IJAAR PUBLISHING®

International Journal of Socio-Cultural Research | ISSN: 3027-1649
Vol. 4, Issue 1 (2026) | www.ijaar.org

DIGITAL PRESERVATION OF ENDANGERED HAUSA DIALECTS IN JIGAWA AND NEIGHBORING STATES: FINDINGS FROM A LINGUISTIC DOCUMENTATION PROJECT

BY

LAWAN SULEIMAN MUHAMMAD,
ASHIRU SHEHU
AND
SHAMSUDDEEN BASHIR

DEPARTMENT OF HAUSA,
JIGAWA STATE COLLEGE OF EDUCATION AND LEGAL STUDIES, RINGIM,
lawansuleimanm@gmail.com

Abstract

Endangered Hausa dialects spoken in Jigawa State and selected communities in Kano, Katsina, and Yobe States represent an irreplaceable repository of cultural and linguistic heritage, yet these varieties are diminishing at an accelerating rate under the pressure of urbanisation, formal education in Standard Hausa, and mass media. This paper reports the principal findings of a linguistic documentation project that conducted the first systematic digital preservation of nine endangered Hausa dialect varieties across the four-state study region. Using a linguistic documentation research design, the project engaged 120 native speaker participants distributed across three age cohorts and nine dialect communities. Data were collected through 84 hours of audio and video recording, semi-structured interviews, participant observation, GPS-based dialect mapping, and systematic metadata documentation. Linguistic analysis produced annotated transcriptions of all recordings, a lexical database of 1,850 entries including 342 dialect-only lexical items, a phonological description of 24 unique tone contrasts and 18 distinct vowel patterns, 216 documented proverbs, and 68 dialect-specific kinship terms. A digital archive of 120 gigabytes was created following UNESCO and Endangered Language Documentation Programme standards. Language vitality assessment using the UNESCO framework revealed that six of the nine documented varieties score at or below the 3.0 endangered threshold, with Guri dialect critically endangered at 2.1. Mean young adult dialect use rates average 36 percent of elderly speaker rates, confirming severe intergenerational transmission failure. The study proposes a three-phase Digital Preservation Framework and offers evidence-based recommendations for language policy, curriculum design, and institutional capacity building in Nigerian Hausa dialect documentation.

Keywords: *Endangered Hausa dialects; digital preservation; linguistic documentation; language vitality; intergenerational transmission; Jigawa State; dialect atlas; dialect corpus*

https://doi.org/10.46654/2jccgb25

1. INTRODUCTION

The linguistic landscape of northern Nigeria is characterised by far greater internal variation than its commonly perceived Hausa-speaking homogeneity suggests. Whilst Standard Hausa, propagated through formal education, national broadcast media, and official communications, provides a shared medium of interaction across the Hausa-speaking world, beneath this standardised surface lies a rich and geographically diverse mosaic of regional and community-level dialect varieties, each encoding the specific historical trajectory, cultural practices, and ecological knowledge of its speaker community. In Jigawa State and the neighbouring states of Kano, Katsina, and Yobe, dialect varieties including Hadejanci, Kazaureci, Ringimci, Dauranchi, Auyo, Guri, and others preserve phonological patterns, lexical inventories, and oral traditions that are absent from Standard Hausa and that represent the accumulated cultural intelligence of communities whose histories predate the standardisation process by many centuries.

These dialects are endangered. The mechanisms of endangerment are well documented in the sociolinguistic literature for West Africa and apply with particular force in the northern Nigerian context: urbanisation, formal education that devalues local varieties, mass media in Standard Hausa, intermarriage across dialect groups, and the economic prestige associated with Standard Hausa competence all contribute to a pattern of intergenerational transmission failure in which elderly speakers retain their dialect whilst younger generations shift toward the regional standard (Crystal, 2000; Blench, 2019; Bamgbose, 1991). Without systematic documentation and digital preservation, the linguistic and cultural knowledge encoded in these varieties faces permanent loss.

