Development of AI-Powered Chatbot for Learning Igbo Language

[*]Ekwealor, Oluchukwu Uzoamaka[1], Betrand, Chidi Ukamaka[2], Ezuruka, Evelyn Ogochukwu[3], Okechukwu, Ogochukwu Patience [4], Uchefuna, Charles Ikenna [5]

[1,3,4]Department of Computer Science, Nnamdi Azikiwe University, Awka, Nigeria

[2]Department of Computer Science, School of Information and Communication Technology,Federal University of Technology, Owerri, Nigeria,

[5]Department of Computer Science, Federal Polytechnic Oko, Nigeria

[*]Corresponding Author

DOI: https://doi.org/10.51244/IJRSI.2026.1304000107

Received: 05 April 2026; Accepted: 10 April 2026; Published: 04 May 2026

ABSTRACT

This study presents the development of an intelligent chatbot system designed for learning the Igbo language, one of Nigeria's major indigenous languages. In the face of globalization and urbanization, which increasingly threaten the survival of African indigenous languages, there is a pressing need for innovative and accessible tools that support both language acquisition and cultural preservation. The proposed chatbot leverages Natural Language Processing (NLP) and conversational Artificial Intelligence (AI) to deliver an interactive and user-friendly learning platform. The system provides personalized learning experiences through conversational interactions, covering key linguistic components such as vocabulary, grammar, and everyday expressions. By simulating natural dialogue, the chatbot enables users to practice reading, writing, and comprehension skills in a supportive, low-pressure environment. Its architecture comprises natural language understanding modules, a knowledge base containing Igbo linguistic rules and vocabulary and a response generation component that produces contextually relevant outputs. This research contributes to the advancement of Computer-Assisted Language Learning for underrepresented languages and offers a scalable framework for developing similar tools for other indigenous African languages. Ultimately, the system promotes broader access to Igbo language education while supporting efforts to preserve its cultural heritage.

INTRODUCTION

Language is a structured system of conventional spoken, written, or symbolic communication through which individuals, as members of a social group, express ideas and participate in cultural exchange (Crystal, 2022). It serves as a primary medium for communication, cultural transmission, and identity formation.

Indigenous languages are among humanity's most valuable yet vulnerable assets. They foster a sense of identity and belonging while preserving traditional knowledge, cultural practices, and worldviews developed over generations (Hinton et al., 2022). In recent decades, global efforts have intensified toward documenting, revitalizing, and preserving endangered languages due to their intrinsic value and the irreversible loss associated with their extinction.

In a globalized world, the preservation and promotion of indigenous languages have become increasingly critical. Dominant languages often overshadow local dialects, a phenomenon referred to as linguistic imperialism, which threatens cultural diversity and the wealth of knowledge embedded in indigenous languages (Phillipson, 2022).

Nigeria exemplifies remarkable linguistic diversity, with over 220 million people and more than 500 indigenous languages across its ethnic groups (Blench, 2022). Among these, Igbo is one of the three major languages, alongside Hausa and Yoruba, serving as a regional lingua franca (Adegbite & Ogunsiji, 2024).

Predominantly spoken in southeastern Nigeria—Abia, Anambra, Ebonyi, Enugu, and Imo—the Igbo language is not only a communication tool but also a medium for transmitting cultural values, traditional practices, philosophical ideas, and communal knowledge among over 25 million speakers worldwide (Emenanjo, 2023).

Despite its significance, Igbo faces considerable challenges in the modern digital era. Globalization, urbanization, and technological advancement have created a sociolinguistic environment in which younger generations are increasingly disconnected from their linguistic heritage (Okwu & Taruvinga, 2021). The dominance of English in Nigeria's education system—rooted in colonial history and sustained post-independence—has further marginalized indigenous languages in formal learning contexts (Bamgbose, 2022). This has contributed to "subtractive bilingualism," where proficiency in a dominant language occurs at the expense of the native language (Cummins, 2021).

Recent studies indicate a decline in Igbo proficiency among urban youth, many of whom prefer English even in informal settings (Nwaozuzu, 2023). This issue is more pronounced in diaspora communities, where exposure to the language is limited and structured learning opportunities are scarce (Achebe & Obioma, 2022).

Additionally, the digital divide exacerbates the situation, as most technological platforms and innovations predominantly support global languages, leaving Igbo underrepresented (Kornai, 2023). This lack of representation discourages engagement among digitally inclined youth, accelerating language decline (Soria & Martinez, 2024).

However, advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) present new opportunities for language learning and preservation (Zawacki-Richter et al., 2022). Chatbots, in particular, have emerged as effective tools for interactive language education, offering personalized, accessible, and engaging learning experiences (Fryer et al., 2022).

