

Assessing the Usability of Artificial Intelligence-Powered Mobile Learning Applications for Teaching the Hausa Language in Nigerian Colleges of Education

Tukur Muhammad¹, Umar Isah², Usman Ibrahim Kaoje³, Sani Muhammad Sani¹

¹Department of Arabic, Adamu Augie College of Education, Argungu, Kebbi State, Nigeria

²Department of Computer Science, Adamu Augie College of Education, Argungu, Kebbi State, Nigeria

³Department of Hausa, Adamu Augie College of Education, Argungu, Kebbi State, Nigeria

Published September 2026

Abstract

Despite the rapid global expansion of Mobile-Assisted Language Learning (MALL) and Artificial Intelligence in Education (AIED), empirical evaluations of AI-driven mobile applications designed for indigenous African languages remain limited. This study evaluated the usability and user experience of AI-powered mobile learning applications used in teaching and learning the Hausa language across selected Colleges of Education in Northern Nigeria. Grounded in the ISO 9241-11 Usability Model and the Technology Acceptance Model (TAM), a mixed-method empirical design combining a cross-sectional survey with task-based usability testing was executed with pre-service teachers ($N = 180$) and Hausa language lecturers ($N = 15$). Quantitative evaluation utilized the System Usability Scale (SUS), a structured Usability Dimensions Questionnaire (UDQ) assessing Learnability, Efficiency, Memorability, Error Tolerance, and Satisfaction, along with four controlled interaction tasks involving Automatic Speech Recognition (ASR) and Natural Language Processing (NLP) modules. Results revealed an overall mean SUS score of 64.20 ($SD = 12.80$), placing current user experience at a marginal level (Grade D). While Learnability scored high ($M = 4.12$, $SD = 0.65$), Error Tolerance ($M = 2.85$, $SD = 0.94$) and Efficiency ($M = 3.42$, $SD = 0.81$) rated lowest, driven by character rendering errors for hooked letters ($\hat{b}$, $\hat{d}$, $\hat{k}$) and miscalibrated ASR pitch detection for Hausa tonal accents. Multiple linear regression indicated that Error Tolerance ($\beta = .424$, $p < .001$) and Efficiency ($\beta = .380$, $p < .001$) significantly predicted overall user satisfaction ($R^2 = .551$). Furthermore, lecturers rated app usability significantly lower than students ($t = 3.482$, $p < .001$). The paper provides a evidence-based software framework for developer optimization and curriculum integration in teacher education.

Keywords: Mobile-Assisted Language Learning (MALL); Artificial Intelligence in Education (AIED); Hausa Language Pedagogy; System Usability Scale (SUS); Human-Computer Interaction (HCI); Colleges of Education; Nigeria

1 Introduction

Integrating digital technology into tertiary language education across Sub-Saharan Africa is essential for educational modernization and cultural preservation (Bitrus, 2026; Oloyede et al., 2023). In Nigerian Colleges of Education (COEs), which operate under the regulatory mandate of the National Commission for Colleges of Education (NCCE) to train primary and junior secondary educators, native language pedagogy—particularly Hausa, Yoruba, and Igbo—occupies a central position (Federal Ministry of Education [FME], 2014; Usman & Abdulhamid, 2021). The Hausa language, spoken by over 70 million native and secondary speakers across West and Central Africa, serves as a primary

medium of socio-cultural interaction and instructional delivery in Northern Nigeria (Alhasan, 2025; Yakasai, 2020).

Despite its demographic prominence, Hausa instruction in COEs has historically relied on conventional, teacher-centric lecturing constrained by a scarcity of interactive instructional materials (Lawal et al., 2024). Mobile-Assisted Language Learning (MALL), combined with Artificial Intelligence (AI) components such as Natural Language Processing (NLP) for automated syntax parsing and Automatic Speech Recognition (ASR) for tone detection, offers a pathway to upgrade language acquisition (Adam & Umar, 2023; Kukulska-Hulme et al., 2023).

However, the educational value of AI-powered mobile apps depends heavily on their usability – the extent to which a system allows users to accomplish goals with effectiveness, efficiency, and satisfaction in a specified context of use (International Organization for Standardization [ISO], 2018). While global commercial language applications (e.g., Duolingo, Babbel) are optimized for major international languages, mobile software adapted for indigenous African languages frequently exhibits UI/UX deficiencies, non-standard soft keyboard layouts for special orthography, and misaligned phonetic algorithms (Awe, 2023; Bello & Muhammad, 2024). Poor usability elevates cognitive load, causes high task error rates, and leads to technology abandonment among educators and pre-service teachers (Garba et al., 2022).