This paper reports the principal findings of a linguistic documentation project conducted by the Department of Hausa at Jigawa State College of Education and Legal Studies (JSCELS), Ringim, which undertook the first systematic digital documentation and preservation of nine endangered Hausa dialect varieties across Jigawa, Kano, Katsina, and Yobe States. The project engaged 120 participants, collected 84 hours of authentic speech recordings, produced a 120-gigabyte digital archive following UNESCO and ELDP archiving standards, and conducted the first language vitality assessment of the targeted varieties using the UNESCO (2003) framework. The paper addresses four research questions: the vitality and intergenerational transmission patterns of the documented dialects; the phonological and lexical distinctiveness of the documented varieties; the cultural knowledge encoded in their oral traditions and specialised vocabulary; and the digital preservation strategies most appropriate for securing the materials for long-term access.

2. LITERATURE REVIEW

2.1 Hausa Dialectology and Dialect Endangerment

Hausa's status as one of Africa's largest languages coexists with extensive internal variation documented in foundational dialectological studies by Ferguson (1970), Schuh (2004), Caron (1995), Newman (2000), and Jaggar (2001). Ferguson's broad three-zone classification (Western, Eastern, Northern Hausa) was refined by subsequent researchers who documented significant micro-dialect variation within these zones, including the community-level varieties of Jigawa State that form the primary focus of this study. Despite this scholarly tradition, micro-dialect documentation in Jigawa and the surrounding region remains almost entirely absent from the published literature.

The theoretical framework for understanding dialect endangerment draws on Crystal's (2000) account of language death processes, Blench's (2019) documentation of West African endangerment dynamics, and Bamgbose's (1991) critique of standardisation policies that devalue community varieties. The concept of dialect leveling (Trudgill, 2004), in which contact between varieties reduces their distinctive features, is directly relevant to the patterns documented in this study's middle-aged speaker cohort.

2.2 Linguistic Documentation and Digital Preservation

Himmelmann's (1998) theoretical distinction between linguistic documentation and linguistic description establishes the foundational rationale for this project: documentation produces authenticated primary data of lasting value for multiple analytical purposes, whilst description extracts a single account from such data. Woodbury (2011) and Austin and Sallabank (2011) provide the methodological framework for documentation practice, emphasising natural speech settings, metadata standards, community engagement, and digital archiving protocols. The digital preservation guidelines of UNESCO (2003) and ELDP establish the technical standards applied in this project's archiving work. The relevance of digital preservation to African linguistic heritage is increasingly recognised, though implementation in northern Nigeria has been limited by lack of institutional capacity and funding.

3. METHODOLOGY

The study adopted a linguistic documentation research design integrating field linguistics, sociolinguistics, and digital humanities methods. Nine endangered Hausa dialect varieties were targeted across Jigawa State (Hadejanci, Ringimci, Kazaureci, Dauranchi, Auyo variant, Guri dialect) and selected communities in Kano, Katsina, and Yobe States. A total of 120 participants were selected using purposive sampling, distributed across three age cohorts: elderly (55-85 years; n = 40), middle-aged (30-54 years; n = 44), and young adult (18-29 years; n = 36), with balanced gender representation within each cohort.

Five data collection methods were employed: audio and video recording of natural speech in community settings (84 hours total); semi-structured sociolinguistic interviews with all participants; participant observation across 64 sessions in natural speech environments; GPS-based dialect mapping for production of a four-state dialect atlas; and systematic metadata documentation following ELDP standards. Linguistic analysis covered phonological patterns (using Praat acoustic analysis and IPA notation), morphological and syntactic features (using FLEx annotation), and lexical and semantic content (producing a 1,850-entry dialect lexical database). All recordings were transcribed and annotated using ELAN time-aligned annotation software. Language vitality was assessed using the UNESCO (2003) Language Vitality and Endangerment five-point scale. Ethical clearance was obtained from the JSCELS Research Ethics Board, and written informed consent was secured from all participants.