Research highlights the effectiveness of chatbots in providing immediate feedback, reducing language anxiety, enabling continuous practice, and creating a non-judgmental learning environment (Huang et al., 2023; Kim & Lee, 2022). Furthermore, their conversational nature aligns with natural language acquisition processes, emphasizing learning through meaningful interaction (García-Martínez et al., 2024).

LITERATURE REVIEW

Language Learning

The process of acquiring or improving proficiency in a second or foreign language has traditionally relied on classroom instruction, interaction with native speakers, and repetitive practice. However, recent pedagogical approaches increasingly incorporate digital tools that enable flexible, autonomous, and personalized learning. Among these, chatbots have emerged as effective instructional tools capable of simulating human conversation, providing real-time feedback, and facilitating continuous interaction beyond classroom constraints. Natural Language Processing (NLP), a subfield of AI focused on enabling machines to understand and generate human language, underpins these innovations. Advances in large language models (LLMs), machine learning, and adaptive systems allow chatbots to interpret learner input, detect errors, and tailor responses effectively. Well-designed chatbots have been shown to enhance linguistic competence while improving learner engagement, motivation, and self-regulated learning.The integration of chatbot technology into indigenous language education represents a critical intersection of technological innovation and cultural preservation (Adelani et al., 2024).

Natural Language Processing for Low-Resource Languages

Low-resource languages such as Igbo face significant computational challenges due to limited digital data and resources. Despite their cultural and historical importance, these languages encounter barriers to technological development and preservation (Chen et al., 2024). Igbo, with over 25 million speakers, presents additional complexities due to its tonal system and dialectal variations.

Recent advancements, including community-driven initiatives like the IgboAPI platform, have begun addressing these limitations by providing essential linguistic datasets (Ogbonna & Ikechukwu, 2024). Developing NLP systems for Igbo requires addressing its unique linguistic characteristics. Its agglutinative morphology complicates tokenization and analysis, while its tonal nature necessitates specialized approaches to text and speech processing.Efforts such as the IgboAPI Dataset, which incorporates multi-dialectal data, represent significant progress in enabling robust NLP applications for the language.

Conversational AI in Educational Contexts

The use of chatbots in education has expanded significantly, particularly due to their ability to support multimodal communication - text, speech, visuals, and gestures (Kim et al., 2023; Kuhail et al., 2022). These features enhance learning effectiveness, especially in language acquisition.

Educational chatbots function as virtual tutors, promoting language skills and communicative competence while offering continuous access beyond traditional classroom hours (Huang et al., 2022; Hwang & Chang, 2021; Zhang et al., 2023).

The application of chatbots in language learning has evolved from simple rule-based systems to sophisticated, context-aware AI systems powered by machine learning and deep learning techniques.

Review of Related Works

Fryer *et al. (*2020) in his research, opined that chatbots can effectively sustain learner engagement, offering consistent availability and reducing social anxiety.

Kuhail *et al*. (2023) reported moderate-to-high effectiveness of educational chatbots, highlighting personalization and feedback as key success factors.

Similarly, Zhang et al. (2024) showed that LLM-based chatbots significantly enhance vocabulary learning through contextual presentation and immediate feedback.

The development of speech recognition systems for African languages has accelerated in recent years, although considerable challenges persist. Siminyu (2021) provided a comprehensive overview of Natural Language Processing (NLP) resources for African languages, emphasizing the critical issue of data scarcity that constrains the advancement of speech recognition technologies. The study highlighted that most African languages lack sufficiently large and diverse speech corpora required to train robust Automatic Speech Recognition (ASR) systems.

The NaijaVoices project, as documented by Adebara et al. (2023), represents a notable initiative aimed at mitigating speech data limitations for Nigerian languages. The project produced approximately 500 hours of speech datasets for Igbo, Yoruba, and Hausa, thereby offering essential resources for ASR development. Nevertheless, the authors acknowledged that these datasets remain inadequate for training state-of-the-art ASR systems comparable to those available for high-resource languages.

Masakhane, a grassroots organization dedicated to advancing NLP for African languages, has made significant contributions to speech recognition research. Abbott and Martinus (2021) documented the organization's efforts to develop ASR systems across multiple African languages, underscoring the value of community-driven approaches and the integration of local expertise in addressing language-specific challenges.

Tonal Language Processing in Computational Systems

The computational processing of tonal languages introduces unique complexities that have been explored by several researchers. Do *et al.* (2019) proposed a comprehensive framework for Vietnamese tone recognition, demonstrating the effective integration of tonal features into ASR systems. Their approach employed multi-task

learning to jointly model segmental and tonal characteristics, resulting in notable improvements in recognition accuracy.