1.1 Statement of the problem

Deploying AI-powered mobile applications for Hausa language teaching in Nigerian Colleges of Education faces several technical and interaction design bottlenecks. Current mobile soft keyboards and NLP input forms struggle to parse specialized Hausa orthographic characters (such as hooked letters *ɓ*, *ɗ*, *ƙ* and glottal stops *ʔ*), resulting in high character input failure rates during automated grammar exercises (Bello & Muhammad, 2024). Additionally, embedded ASR pronunciation evaluation modules—often trained on standard non-tonal speech corpora—fail to evaluate Hausa tonal variations accurately (*càakāa* vs. *cáakāa*), leading to false negative scores and user frustration (Abubakar et al., 2023). Coupled with intermittent network coverage and a lack of instructor management interfaces, these usability flaws hinder effective application adoption. Without empirical HCI evaluation data, developers and education stakeholders lack the insights needed to refine indigenous AI language software.

1.2 Research objectives

1. To measure the baseline System Usability Scale (SUS) score of AI-powered mobile Hausa learning applications among pre-service teachers and lecturers in selected Nigerian COEs.
2. To evaluate user ratings across five core ISO 9241-11 usability sub-constructs: Learnability, Efficiency, Memorability, Error Tolerance, and Satisfaction.
3. To identify the specific interface, orthographic, and ASR performance barriers encountered during controlled task execution.
4. To determine the empirical effect of core usability dimensions on users' overall satisfaction and continued adoption intention.

1.3 Research hypotheses

- H_{01} : Usability sub-constructs (Learnability, Efficiency, Memorability, Error Tolerance) do not significantly predict overall user satisfaction with AI mobile Hausa learning tools.

- H_{02} : There is no statistically significant difference in perceived System Usability Scale (SUS) scores between student pre-service teachers and Hausa language lecturers.

2 Theoretical framework

This study synthesizes the ISO 9241-11 Ergonomic Framework (ISO, 2018; Nielsen, 1994) with the Technology Acceptance Model (TAM) (Davis, 1989; Venkatesh et al., 2012), as shown in Figure 1.

