

ENDANGERED LANGUAGES IN BORNEO: A SYSTEMATIC REVIEW OF CLASSIFICATIONS AND PRESERVATION STRATEGIES

Zulfadzlee Zulkiflee

Universiti Brunei Darussalam

24H8045@ubd.edu.bn

ABSTRACT

Languages across Borneo, including Belait, Penan, Bidayuh, and minority Narum variants in the Baram-Tinjar region, are rapidly endangered due to language shift, limited intergenerational transmission, and the dominance of regional languages. Despite Borneo's linguistic diversity and significance to Austronesian classification, a comprehensive systematic synthesis of language classifications and preservation efforts remains limited. This study addresses the research gap by systematically reviewing Scopus-indexed literature on language classification and preservation strategies for endangered languages in Borneo. A qualitative systematic review was conducted following PRISMA guidelines, with a Boolean search applied across the Scopus database. Inclusion criteria encompassed studies addressing language classification (phonological features, sociolinguistic vitality frameworks, linguistic relationships) and community-based preservation initiatives published between 2000 and 2026. From 509 initial records, 11 studies met all inclusion criteria. They were subjected to a thematic synthesis organised around two research questions: (RQ1) how endangered languages in Borneo have been classified using comparative-historical, lexicostatistical, or typological methodologies; and (RQ2) what preservation strategies have been employed and what factors determine their effectiveness. Findings reveal that classification in the reviewed corpus draws on descriptive linguistics, orthography development, UNESCO vitality frameworks, and social network analysis, with the outcomes emerging from studies triangulating structural, functional, and social evidence. Documentation, community-based revitalisation, and technological interventions constitute the preservation practice, with community ownership consistently distinguishing sustainable from short-lived initiatives. Regional disparities are evident, with Sarawak exhibiting stronger documentation infrastructure than Brunei, Sabah, and Kalimantan. This review underscores the critical need for integrated, coordinated documentation efforts and evidence-based language policies across Borneo's diverse political contexts.

Keywords: endangered languages, Borneo, language classification, preservation strategies, systematic review

ABSTRAK

Bahasa-bahasa di seluruh Borneo, termasuk Belait, Penan, Bidayuh, dan varian minoriti Narum di kawasan Baram-Tinjar, semakin terancam akibat peralihan bahasa, penghantaran antara generasi yang terhad, serta dominasi bahasa-bahasa serantau. Walaupun kepelbagaian linguistik Borneo dan kepentingannya dalam klasifikasi bahasa-bahasa Austronesia diakui secara meluas, kajian literatur bersistematis yang komprehensif mengenai klasifikasi dan pemeliharaan bahasa masih terbatas. Kajian ini menangani jurang penyelidikan tersebut dengan meneliti secara sistematis kajian berindeks Scopus tentang klasifikasi bahasa dan strategi pemeliharaan bahasa-bahasa terancam di Borneo. Kajian sorotan literatur bersistematis kualitatif telah dijalankan mengikut garis panduan PRISMA, dengan menggunakan teknik carian Boolean yang diaplikasikan ke dalam pangkalan data Scopus. Kriteria inklusi meliputi kajian yang membahaskan klasifikasi bahasa (kerangka vitaliti sosiolinguistik, hubungan linguistik) serta inisiatif pemeliharaan berasaskan komuniti yang diterbitkan dari tahun 2000 hingga 2026. Daripada 509 rekod awal, sebanyak 11 kajian memenuhi semua kriteria inklusi dan dianalisis secara tematik berdasarkan dua soalan kajian: (RQ1) bagaimana bahasa-bahasa terancam di Borneo diklasifikasikan menggunakan metodologi linguistik komparatif, leksikostatistik, atau tipologi; dan (RQ2) apakah strategi pemeliharaan yang digunakan serta faktor-faktor yang menentukan keberkesanannya. Dapatan kajian memperlihatkan bahawa klasifikasi dalam korpus yang dikaji menggunakan analisis linguistik deskriptif, pembangunan ortografi, kerangka vitaliti UNESCO dan analisis hubungan sosial. Hasil kajian yang signifikan dapat dilihat daripada kajian yang menggabungkan bukti struktural, fungsional dan sosial. Pendokumentasian, revitalisasi berasaskan komuniti, dan intervensi teknologi membentuk tiga tonggak pemeliharaan bahasa, dengan hubungan komuniti secara konsisten membezakan inisiatif yang mampan daripada inisiatif jangka pendek. Variasi serantau turut dikenal pasti, iaitu Sarawak menunjukkan inisiatif pendokumentasian yang lebih kukuh berbanding Brunei, Sabah, dan Kalimantan. Kajian ini menegaskan keperluan mendesak terhadap usaha pendokumentasian dan dasar bahasa berasaskan bukti merentasi konteks politik di Borneo.