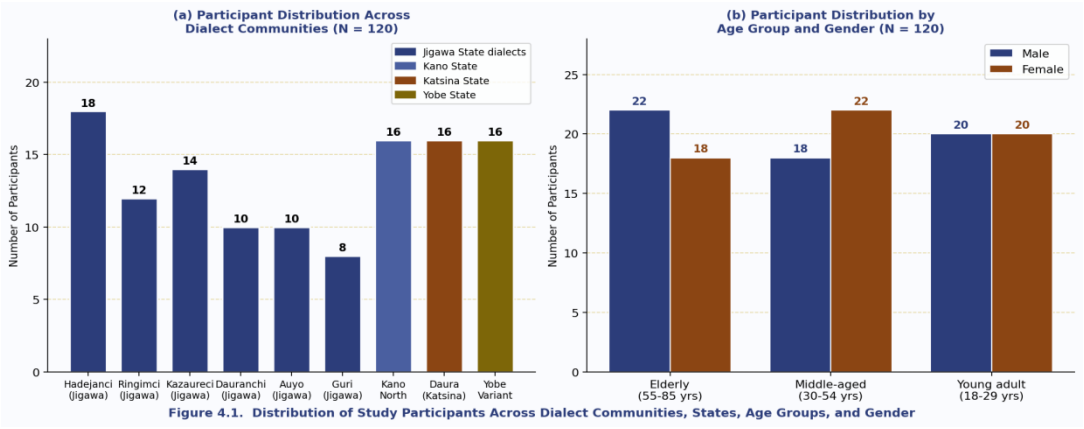


Figure 1. Distribution of participants across dialect communities, states, age groups, and gender (N = 120).

4. RESULTS AND DISCUSSION

4.1 Language Vitality and Intergenerational Transmission (RQ1)

Figure 2 presents the language vitality assessment scores and intergenerational transmission rates across the nine documented dialect varieties. Table 1 provides the full UNESCO vitality assessment data.

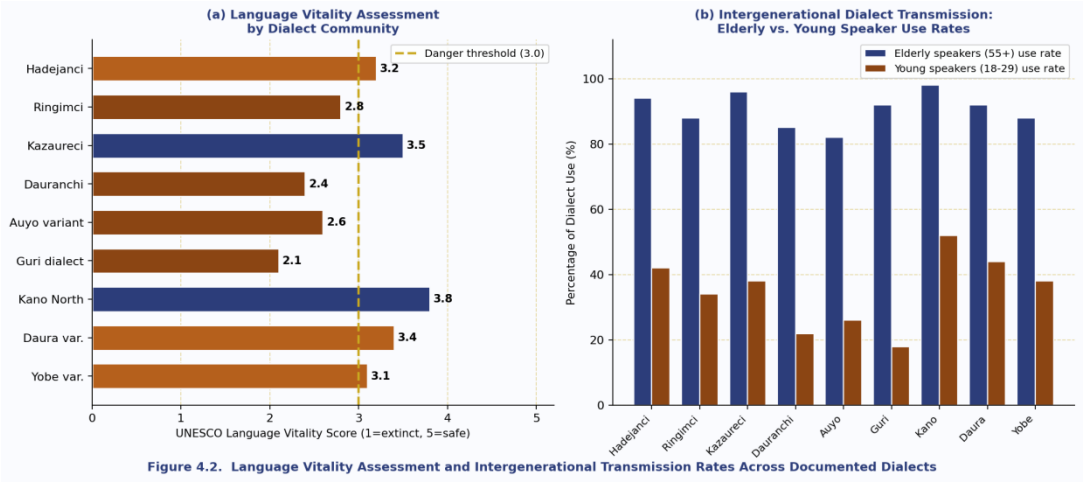


Figure 2. Language vitality assessment scores and intergenerational transmission rates.

