented as follows: “Language Development AND Nigeria”, “Indigenous languages AND digital presence”, “Language preservation AND online platforms”, “Nigerian languages AND orthography development”, “Digital technology AND language visibility”, “ Nigerian language policy AND implementation”. To allow for easy interpretation of the study’s examination of the relevant literature included in the study, they were classified based on the study’s conceptual framework – Cultural representation, language accessibility, community engagement, challenges and opportunities.

### 3. CONCEPTUAL FRAMEWORK

#### *Language Accessibility*

Over 80% of the online information are preserved in English language, making English the dominant language of the web (Abdullahi and Ekuobase, 2024). The limited access to online information in some languages was noted in Mohamed *et al.* (2019), results from under-representation of those native languages on the web. Shaukat *et al.* (2022) attributed this challenge to little or lack of attention in resource development efforts towards low-resource languages as opposed to high-resource language like English and other few western languages where most of the efforts and attention has been directed over the years. For example, Sahu and Pal (2023) built a text collection in Snaskri native language when it became obvious that such access and retrieval system do not exist for the language owing to linguistic resource scarcity. Similarly, Gautad and Puttkammer (2023) built a labelled dataset for four South African languages to support NLP studies and cross-lingual retrieval evaluation. This efforts are however meant to support or aid fine tuning which is still not a better alternative to original native language resource.

In terms of information access on the web, Information Retrieval System (IRS) remain a critical application that enable search and retrieval of web information. However, 1) the capabilities of this application is limited by information overload on the web – an exponential growth in information that make the search and retrieval more difficult. 2) The linguistic proficiency in the language the information is preserved on the web. With the growth of web content in languages other than English language, relevant information may not necessarily be available in English but in other languages (Kim *et al.*, 2024). Restricting user queries to one language such as English language is insufficient for retrieving relevant information online. Considering language restricted nature of the existing retrieval system (IR) – where the search query, the document collection and the language of the retrieved information are all in English language, automatically excludes non-English speakers from online information access. Although efforts are underway to develop a language independent information retrieval (IRS) as evident the study by Abdullahi and Ekuobase (2024).

#### *Cultural Representation*

Language and culture are two inseparable linguistic concepts. Language is an important component of culture and a veritable instrument with which individual associate or interacts with others, to sharpen their social behavior and subsequent integration into the larger society (Latief *et al.*, 2020). Culture on the other hand is an all-encompassing concept that as to do with belief, values, norms, knowledge, language, and ideology of a giving society to mention a few. However, a description of human culture will be incomplete without language. Language is used to communicate societal norms, values and other cultural components highlighted above from one generation to another. Thus, not only is language a means of transferring culture from one generation to another, but also a means of nurturing culture and cultural identity of man. Needless to say, no human society can co-exist and progress culture wise without an instrument for promoting their cultural heritage. A critical tool that enable these possibilities in every human society is language. Hence, a society whose language is neither written nor spoken regularly and consistently for whatever influence is at the verge of collapse or endangered.

Language Endangerment is a broad concept, depending on the perspective in which it is viewed. Some literature defined endangered language as any language that is at risk or under threat of disappearing or lost (Extinction) as to be endangered language. UNESCO (2003) as cited in (Filipovi' and Pütz, 2016) provided a brief, and comprehensive definition of endangered language, which captures the key issues of this concepts – linguistic usage, intergeneration transfer and number of speakers:

*“A language is endangered when its speakers cease to use it, use it in fewer and fewer domains, and use fewer of its registers and speaking styles, and/or stop passing it on to the next generation”*

In other words, some authors described endangered language as that language which is not spoken as either first language or second language by its natives and referred to it as MORIBUND language. Additional definition as provided under this perspective is hereby presented as:

A language is said to be endangered when it is not learned as first or second language by younger generation, and when the remaining fluent speakers are senior citizens, (Filipovi' *et al.*, 2016). *Moribund* is another name linguistic expert used to describe language that is not learned as mother tongue. This implies, no intergeneration transmission. Thus, leaving the language as critically endangered (just about going into extinction). Language endangerment is a serious global problem which has attracted attention of international communities as it requires a concerted effort from government, United Nations (UN) international organizations UNESCO, World Bank and research community across the globe. In fact, it has been predicted that 90% of the languages worldwide are likely going into extinction in the year 2100 if statistically, a language will go into extinction every two weeks according to Koffi (2021). If this rate at which world languages dies continues, the aftermath will be disastrous. The extent of this menace prompted the declaration of international year of endangered languages in 2019 by UN.