Kata kunci: bahasa terancam, Borneo, klasifikasi bahasa, strategi pemeliharaan, literatur bersistematis

INTRODUCTION

Borneo is home to one of the most linguistically diverse regions in Southeast Asia, encompassing hundreds of distinct languages distributed across Brunei Darussalam, the Malaysian states of Sarawak and Sabah, and the Indonesian territory of Kalimantan (Adelaar & Himmelmann, 2004). These languages belong predominantly to the Austronesian family, with concentration within the Malayo-Polynesian and Land

Dayak subgroups, as well as several non-Austronesian Papuan-contact languages in eastern Kalimantan (Blust, 2013). The linguistic heritage of Borneo represents an irreplaceable archive of human cognitive and cultural history; yet, increasing numbers of these languages face severe endangerment (Moseley, 2010). Language endangerment in Borneo is driven by complex interactions among sociolinguistic forces, including the dominance of national languages such as Malay and Indonesian, rural-to-urban migration, and limited intergenerational transmission. Languages such as Belait and Tutong in Brunei, Narum and Bakong in Sarawak, and numerous Dayak varieties across Kalimantan are estimated to have fewer than 100 active speakers, with younger generations increasingly shifting to regionally dominant languages (Smith, 2017; Omar & Norahim, 2020).

Despite the scale and urgency of language loss in Borneo, the scholarly literature on endangered language classification and preservation remains limited. Studies have tended to focus narrowly on either linguistic classification, phonological reconstruction, proto-forms, and subgrouping or on sociolinguistic and pedagogical aspects of preservation, without systematically integrating both dimensions. This field limits the creation of evidence-based preservation models that are both linguistically rigorous and practically effective. Moreover, while individual documentation projects and community-based revitalisation initiatives have been reported across the region, a comprehensive systematic review that maps and evaluates these efforts across Borneo's diverse political and cultural contexts has not yet been conducted. The present study addresses this gap through a qualitative systematic review of peer-reviewed literature, articles, journals, and documentation projects published between 2000 and 2026. Two principal research questions guide the review:

RQ1: How have endangered languages in Borneo been classified using comparative-historical, lexicostatistical, and typological methodologies in existing Scopus-indexed literature?

RQ2: What preservation and documentation strategies have been employed for endangered languages across Borneo's diverse political and cultural contexts, and what factors determine their effectiveness and sustainability?

To address these questions, the systematic review synthesises existing approaches to language classification across Borneo's endangered languages and evaluates the range of preservation strategies documented in the literature. By integrating these two dimensions, the review offers a more holistic understanding of what is known, what works, and what critical gaps remain in efforts to sustain Borneo's endangered linguistic heritage. The findings are particularly relevant to linguists, language planners, community organisations, and government bodies engaged in language maintenance across the Brunei-Sarawak-Sabah-Kalimantan region.

LITERATURE REVIEW

The issue of endangered languages in Borneo has increasingly attracted scholarly attention due to the region's exceptional linguistic diversity and the accelerating processes of language shift, attrition, and loss. Situated within the broader Austronesian linguistic landscape, Borneo is home to numerous minority and indigenous languages, many of which are under-documented and critically endangered. Recent literature, particularly that focusing on language documentation, digital innovation, and minority language sustainability, provides important insights into both the classification challenges and preservation strategies relevant to this study. Studies such as Abingosa et al. (2025) emphasise the urgent need for systematic language documentation as a foundational step toward preservation. Documentation efforts typically involve the collection of lexical, grammatical, and oral narrative data, which serve both as academic resources and as community heritage archives.

However, previous studies also reveal a critical gap in documentation practices: many projects remain externally driven and lack long-term sustainability. Umar et al. (2025) argue that the visibility of indigenous languages in formal education systems remains limited, contributing to weak intergenerational transmission. This issue is particularly relevant in Borneo, where languages such as Belait, Penan, and Narum are experiencing declining numbers of speakers. Furthermore, the inclusion of marginalised language communities, such as sign language users, underscores the broader conceptualisation of linguistic endangerment. Alisoy (2025) emphasise that language documentation must extend beyond spoken languages to encompass diverse communicative systems, underscoring the need for inclusive, multidisciplinary approaches.