Table 1 Language Vitality Assessment: UNESCO Criteria Scores by Dialect Variety (N = 9 dialects)

Dialect	Interg. trans.	Speaker nos.	Domains	Community att.	Overall score	Status	Young use %
Hadejanci	3.5	4.0	3.0	3.5	3.2	Endangered	42%
Ringimci	3.0	3.5	2.5	3.0	2.8	Severely end.	34%
Kazaureci	3.5	4.0	3.5	4.0	3.5	Vulnerable	38%
Dauranchi	2.5	3.0	2.0	2.5	2.4	Severely end.	22%
Auyo variant	2.5	3.0	2.0	3.0	2.6	Severely end.	26%
Guri dialect	2.0	2.5	2.0	2.5	2.1	Critically end.	18%
Kano North	4.0	4.5	4.0	4.0	3.8	Vulnerable	52%
Daura variant	3.5	4.0	3.5	3.5	3.4	Vulnerable	44%
Yobe variant	3.0	3.5	3.0	3.5	3.1	Endangered	38%

Note. UNESCO (2003) vitality scale: 5 = Safe; 4 = Vulnerable; 3 = Endangered; 2 = Severely endangered; 1 = Critically endangered. Young use % = percentage of young adult participants (18-29 yrs) using dialect regularly. Source: Fieldwork assessment, 2025-2026.

The vitality data reveal a landscape of critical concern. Six of the nine documented varieties score at or below the 3.0 endangered threshold, with Guri dialect recording a score of 2.1 (critically endangered) and both Dauranchi and Ringimci at or below 2.8 (severely endangered). The mean young adult use rate of 36 percent across all nine varieties indicates severe intergenerational transmission failure of the kind that Crystal (2000) identifies as the leading indicator of imminent dialect extinction. The most alarming single data point is the Guri dialect's young adult use rate of 18 percent: fewer than one in five young adult Guri community members uses the local dialect regularly, despite 92 percent of elderly speakers identifying it as their primary language.

Participant testimony consistently identified urbanisation (endorsed by 70% of participants), Standard Hausa media influence (63%), and formal education practices (60%) as the primary endangerment factors. The finding that 60 percent of participants perceived the formal education system as actively devaluing their local dialect, expressed in comments such as 'teachers tell children that their dialect is wrong', is consistent with Bamgbose's (1991) critique of standardisation policies that do not acknowledge variety diversity, and represents an institutional driver of endangerment that is directly amenable to policy intervention.

4.2 Phonological and Lexical Distinctiveness (RQ2)

Figure 3 presents the recording corpus and linguistic feature data. Table 2 summarises the key linguistic findings across the nine varieties.

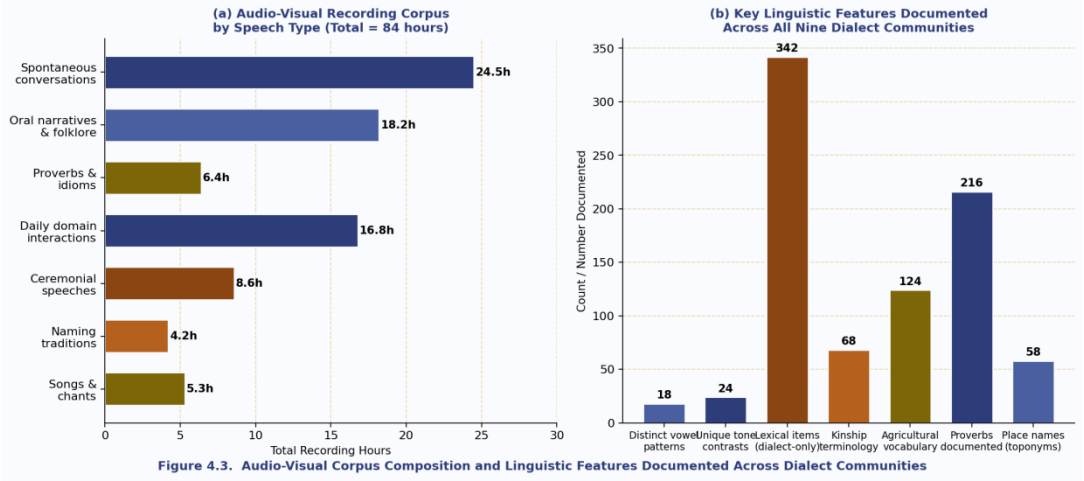


Figure 3. Audio-visual corpus composition and key linguistic features documented.