### *Community Engagement*

Many literature have noted language speaker's attitude as one of the significant factors in language endangerment particularly Parent (Hameed, 2022). Parent's attitude has been identified as one of the root cause of language endangerment (Fidyati, 2021; Hameed, 2022). This is true because the primary socialization of a child is traced to the mother (Fidyati, 2021) – one of the basic and first skill a child learn from the mother is language, which is the fundamental component of a culture. No doubt, the first language a child encountered from the mother is known as mother tongue. The assumption is that any language spoken to a child by the mother as first language is regarded as the child's mother tongue (Brock-Utne, 2012). However, if the mother's attitude towards the speaking of her indigenous language is negative, then, the child may grow up with language other than his/her indigenous language vise a visa mother tongue. The consequence of parents not nurturing their cultural norms and values by speaking to their children their indigenous language can results into poor or lack of intergeneration language transmission – one of the strong characteristics of an endangered language (Immidisetti, 2021). Hence, the number of speakers of such language continue to diminish and in the long run go into extinction, if the trend progresses.

A similar scenario is the domination of one language over another in form of colonialism, where the dominant language influence in terms of socio-economic, political opportunities and privileges has caused other minority languages to be seen as inferior and with time consumes them. A dominant language (Dewri and Haokip, 2021) is a language commonly spoken in a particular society by over 85% of its population as a medium for business transactions, inter-personal interactions, politics and in media houses. Dominant language creates an insecure environment for minority languages to survive either because of the fear of linguistic incompetence or inferiority status of their native language (Mohammed, 2020; Fidyati, 2021). Therefore, the smaller the speakers of a particular language the narrow its chances of survival and the brighter the chances of the dominant language consuming the smaller languages (Immidisetti, 2021). The study also noted that, colonialism has negative impact on indigenous language growth and development worldwide due to oppression. Thus, depriving people from living the ways they choose to and practice their cultural norms and values within their society. The nature of oppression depends on economic viability of a particular country (Hart, 2015) – marginalizing and degrading their indigenous languages.

Furthermore, with Globalization, European economy is still benefiting substantially from the globalization influence in terms of cultural empowerment, economic, political and social integration at the detriment of other indigenous world languages. ICT being a driving force of globalization has not favored some individual, group, society, country or region in terms of usage, access due to linguistic barriers – Digital divide and linguistic divide. Hence, there is a sharp divide between those that have access and those denied access. This further exacerbates the threat to and survival of the world's indigenous languages particularly Nigerian languages.

## **4. CHALLENGES AND OPPORTUNITIES**

It is an obvious fact, that Nigeria's linguistic landscape is faced with challenges, from orthography development, language formalization, linguistic divide, to poor language policy framework implementation. This was noted in the study by Adibe and Okeke (2025), that the Nigerian government commitment to language planning and development is biased towards the English language and a few indigenous languages, such as Hausa, Yoruba, and Igbo. This bias is evident even in the implementation of the envisioned policy substituting indigenous languages as a medium of instruction at lower education levels, which has yet to be realized. The inconsistencies in language policy implementation and limited language planning for diverse minority indigenous

languages raise serious concerns about the government's commitment to advancing indigenous languages development, promotion, and revival of its rich cultural heritage. These persistent policy failures arise from a lack of standardization, the influence of globalization, and limited linguistic resource availability, despite the anticipated role of indigenous language play in education and public life. Further prerequisite for language inclusivity and accessibility especially in this digital revolution is the availability of digital resources in indigenous languages.