The classification of Bornean languages remains a contested domain due to complex historical, geographical, and contact-induced factors. Traditional comparative methods have been instrumental in establishing genealogical relationships among Austronesian languages; however, recent scholars call for more integrative approaches. Umar et al. (2025) highlight that language classification must consider not only structural features but also sociolinguistic dynamics, including language contact, borrowing, and convergence. In Borneo, prolonged interaction among communities has resulted in overlapping linguistic features, complicating clear-cut classifications. These methods allow researchers to model language evolution quantitatively, providing new perspectives on subgrouping and divergence. Nevertheless, their effectiveness depends on the availability of high-quality linguistic data, which remains scarce for many Bornean languages. Thus, the literature suggests that classification should not be viewed as a purely technical exercise but as a dynamic process that integrates linguistic, historical, and ethnographic evidence. This aligns with current trends in historical linguistics that emphasise interdisciplinary methodologies.

A significant theme emerging from recent studies is the role of digital technologies in language preservation. Abingosa et al. (2025) demonstrate how mobile applications can facilitate language learning and documentation, particularly among younger generations. Such tools not only enhance accessibility but also promote community engagement. Similarly, Tasie (2025) explores the potential of artificial intelligence (AI) to support digital communication for minority languages. Their findings suggest that AI-driven platforms can assist in language revitalisation by enabling automated translation, speech recognition, and corpus development. However, the authors caution that technological solutions must be culturally sensitive and community-oriented to avoid reinforcing linguistic inequalities. Digital archives and online repositories are also increasingly recognised as essential components of preservation strategies. These platforms provide long-term storage and global accessibility of linguistic data, thereby supporting both academic research and community revitalisation efforts. Despite these advancements, the literature points to several limitations. Digital initiatives often face challenges related to funding, technical expertise, and infrastructure, particularly in remote regions of Borneo. Moreover, the sustainability of such projects depends on continuous community involvement and institutional support.

Another critical strand in the literature emphasises the importance of community participation in language preservation. Contemporary approaches advocate collaborative models in which local communities are actively involved in documentation, analysis, and revitalisation. Sharma & Ram (2025) argue that empowering indigenous communities through education and capacity-building initiatives is essential for sustainable language maintenance. This includes training community members in linguistic documentation techniques and integrating local languages into school curricula. In addition, participatory approaches recognise the role of oral traditions, cultural practices, and identity in language preservation. By situating language within its cultural context, these approaches foster a sense of ownership and pride among speakers, which is crucial for intergenerational transmission. For Borneo, this perspective is particularly relevant given the strong link between language, land, and cultural identity among indigenous groups. Preservation strategies that fail to account for these connections risk being ineffective.

Synthesizing the literature, it becomes evident that effective preservation of endangered languages in Borneo requires an integrated framework that combines rigorous classification methods with innovative and community-driven strategies. While traditional linguistic analysis remains crucial, it must be complemented by digital tools, participatory approaches, and interdisciplinary collaboration. Importantly, the literature suggests that preservation efforts should move beyond reactive documentation towards proactive revitalisation. This involves not only recording languages but also creating environments in which they can be actively used and transmitted.