Similarly, research by Michaud on Mandarin (2020), Chinese tone recognition provided valuable insights into acoustic modeling for tonal languages. The study revealed that incorporating tonal information into deep neural network architectures significantly enhances recognition performance, particularly in conversational contexts. These findings are highly relevant to Igbo language processing, given the shared tonal properties between the languages.

Bontogon (2021) designed a mobile application for learning the Cebuano language, integrating both speech recognition and speech synthesis functionalities. The system achieved a speech recognition accuracy of 78% and demonstrated measurable improvements in learners' pronunciation and vocabulary acquisition.

The IgboAPI project, documented by Okafor and Udechukwu (2023), sought to address the scarcity of accessible lexical resources for Igbo language learners. The project introduced a comprehensive dictionary API with pronunciation support; however, it lacked interactive conversational features. The authors suggested that integrating chatbot functionality could significantly enhance the usability and pedagogical value of such resources.

Recent work by Onyenwe et al. (2022) on Igbo corpus development has contributed valuable linguistic data for computational applications. Their corpus comprises over 100,000 annotated words with morphological information, providing a strong basis for training language models. However, the absence of speech data limits its applicability in speech-enabled systems.

Bibauw et al. (2022) proposed a comprehensive evaluation framework that combines quantitative metrics, such as improvements in language proficiency, with qualitative measures, including user engagement and experience.

Furthermore, Chen et al. (2024) emphasized the importance of culturally appropriate evaluation strategies for AI-driven language learning systems. Their findings indicate that assessment frameworks designed for widely spoken languages may not adequately capture the effectiveness of systems tailored to indigenous languages.

Operation of the System

The proposed Igbo language learning chatbot is an AI-driven interactive system designed to improve users' comprehension and communication skills. It allows users to input words, phrases, or sentences and receive accurate translations, corrected spellings, and contextual examples.

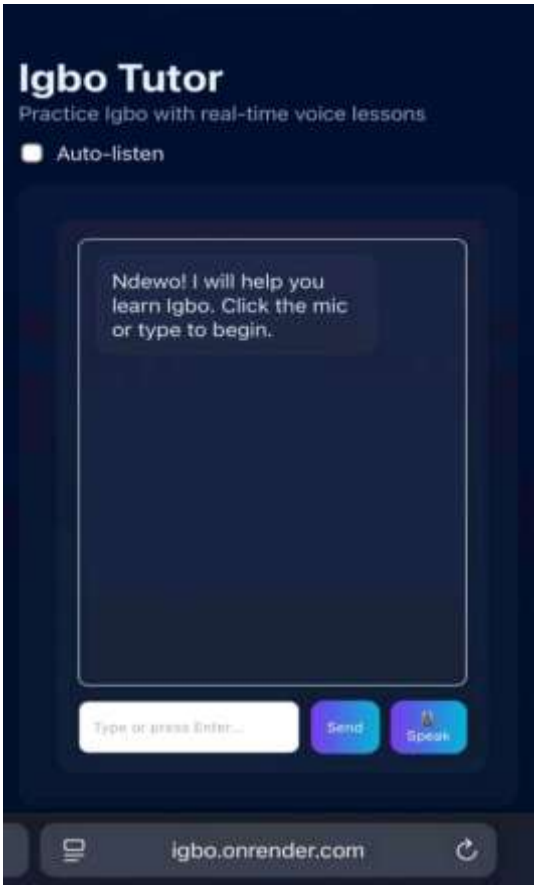
The system includes a task-based learning module that provides short exercises to assess comprehension. By leveraging NLP techniques, the chatbot interprets user input and delivers intelligent feedback.

The platform bridges the gap between technology and indigenous language education by offering an accessible and efficient learning environment. Its dialogue-based design supports natural interaction, enabling users to practice reading, writing, and comprehension in a low-pressure setting. Key features include interactive exercises and real-time feedback.

The system architecture integrates natural language understanding modules, a comprehensive knowledge base of Igbo linguistic rules, and a response generation engine. It also incorporates a rich Igbo language corpus with phonetic transcriptions to address tonal nuances. Evaluation results indicate that the system effectively engages users and supports progressive language acquisition.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue IV April 2026

System Sample Output



CONCLUSION

This paper successfully demonstrates the feasibility and potential of using modern artificial intelligence technologies to preserve and promote indigenous African languages through digital learning platforms. The developed Igbo language learning chatbot represents a significant contribution to language technology for Igbo, addressing the critical need for accessible and interactive language learning tools.