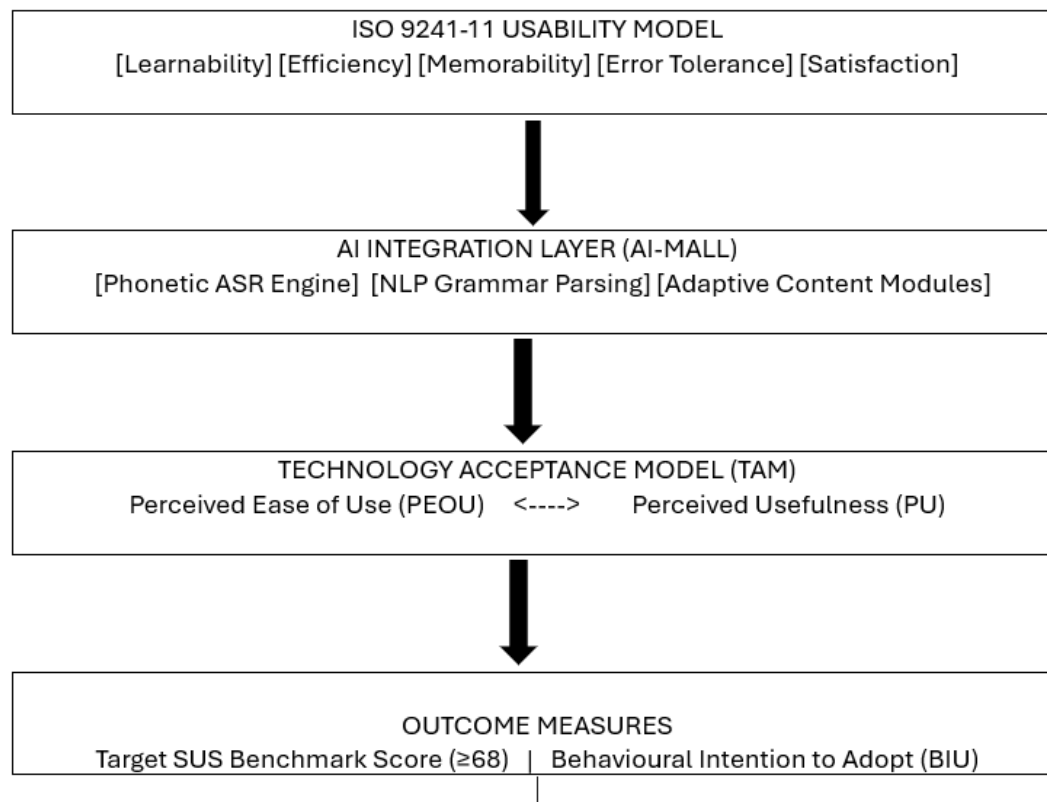


Figure 1. Theoretical framework of the study.

1. **ISO 9241-11 Framework:** Defines usability through three primary dimensions: *Effectiveness* (accuracy and task completion), *Efficiency* (resources expended relative to output), and *Satisfaction* (user attitudes). Nielsen (1994) operationalized these into five measurable components: Learnability, Efficiency, Memorability, Error Tolerance (low error rate), and Satisfaction.
2. **Technology Acceptance Model (TAM):** TAM posits that Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) determine an individual's Behavioral Intention to Use (BIU) a system. In this context, interface usability directly influences PEOU, which in turn influences lecturers' and pre-service teachers' willingness to integrate AI applications into classroom practice.

3 Methodology

3.1 Research design

The study adopted a convergent parallel mixed-methods empirical design combining a quantitative cross-sectional survey with task-based laboratory usability testing.

3.2 Population and sampling

The target population consisted of NCE II and III Hausa language pre-service teachers and academic staff across public Colleges of Education in North-Western Nigeria. A multi-stage sampling approach was used (Table 1):

- Stage 1: Stratified random selection of three public COEs across the region.
- Stage 2: Simple random sampling to select 180 pre-service teachers specializing in Hausa combinations.
- Stage 3: Purposive sampling to select 15 active Hausa language lecturers.

Table 1. Characteristics of the study sample.

Participant group	Sample size (n)	Male (n)	Female (n)	Mean age (years)
Pre-service Teachers (Students)	180	98	82	22.4 ± 2.1
Hausa Language Lecturers	15	11	4	41.6 ± 5.8
Total Study Sample	195	109	86	–

3.3 Evaluated AI applications

Participants interacted with two widely deployed Android-based Hausa applications incorporating AI features (*HausaTutor AI* and *Murna Interactive AI*). These apps contained:

- ASR Phonology Module: Evaluates student voice recordings for tonal correctness.
- NLP Grammar Checker: Analyzes written submissions and flags orthographic errors.
- Adaptive Quiz Engine: Dynamically alters question difficulty based on prior user error logs.

3.4 Data collection instruments

1. System Usability Scale (SUS): A 10-item standardized psychometric questionnaire using a 5-point Likert response scale (1 = Strongly Disagree to 5 = Strongly Agree). Scores were calculated using Brooke's (1996) standard formula to produce a composite usability score ranging from 0 to 100.
2. Usability Dimensions Questionnaire (UDQ): A 25-item structured instrument measuring the five sub-constructs (5 items per construct) on a 5-point Likert scale. Reliability testing yielded a Cronbach's alpha of $\alpha = .89$.
3. Structured Task Observation Protocol: Participants performed four controlled interaction tasks:
Task 1: Navigating system onboarding and opening the Hausa phonology module.
Task 2: Entering Hausa hooked-letter passages (*ḥ, ḍ, ḱ*) into the NLP grammar checker.
Task 3: Recording and submitting voice samples for ASR tonal pronunciation feedback.
Task 4: Executing an offline sync for an adaptive reading unit under low-bandwidth network emulation.

3.5 Data analysis procedures

Data were analyzed using IBM SPSS Statistics (Version 27.0). Descriptive statistics (Mean, Standard Deviation, Percentages) summarized demographic variables, SUS indices, and dimension scores. Inferential statistical procedures included multiple linear regression to model predictors of overall user satisfaction (H_{01}) and an independent-samples t -test to evaluate SUS variance between students and faculty (H_{02}).

4 Results

4.1 System Usability Scale (SUS) score

According to standard HCI evaluation benchmarks (Brooke, 1996; Sauro, 2011), a SUS score above 68.0 reflects acceptable usability, while scores below 68.0 signal marginal or unacceptable performance.