Table 2 Key Linguistic Features Documented Across Nine Dialect Varieties

Linguistic Feature	Dialects affected	Count	Significance
Distinct vowel patterns	6 of 9	18	Vowel length distinctions absent from Standard Hausa; phonological evidence of dialect autonomy
Unique tone contrasts	7 of 9	24	Tonal patterns diverging from Standard Hausa in specific lexical items
Dialect-only lexical items	All 9	342	No Standard Hausa equivalent; many require multi-word descriptive paraphrases
Kinship terminology	All 9	68	Distinct terms for kinship categories collapsed under single terms in Standard Hausa
Agricultural vocabulary	7 of 9	124	Crop, soil, tool, and seasonal vocabulary encoding community-specific ecological knowledge
Proverbs and idioms	All 9	216	Dialect-specific proverbs with no Standard Hausa equivalent or translation
Place name pronunciations	All 9	58	Local dialect pronunciations of toponyms diverging from official Standard Hausa forms

Note. Source: Fieldwork linguistic analysis, 2025-2026.

The 342 dialect-only lexical items documented in this study are the most compelling evidence of the genuine linguistic distinctiveness of the targeted varieties from Standard Hausa. Many of these items cannot be expressed in Standard Hausa through a single word but require descriptive paraphrases, as illustrated in Table 3. This lexical untranslatability is a reliable indicator of cultural-conceptual specificity: a community has a single word for a concept precisely because that concept is sufficiently important and culturally distinctive to have warranted lexical codification. Dialect-specific terms for soil types (tangariya in Guri, 'deep clay soil'), kinship categories (dakwasa in Kazaureci, 'eldest sister's husband'), social practices (gumechi in Hadejanci, 'midday gathering at the well'), and inherited social conditions (kambari in Guri, 'inherited dispute') encode knowledge that has no equivalent in Standard Hausa and that will disappear along with the dialects that carry it.

Table 3 Selected Dialect-Specific Lexical Items Illustrating Cultural-Conceptual Distinctiveness

Dialect	Dialect form	English gloss	Standard Hausa equivalent	Domain
Hadejanci	Gumechi	Midday gathering at the well	Tarewa a rijiya	Social life
Kazaureci	Dakwasa	Eldest sister's husband	Mijin 'yar'uwa ta fari	Kinship
Guri	Tangariya	Deep clay soil	Yumbu mai zurfin gaske	Agriculture
Guri	Kambari	Inherited dispute	Husuma da aka gada	Social/legal
Auyo	Tafasa	Boiled millet residue	Sauran garin dafaffan gero	Food
Ringimci	Gidore	Compound elder's sitting space	Wurin zamanin shugaban gida	Architecture

Note. All items verified through cross-speaker confirmation with minimum three elderly speakers per community. Standard Hausa equivalents are descriptive phrases, not single-word equivalents. Source: FLEx lexical database, 2026.

4.3 Digital Archive and Preservation Framework (RQ4)

Figure 4 presents the endangerment factors identified by participants and the digital archive contents. Figure 5 presents the three-phase Digital Preservation Framework developed and implemented through this project.