#### *Digitization and Linguistic Divide*

Digitization or digital age is an information age otherwise known as modern era characterized by access and control of information. A powerful tool that enable this possibility is an integrated technology comprising computer and ICT. The provision and accessibility of this tool has been restricted to some group of people, communities, countries and regions, either as a result of country's income level, literacy level, and non-availability of infrastructure, technological skills, educational background, social support or linguistic competence. An all-encompassing term that describe this phenomenon is commonly known as "digital divide" – defines inequality of access to ICT and the global networks of computers/ internet (Aissaoui, 2020; Zdjeler and Hrustek, 2021). In this information society, uneven participation in and access to the digital technology and internet can be seen as a threat not only to the emergence of social inequalities but more importantly as a global threat to world's indigenous languages endangerment. When viewed from the linguistic perspective we have "linguistic divide" – where ICT and internet accessibility is hindered by linguistic barrier. The question is, how can a typical monolingual native, though, not illiterate in their respective native languages participate and benefits from the potentials offered by these great technologies (ICT and computer or internet) when the language of internet and the web resources remain dominated by English language (Ugwu, 2020).

The above two related concepts, *digital divide* and *linguistic divide* respectively describes social problems and can be summarized under one heading – "Social exclusion" (Zdjeler and Hrustek, 2021). This implies an individual, group or society who in one way or the other lacks opportunity or privilege to participate or enjoy all the social benefits derivable from ICT and the internet technologies. Furthermore, viewing the impact of globalization from the perspectives of socio-cultural, political and economic empowerment – an indicator of economic growth and development of every nation, we can infer that globalization is more of western culture empowerment as English language is still maintaining dominance as global and internet language. In view of the foregoing, a number of African countries particularly Nigeria colonized by the west (British) speaks English language not only as its lingua franca, but also as language of colonization. The consequence of this development is the relegation (losing prestige particularly to the dominant language) of status of the world's indigenous languages especially in a multi-cultural and multilingual region like Africa (Immidisetti, 2021). Thus, this further conditioned citizens of the British colony to learn English language as second language being a language of instruction in both pre-primary, primary, secondary and higher education levels. This of course, is no doubt a major threat to the continued existence and survival of the indigenous languages globally (Zdjeler and Hrustek, 2021). However, the rate at which this threat to indigenous language survival progresses from one stage to another will lie solely on intergeneration transfer of such language – from older generation to younger generation. Furthermore, it was understood from the study of Shikali and Mokhosi (2020) that language modelling is one basic requirement for language resource development task and it relies heavily on adequate pre-trained dataset for quality representation of words for NLP task. However, most of the under-represented languages suffered similar setback of scarce pre-trained dataset necessary for NLP research studies. Without any serious motivation and research efforts at developing a linguistic tool such as large-scale pre-trained datasets, NLP research studies in under-resource languages may suffer grave setback (Cunliffe *et al.*, 2022). This may impact negatively on linguistic and cultural diversity as well as the economic development of the affected languages in terms of online trade in the global digital economy (Jeon *et al.*, 2021).

#### *Opportunities for Language Growth*

Digital technology platforms such as social media, and corpus development for indigenous languages could significantly contribute to promotion and ultimately leveraging the preservation of Nigerian linguistic and cultural heritage. Community participation through grassroots support in the form of cultural festivities, listening to local music, movies, and restricting the family from speaking any other language other than the mother-tong within themselves at home, can motivate intergeneration transfer and community participation. Overall, more opportunities can be enjoyed with government support through language policy that promotes multilingual education, language revitalization, and cultural activities can leverage Nigeria's linguistic heritage preservation.