REFERENCE

1. Achebe, C. N., & Obioma, K. R. (2022). Language maintenance among Igbo diaspora communities: Challenges and opportunities. Journal of African Diaspora Languages, 8(2), 45-62.
2. Adebara, I., Muhammad, S. H., Abdulmumin, I., Awokoya, A., Adelani, D. I., Ibrahim, A., AbdulKabir, M., Bakare, T., Ahmad, I. S., & Ruder, S. (2023). NaijaVoices: A Nigerian English speech corpus for speech recognition research. Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2541-2556.
3. Adegbite, W., & Ogunsiji, A. (2024). Nigerian multilingualism in the digital age: Policy implications and community responses. Language Policy and Planning in Africa, 15(1), 78-95.
4. Adelani, D. I., Alabi, J., Fan, A., Kreutzer, J., Shen, X., Reid, M., & Ruder, S. (2024). The IgboAPI dataset: Empowering Igbo language technologies through multi-dialectal enrichment. arXiv preprint arXiv:2405.00997.
5. Bamgbose, A. (2022). Language and exclusion: The consequences of language policies in Africa. Multilingual Matters.
6. Bibauw, S., François, T., & Desmet, P. (2022). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of chatbot-assisted language learning. Computer Assisted Language Learning, 35(4), 827-863.
7. Blench, R. (2022). The classification of Nigerian languages: Current status and future directions. Kay Williamson Educational Foundation.
8. Chen, L., Wang, M., & Johnson, K. (2024). Opportunities and challenges of large language models for low-resource languages in humanities research. arXiv preprint arXiv:2412.04497.
9. Crystal, D. (2022). The Cambridge encyclopedia of language (4th ed.). Cambridge University Press.
10. Cummins, J. (2021). Rethinking the education of multilingual learners: A critical analysis of theoretical concepts. Multilingual Matters.
11. Emenanjo, E. N. (2023). Igbo language in the 21st century: Challenges and opportunities. Africana Publishers.
12. Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2022). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. Computers in Human Behavior, 135, 107356.
13. García-Martínez, I., Tadeu, P., & Belmonte, J. L. (2024). Chatbots in language learning: A systematic review of research, applications, and challenges. Educational Technology Research and Development, 72(2), 245-278.
14. Hinton, L., Huss, L., & Roche, G. (Eds.). (2022). The Routledge handbook of language revitalization (2nd ed.). Routledge.
15. Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. Journal of Computer Assisted Learning, 38(1), 237-257.
16. Hwang, G. J., & Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. Interactive Learning Environments, 31(8), 4099-4112.
17. Kim, N. Y., Cha, Y., & Kim, H. S. (2023). The use of chatbots in university EFL settings: Research trends and pedagogical implications. Frontiers in Psychology, 14, 1131506.
18. Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2022). Interacting with educational chatbots: A systematic review. Education and Information Technologies, 28(1), 973-1018.
19. Nwaozuzu, G. I. (2023). Urban youth and Igbo language vitality: A sociolinguistic study of language attitudes in southeastern Nigeria. International Journal of the Sociology of Language, 2023(279), 67-85.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue IV April 2026

20. Ogbonna, C. K., & Ikechukwu, M. A. (2024). Design and implementation of an automated web-based Igbo text analyzer using natural language processing (NLP) tools. ResearchGate Publication, 384461055.
21. Okafor, C. N., & Udechukwu, A. (2023). IgboAPI: A comprehensive lexical resource for Igbo language learning applications. African Journal of Information and Communication Technology, 19(3), 78-95.
22. Okwu, O. I., & Taruvinga, A. (2021). Language shift among African immigrant communities: The case of Igbo speakers in South Africa. Language Matters
23. Onyenwe, I. E., Uchechukwu, C., & Hepple, M. (2022). A large annotated Igbo corpus for natural language processing. Language Resources and Evaluation, 56(2), 487-520.
24. Phillipson, R. (2022). Linguistic imperialism (3rd ed.). Oxford University Press.
25. Zhang, B., Wu, L., & Anderson, T. (2023). Integrating chatbots in education: Insights from the chatbot-human interaction satisfaction model (CHISM). International Journal of Educational Technology in Higher Education, 20(1), 1-25.
26. Soria, K. M., & Martinez, R. (2024). Digital divides and indigenous language preservation: Technology access and language maintenance among Native American communities. International Journal of Indigenous Education.
27. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2022). Systematic review of research on artificial intelligence applications in higher education–where are the educators? International Journal of Educational Technology in Higher Education