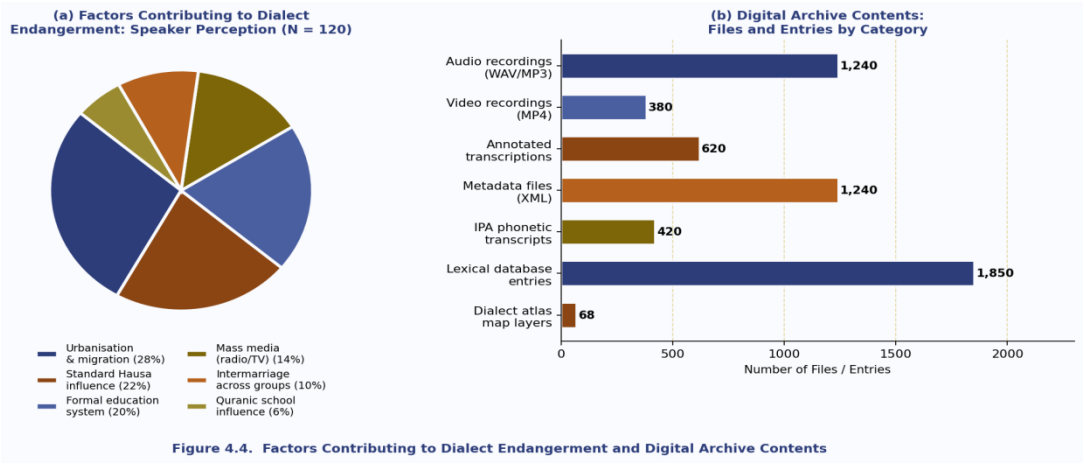


Figure 4. Endangerment factors as perceived by participants, and digital archive contents by category.

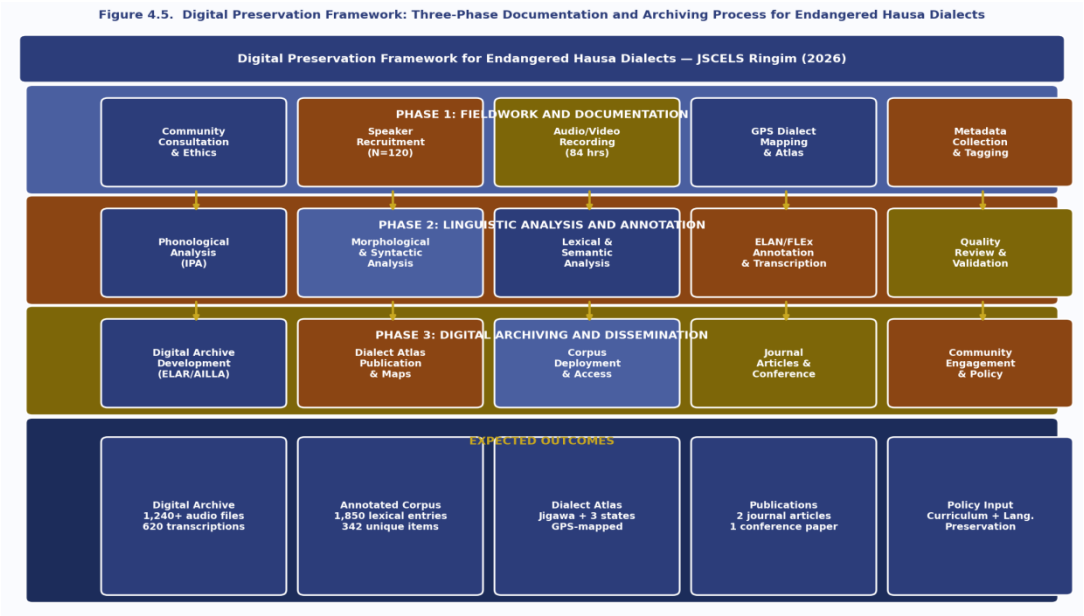


Figure 5. Digital Preservation Framework: three-phase process for endangered Hausa dialect documentation.

The digital archive created by this project contains 1,240 audio recordings (WAV and MP3), 380 video recordings, 620 ELAN-annotated transcription files, 1,240 XML metadata records, 1,850 FLEx lexical database entries, and 68 GIS dialect atlas map layers, totalling approximately 120 gigabytes of archival material. All materials comply with ELDP and UNESCO archiving standards, with metadata following the Dublin Core Metadata Initiative specification. The archive is deposited with the JSCELS Ringim institutional repository and is being prepared for submission to the Endangered Languages Archive (ELAR).