## **5. THE WAY FORWARD**

Language Revitalization (LR) basically affects the languages that are not frequently spoken in everyday interactions and those that are spoken occasionally within a restricted domain (Bromham *et al.*, 2021). Language restoration could be achieved through teaching, dissemination, documentation and improved intergeneration transmission. LR, restoration, revival, reversal, reclamation, and documentation are all related concepts that described the processes of rescuing endangered languages from extinction (Bromham *et al.*, 2021). Language extinction is eminent if no serious efforts are invested in the revitalization or restoration of multiplicity of endangered world indigenous languages, especially, as the English language is expanding to assume the global language due to

colonization and globalization which is capable of catapulting several indigenous languages into extinction (Bromham *et al.*, 2021). As opined by Austin and Sallabank (2013), language speakers and community of language speaker's attitude plays a major role when it comes to language revival and preservation. An instance of this, is a poor intergeneration transmission by the older generation, perhaps because of negative attitude towards their indigenous language especially when they find themselves in the mist of dominant language speakers (Hameed, 2022). They feel inferior or ashamed to speak their language for fear of unequal treatment, attention, security, stigmatization, socio-economic benefits or simple because their controlled politically and economically by the dominant language (Koffi, 2021). This negative attitude diminished linguistic value and detrimental to an indigenous language survival. A concept that characterized this behavior by the minority language speakers is known as "linguistic insecurity" or "linguistic self-hatred" – a scenario where the minority speakers themselves succumb to believe of dominant language speakers that their language is truly inferior. In other to achieved LR, indigenous languages speakers and their community, need to develop a positive attitude towards their language and as a matter of fact encourage younger generations by disseminating their cultural norms, values and heritage by organizing cultural festivals to showcase their reach cultural endowment.

Intergeneration transmission being one of the key indicators of language endangerment should be given serious attention by way of discouraging senior citizens, particularly parents, from not speaking to their children at home any language other than their native language. As doing so, will not only boost the morale of the younger generations, but will also regulates and preserves their language and cultural identity, a system and treasure of knowledge (Hameed, 2022). Hence, the survival of any language lies not only on the community and speakers of the language but whether or not it is passed to the next generations and progresses as such (Hameed, 2022). Language documentation is noted in (Bromham *et al.*, 2021) as an alternative measure for indigenous language preservation, maintenance for reference and historic purposes. The essence is to preserve linguistic and cultural treasure of the minority community should the speakers of the language be wiped out in the event of natural disaster such as tsunami, earth quake, pandemic or war. The survivors from any of these disasters could take advantage of the historic documented reports as a reference for revitalizing their cultural heritage. However, besides time consuming of this approach, a documentarian may lack linguistic fluency of the language to be documented and hence, the situation worsened (Prud'Hommeaux, *et al.*, 2021). This has been the common scenario in the case of endangered languages documentation.

## 6. CONCLUSION

The readiness of Nigerian indigenous minority languages for inclusivity and online accessibility requires a concerted efforts and targeted strategies to address the unique challenges hindering their growth and development. By developing inclusive language policies, investing in indigenous language development, and implementing community support initiatives, inclusivity and accessibility for diverse indigenous Nigerian languages become promising. Ultimately, this can contribute to preserving the country's rich cultural heritage and enhancing socio-cultural equity and cohesion.