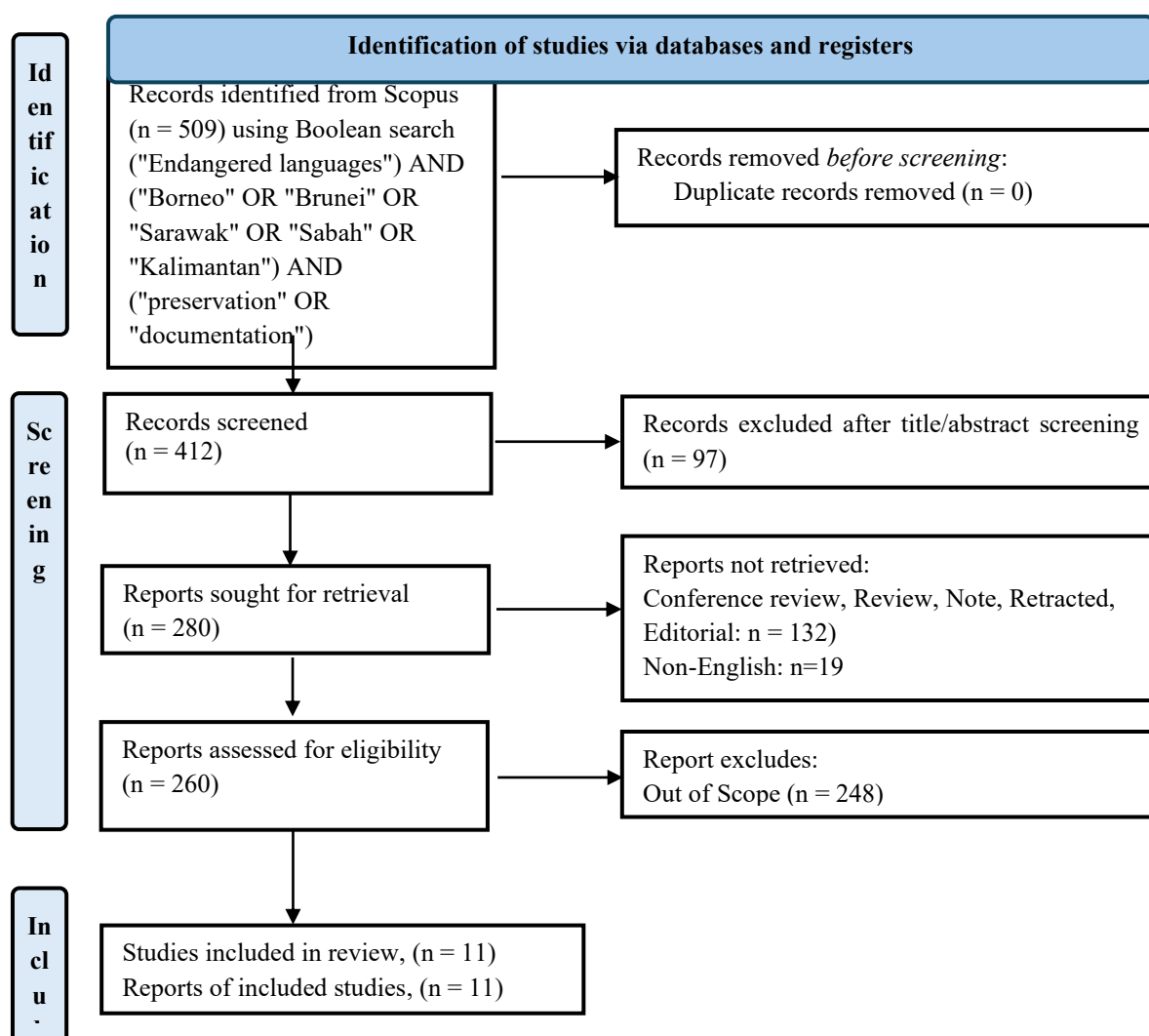
METHODOLOGY

This study adopted a qualitative systematic review design to synthesise existing research on language classification and preservation of endangered languages in Borneo. The systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Ghamrawi et al., 2025) to ensure transparency, reproducibility, and methodological rigour. The search employed the following Boolean string: ("Endangered languages") AND ("Borneo" OR "Brunei" OR "Sarawak" OR "Sabah" OR "Kalimantan") AND ("preservation" OR "documentation"). This string was designed to capture studies addressing the two core dimensions of the review: language classification and preservation practice. Searches were restricted to English- and Malay-language publications published between 2000

and 2026. The reference lists of included studies were hand-searched to identify additional relevant publications not retrieved during the database search.

Studies were included if they: (1) focused on one or more endangered or minority languages indigenous to Borneo; (2) addressed language classification through comparative-historical, typological, or addressed language preservation through documentation, revitalisation and community involvement; (3) were published in peer-reviewed journals and edited volumes; and (4) were published between 2000 and 2026. Studies were excluded if they: (1) focused exclusively on dominant languages such as Standard Malay and Bahasa Indonesia, without reference to endangered minority languages; (2) lacked sufficient methodological information to allow quality assessment; (3) were conference abstracts without accompanying full texts; or (4) addressed Borneo only incidentally within broader Southeast Asian or pan-Austronesian discussions without providing language-specific data.