Table 2. System Usability Scale scores by cohort.

Evaluated cohort	Mean SUS score	Standard deviation (SD)	Benchmark / adjectival rating
Composite Sample (Total)	64.20	12.80	Grade D / Marginal (Below Average)
Student Sub-group	65.10	12.10	Marginal
Lecturer Sub-group	53.40	14.50	Unacceptable / Poor

The overall mean SUS score of 64.20 (Table 2) indicates significant interface and interaction barriers in the evaluated mobile applications.

4.2 Evaluation of usability sub-constructs

Usability sub-constructs were scored on a 1.00–5.00 scale (1.00–1.80 = Very Low, 1.81–2.60 = Low, 2.61–3.40 = Moderate, 3.41–4.20 = High, 4.21–5.00 = Very High).

Table 3. Ratings of the usability sub-constructs.

Usability sub-construct	Items	Mean (M)	Std. deviation (SD)	Adjectival rating
Learnability	5	4.12	0.65	High
Memorability	5	3.76	0.72	High
Efficiency	5	3.42	0.81	Moderate
Satisfaction	5	3.10	0.89	Moderate
Error Tolerance	5	2.85	0.94	Moderate / Low

Learnability rated highest ($M = 4.12$; Table 3), showing that users quickly master basic menu structures. In contrast, Error Tolerance scored lowest ($M = 2.85$), reflecting system friction during character entry and automated speech evaluation.

4.3 Task execution performance and error analysis

Table 4. Task execution performance and error analysis.

Task	Success rate (%)	Mean time (s)	Primary user failure / interaction error mode
T1: Module Onboarding	94.4%	24.5 ± 6.2	Minor visual menu confusion
T2: NLP Text Input	66.1%	78.2 ± 18.4	Lack of soft keyboard layout for hooked letters (<i>h, d, k</i>)
T3: ASR Speech Feedback	52.8%	112.0 ± 31.1	ASR rejection of correct tonal pitch due to accent miscalibration
T4: Offline Synchronization	38.3%	145.6 ± 42.0	Timeout failures; ambiguous progress bar state

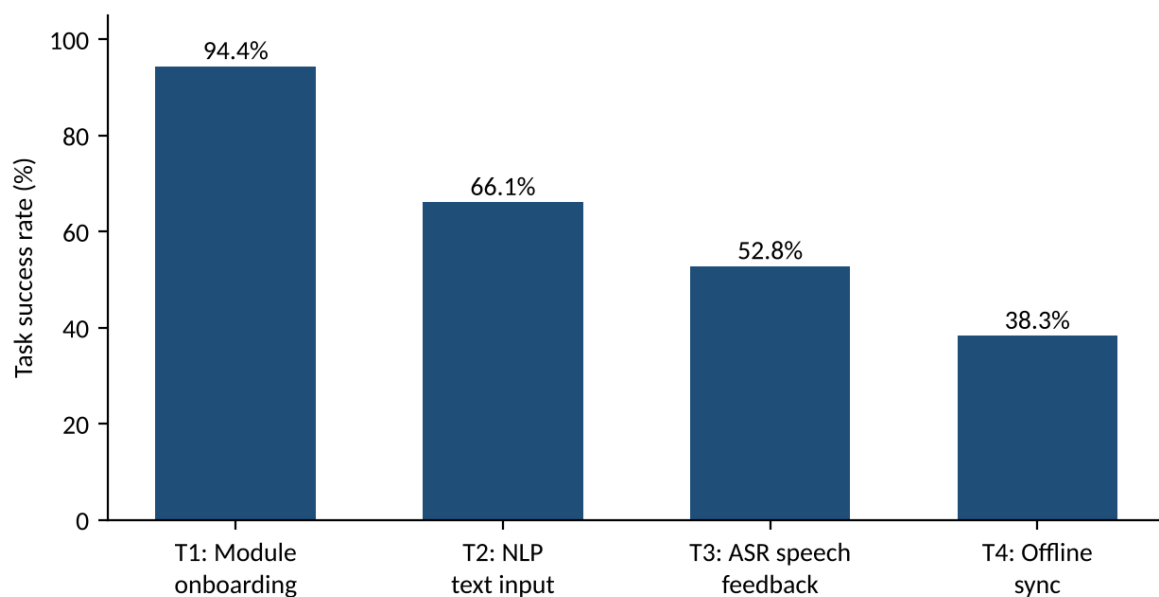


Figure 2. Task success rates across the four controlled tasks.