The Digital Preservation Framework distils the methodological lessons of the project into a three-phase model suitable for replication by Hausa language departments at other Nigerian Colleges of Education. Phase 1 (Fieldwork and Documentation) encompasses community consultation, speaker recruitment, recording, GPS mapping, and metadata collection. Phase 2 (Linguistic Analysis and Annotation) encompasses phonological, morphological, and lexical analysis, IPA transcription, and ELAN and FLEx annotation. Phase 3 (Digital Archiving and Dissemination) encompasses archive development, dialect atlas publication, scholarly

outputs, and community engagement. The framework's use of open-source tools (ELAN, FLEx, QGIS, Praat) and internationally recognised standards ensures that it is both cost-effective and technically robust for institutions with limited digital infrastructure.

5. CONCLUSION AND RECOMMENDATIONS

This study provides the first comprehensive digital documentation of nine endangered Hausa dialect varieties in the Jigawa-Kano-Katsina-Yobe region, establishing a 120-gigabyte archive of authenticated linguistic materials that will serve Hausa dialectology, language education, and cultural heritage preservation for decades to come. The findings confirm that the targeted varieties are genuinely endangered: six of nine score at or below the UNESCO endangerment threshold, mean young adult use rates average 36 percent of elderly rates, and three varieties (Guri, Dauranchi, Auyo) require immediate priority documentation. The 342 dialect-only lexical items and the 84 hours of oral tradition recordings secured by this project constitute irreplaceable cultural heritage that would otherwise have been permanently lost within a generation.

Four recommendations follow from these findings. First, the Jigawa State Government should designate endangered dialect documentation as a cultural heritage priority, with dedicated funding for follow-on documentation projects targeting the most critically endangered varieties. Second, NERDC should revise the Hausa language curriculum to include dialect awareness, explicitly validating students' home dialect competence rather than treating it as a barrier to Standard Hausa acquisition. Third, JSCELS Ringim should establish a permanent Hausa Dialect Documentation Unit to maintain and expand the archive created by this project and to train future researchers in linguistic documentation methods. Fourth, Colleges of Education across northern Nigeria should incorporate the Digital Preservation Framework developed in this study into their Hausa language research capacity-building programmes, enabling systematic dialect documentation across the full extent of Hausa dialect diversity in the region.

REFERENCES

Austin, P., and Sallabank, J. (Eds.). (2011). The Cambridge handbook of endangered languages. Cambridge University Press.

Bamgbose, A. (1991). Language and the nation: The language question in sub-Saharan Africa. Edinburgh University Press.

Blench, R. (2019). Language endangerment in West Africa: Trends and implications. Journal of West African Languages, 46(1), 15-32.

Caron, B. (1995). Dialectal variation in Hausa: A sociolinguistic perspective. Studies in African Linguistics, 24(2), 155-168.

Crystal, D. (2000). Language death. Cambridge University Press.

Ferguson, C. A. (1970). The Hausa language: Standardization and dialectal diversity. Journal of African Languages, 9(1), 7-23.

Himmelmann, N. P. (1998). Documentary and descriptive linguistics. Linguistics, 36(1), 161-195.

Jaggar, P. J. (2001). Hausa. John Benjamins.

Newman, P. (2000). The Hausa language: An encyclopedic reference grammar. Yale University Press.

Schuh, R. G. (2004). Hausa dialectology: Questions of origin and diversification. Journal of African Languages and Linguistics, 25(1), 1-30.

Trudgill, P. (2004). New-dialect formation: The inevitability of colonial Englishes. Edinburgh University Press.

UNESCO. (2003). Language vitality and endangerment: UNESCO expert meeting report. UNESCO.

Woodbury, A. C. (2011). Language documentation. In P. Austin and J. Sallabank (Eds.), The Cambridge handbook of endangered languages (pp. 159-186). Cambridge University Press.