## REFERENCES

1. Abdullahi, U. B. and Ekuobase, G.O. (2024). A Lingual Agnostic Information Retrieval System. The Scientific World Journal Volume 2024, Article ID 6949281, p37, <https://doi.org/10.1155/2024/6949281>.
2. Adibe, O.P. and Okeke, T.O. (2025). Language planning and policy implementation in Nigeria: Challenges and the way forward. *Icheke Journal of the Faculty of Humanities*, 23(1), 213 – 234.
3. Aissaoui, N. (2020). The Digital Divide: a Literature Review and Some Direction for Future Research in Light of COVID-19, *Global Knowledge, Memory and Communication: Advanced Online Publication*, <https://doi.org/10.1108/GKMC-06-2020-0075>.
4. Akporokah, O.J. and Okagbare, V.E. (2025). Digital media interaction and indigenous language preservation in Delta state tertiary institutions. *RKGP Journal of Social Sciences and Arts*, 1(1), 01 – 10, <https://doi.org/10.32001/RKGP-JSSA2025.011>.
5. Austin, K. P. and Sallabank, J. (2013). Endangered Languages: An Introduction, *Journal Multilingual and Multicultural Development*, 34(4), 313–316.
6. Blench, R. (2020 Eds.). *An Atlas of Nigerian Languages*, McDonald Institute for Archaeological Research University of Cambridge, United Kingdom, pp. 198.
7. Bromham, L., Dinnage, R., Skirgord, H., Ritchie, A., Cardillo, M., Meakins, F., Greenhill, S. and Hua, X. (2021). Global Predictors of Language Endangerment and the Future of Linguistic Diversity, *Nature Ecology and Evolution*, 25, 21, 1 – 23
8. Brock-Utne, B. (2012). Language and Inequality: Global Challenge to Education, *A Journal of Comparative and International Education*, 42(5), 1– 21.
9. Cunliffe, D., Vlachidis, A., Williams, D. and Tudhope, D. (2022). Natural language processing for under-resourced languages: developing a Welsh natural language toolkit.

10. Dewri, P. and Haokip, M. D. (2021). Language Endangerment: A Case Study of Tiwa, an Endangered Language of North East India. *Journal of English Language and Literature*, 8(3), 81 – 93.
11. Fidyati, I. (2021). The Factors that Influence the Language Shift among Parents, *Journal of English Language and Education*, 6(2), 202 – 210.
12. Filipovic, L. and Putz, M. (2016). *Endangered Languages and Languages in Danger: Issues of Documentation, Policy, and Language Right*, John Benjamins Publishing Company, Amsterdam.
13. Hameed, A. (2022). Mother, Mother Tongue, and Language Endangerment Process: An Exploratory Study. *Theory and Practice in Language Studies*, 12(4), 726 –735.
14. Hart, A. M. (2015). *International Encyclopedia of the Social and Behavioral Sciences*, 804 – 810.
15. Immidisetti, S. (2021). *The State of Language, Endangerment, and Policy in India: A Forking Path*. Honors College Thesis, 325pp.
16. Kim, J., Chung, S., and Chi, S. (2024). Information retrieval from multilingual construction documents using pretrained language models. *Journal of Construction Engineering and Management*, 150(6), <https://doi.org/10.1061/JCEMD4.COENG-14273>.
17. Koffi, E. (2021). Language Endangerment Threatens Phonetic Diversity. *Acoustical Society of America*, 17(2), 23– 31.
18. Latief, M. A., Saleh, N. J. and Pammu, A. (2020). The Effectiveness of Machine Translation to Improve the System of translating Language on Cultural Context. *Conference Series Earth and Environmental Science* 575, 1– 10.
19. Mohamed, E., Elmougy, S. and Aref, M. (2019). Towards multi-lingual based on internet linguistic diversity measurement. *Ain Shams Engineering Journal*, 10(3), 489-497.
20. Mohammed, A. and Abdalla, M. (2020). English Language and Globalization. *International Journal of Novel Research in Education and Learning*, 7(1), 5 –11.
21. Prud'hommeaux, E., Jimerson, R., Hatcher, R. and Michelson, K. (2021). Automatic Speech Recognition for Supporting Endangered Language Documentation, *Language Documentation and Conservation*, 15, 49 – 513.
22. Shaukat, S., Shaukat, A., Shazad, K. and Daud, A. (2022). Using TREC for developing semantic information retrieval benchmark for Urdu. *Information Processing and Management*, 59(3), <https://doi.org/10.1016/j.ipm.2022.102939>.
23. Ugwu, E. O. (2020). English Language, Globalization and the Future of the Nigerian Indigenous Language. *European Scientific Journal*, 16(14), 1–38.
24. Zdjelar, Z. and Hirustek, Z. N. (2021). Digital Divide and E-Inclusion as Challenges of the Information Society – Research Review. *JIOS*, 45(2), 601– 638.