4.4 Hypotheses testing

4.4.1 Testing hypothesis 1 (H_{01})

Multiple linear regression assessed whether Learnability, Efficiency, Memorability, and Error Tolerance significantly predicted Overall Satisfaction (Y), as specified in Equation (1); the estimates are given in Table 5.

$$\text{Satisfaction} = \beta_0 + \beta_1(\text{Learnability}) + \beta_2(\text{Efficiency}) + \beta_3(\text{Memorability}) + \beta_4(\text{Error Tolerance}) + \varepsilon$$

Table 5. Multiple linear regression of overall satisfaction on the usability sub-constructs.

Predictor variable	Unstandardized B	Std. error	Standardized β	t-value	p-value
(Constant)	0.412	0.182	–	2.264	.025
Learnability	0.115	0.068	0.102	1.691	.092 (NS)
Memorability	0.142	0.071	0.128	2.000	.047

Predictor variable	Unstandardized B	Std. error	Standardized β	t-value	p-value
Efficiency	0.368	0.062	0.380	5.935	< .001
Error Tolerance	0.421	0.058	0.424	7.258	< .001

Model Summary: $R = .742$, $R^2 = .551$, Adjusted $R^2 = .539$, $F(4, 190) = 58.32$, $p < .001$.

The model accounted for 55.1% of the variance in overall user satisfaction. Error Tolerance ($\beta = .424$, $p < .001$) and Efficiency ($\beta = .380$, $p < .001$) emerged as the primary statistically significant predictors. Consequently, H_{01} was rejected.

4.4.2 Testing hypothesis 2 (H_{02})

An independent-samples t -test evaluated perceived usability differences between pre-service teachers and lecturers (Table 6).

Table 6. Independent-samples t -test of SUS scores by group.

Group	N	Mean SUS	SD	Mean diff.	t-value	df	p-value
Students	180	65.10	12.10	11.70	3.482	193	< .001
Lecturers	15	53.40	14.50				

A statistically significant difference emerged between students ($M = 65.10$, $SD = 12.10$) and lecturers ($M = 53.40$, $SD = 14.50$); $t(193) = 3.482$, $p < .001$. Lecturers rated application usability significantly lower, citing a lack of instructor authoring controls, missing diacritic options, and inadequate learning analytics. Consequently, H_{02} was rejected.

5 Discussion

The empirical findings demonstrate that current AI-powered Hausa mobile learning apps suffer from usability gaps, evidenced by the overall SUS score of 64.20 falling below international acceptability standards (68.0) (Brooke, 1996; Sauro, 2011).

While a high Learnability score ($M = 4.12$) indicates that digitally literate pre-service teachers easily grasp basic navigation, severe deficiencies in Error Tolerance ($M = 2.85$) and Efficiency ($M = 3.42$) impair learning tasks. The task analysis (Table 4 and Figure 2) clarifies these operational bottlenecks: 33.9% of participants failed Task 2 due to missing in-app virtual soft keyboards pre-configured with Hausa hooked letters (β , d , k), forcing reliance on awkward long-press actions or English substitutes (Bello & Muhammad, 2024).

Task 3 performance (52.8% completion rate) highlights algorithmic limitations in current AI language engines. Standard ASR algorithms misinterpret Hausa tonal contours, resulting in false negative speech evaluations (Abubakar et al., 2023). Furthermore, the significant difference in SUS scores between students ($M = 65.10$) and faculty ($M = 53.40$) reveals an architectural gap: current tools function as student revision apps rather than complete pedagogical systems equipped with instructor dashboards, aligned with earlier observations by Lawal et al. (2024) and Usman and Abdulhamid (2021).

6 Recommendations and conclusion

6.1 Recommendations

1. **Linguistic Soft Keyboard Integration:** Mobile app developers must embed custom soft keyboards featuring dedicated keys for special Hausa characters (*ḥ, ḍ, ḳ, Title/Hooked capitals*) to streamline text entry.
2. **Acoustic Model Recalibration:** AI development teams should fine-tune speech evaluation engines using localized Hausa speech datasets to ensure accurate tone parsing across regional accents.
3. **Instructor Dashboard Development:** Software designers should create instructor management web and mobile interfaces that allow lecturers to assign exercises, track progress, and adapt curriculum content.
4. **Offline Caching Optimization:** Mobile architectures should prioritize offline-first data synchronization to remain operational during network outages in rural tertiary campuses.

6.2 Conclusion

AI-powered Mobile-Assisted Language Learning holds potential for advancing indigenous language pedagogy in Nigerian teacher education. However, technical availability alone does not ensure classroom success. As demonstrated in this empirical evaluation, current mobile applications are limited by orthographic input barriers, miscalibrated speech evaluation, low error tolerance, and minimal instructor controls. Addressing these UI/UX and algorithmic challenges will enable software developers and educational leaders to deliver usable, effective digital platforms that support indigenous language preservation and pre-service teacher training in Nigeria.

References

- [1] Abubakar MS, Haruna U, Salisu A (2023). Acoustic modeling of tone recognition for Hausa speech synthesis and recognition engines. *Journal of African Language Technology*, 4(1), 45–58.
- [2] Adam I, Umar AM (2023). Artificial intelligence in Nigerian higher education: Opportunities, infrastructure challenges, and policy imperatives. *Nigerian Journal of Computer Science and Education*, 9(2), 112–126.
- [3] Alhasan K (2025). Using Hausa language in the assessment of students: A case study of higher institution in Bauchi, Nigeria. *Kashf Journal of Multidisciplinary Research*, 2(6), 7–11. <https://doi.org/10.71146/kjmr486.0>
- [4] Awe KI (2023). Assessing the usability factors influencing students' utilization of mobile devices for learning in FCT-Abuja (Unpublished master's thesis). *Federal University of Technology*, Minna, Nigeria.
- [5] Bello AY, Muhammad B (2024). Character encoding and key mapping challenges in indigenous African scripts on mobile virtual soft keyboards. *African Journal of Human-Computer Interaction*, 8(1), 19–34.
- [6] Bitrus HMIDB (2026). Effectiveness of mobile-assisted applications in teaching Hausa reading and writing among junior secondary school students in Plateau Central Senatorial Zone, Nigeria. *International Journal of Multidisciplinary Research, Science, Technology and Innovation*, 11(2), 560–571.
- [7] Brooke J (1996). SUS: A 'quick and dirty' usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.
- [8] Davis FD (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- [9] Federal Ministry of Education (2014). National policy on education (6th ed.). *NERDC Press*.
- [10] Garba S, Yakubu M, Danjuma A (2022). Usability of digital learning platforms in Northern Nigerian tertiary institutions during network disruptions. *West African Journal of Educational Technology*, 14(3), 88–101.

-
- [11] International Organization for Standardization (2018). Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts (ISO Standard No. 9241-11:2018). <https://www.iso.org/standard/63500.html>
 - [12] Kukulska-Hulme A, Lee H, Norris L (2023). Mobile-Assisted Language Learning in global contexts: Emerging technologies and pedagogical ergonomics. *ReCALL*, 35(2), 140–156. <https://doi.org/10.1017/S095834402200018X>
 - [13] Lawal A, Cavus N, Lawan AA, Sani I (2024). Hausar Kurma: Development and evaluation of interactive mobile app for the English-Hausa sign language alphabet. *IEEE Access*, 12, 46012–46023. <https://doi.org/10.1109/ACCESS.2024.3381538>
 - [14] Nielsen J (1994). Usability engineering. *Morgan Kaufmann Publishers*.
 - [15] Oloyede EO, Adeleke AO, Bello TO (2023). Integrating artificial intelligence into language education in Nigerian tertiary institutions: Readiness and perceptions. *Journal of Educational Technology Systems*, 52(1), 65–84. <https://doi.org/10.1177/00472395231189012>
 - [16] Sauro J (2011). A practical guide to the System Usability Scale: Measuring usability handy guide. *Measuring Usability LLC*.
 - [17] Usman T, Abdulhamid M (2021). Pedagogy and technology in Hausa language education: Challenges and prospects in Colleges of Education in North-West Nigeria. *Dutse Journal of Humanities and Social Sciences*, 6(2), 102–115.
 - [18] Venkatesh V, Thong JY, Xu X (2012). Consumer acceptance and use of information technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, 36(1), 157–178. <https://doi.org/10.2307/41410412>
 - [19] Yakasai HM (2020). Hausa orthography and digital communication: Adapting traditional scripts to modern computing platforms. *Bayero Journal of Linguistics and Communication*, 11(1), 15–29.