The Scopus database search using the specified Boolean strategy yielded 509 records. No duplicate records were identified or removed before screening. During the title and abstract screening phase, 97 records were excluded for lack of relevance to endangered-language classification and preservation in the Borneo context. This resulted in 412 records proceeding to the screening stage, of which 280 reports were sought for retrieval. However, 151 reports were not retrieved, comprising conference reviews, general reviews, notes, retracted articles, and editorials ($n = 132$), as well as non-English publications ($n = 19$). Consequently, 260 full-text reports were successfully assessed for eligibility. At this stage, a substantial number of studies ($n = 248$) were excluded for being out of scope, particularly those that did not directly address endangered languages, classification frameworks, or preservation strategies in Borneo. Ultimately, 11 studies met all inclusion criteria and were retained for the final synthesis. These studies were subsequently subjected to quality assessment across five dimensions: research design clarity, methodological transparency, data adequacy, analytical rigour, and relevance of findings. The PRISMA flow for this review is presented as follows:



FINDINGS

The systematic review identified 11 Scopus-indexed studies meeting the inclusion criteria (see Table 1), representing a diverse body of research on endangered languages across Borneo from 2000 to 2026, with a notable concentration in the post-2019 period. Geographically, the studies are distributed across Sarawak, Brunei, and Sabah, with the strongest concentration in Sarawak, followed by Brunei, and a smaller representation from Sabah. The predominance of Sarawak-based studies reflects the relative strength of documentation infrastructure and linguistic research traditions in that state (Norahim & Badawi, 2024). Methodologically, the corpus is diverse, incorporating surveys, descriptive linguistic analyses, documentary fieldwork, collaborative orthography development, social network analysis, and comparative policy review. The synthesis reveals two thematic domains aligned with the study's research questions: (1) language classification approaches and endangerment assessment, and (2) preservation and revitalisation strategies.

Table 1. Summary of included Scopus studies on endangered languages in Borneo.

Author (Year)	Context / Focus	Method & Sample	Thematic Code(s)	Technology	Key Outcomes
Smith & Smith (2017)	Sabah & Sarawak (138 languages)	Descriptive review	Language revitalisation; Identity; Indigenous language development	Writing systems; L1 literacy programmes	Indigenous communities developing writing systems and early childhood education; revitalisation awareness growing; shift toward Malay documented.
Lendik & Yuit (2021)	Sarawak (Kenyah Long Wat)	Descriptive linguistic analysis	Language documentation; Endangered features; Linguistic identity	None reported	Epithet forms documented as endangered discourse feature; national language dominance drives erosion; ideological recalibration needed.
McLellan & Jones (2015)	Borneo (Brunei, Sarawak)	Comparative policy analysis	Language policy; Community-based revitalisation; Revitalisation strategies	Policy instruments	Community-based bottom-up initiatives more sustainable than top-down policy; local participation is critical for long-term outcomes.
Burkhardt & Burkhardt (2019)	Sarawak (Berawan; <4,000 speakers; 3 communities)	Collaborative orthography development	Language documentation; Orthography; Revitalisation	Writing system development tool	Unified orthography developed collaboratively across three communities; adequate testing and time critical; new implementation factors proposed.
Sercombe (2010)	Borneo (Eastern Penan)	Qualitative descriptive	Language endangerment; Literacy; Documentation	Lexicographic tool; L1 literacy materials	Eastern Penan most socioeconomically marginalised in Borneo; lexicographic project supports L1 literacy, maintenance, and second language learning.
Coluzzi (2010)	Brunei (Iban & Murut/Lun Bawang, Temburong)	Survey; 2008 fieldwork	Language shift; Sociolinguistic vitality; Endangerment	Survey instrument	Significant shift toward Malay among youth; Iban more endangered than Murut; urban migration accelerates shift.

Author (Year)	Context / Focus	Method & Sample	Thematic Code(s)	Technology	Key Outcomes
Mohamed & Hashim (2012)	Sarawak (Sihan/Sian community)	Field survey, observation, interviews	Language vitality; UNESCO criteria; Endangerment classification	Survey instrument; UNESCO 9-criteria framework	Sihan fails all 9 UNESCO vitality criteria; severely threatened; community attitudes and transmission both critically low.
Norahim & Badawi (2024)	Sarawak (~51 indigenous languages)	Documentary linguistics; Descriptive review	Language documentation; Oral literature; Under-resourced languages	Documentation platforms; proposed digital archives	Majority of Sarawak languages undescribed; oral literature unsystematically documented; multidimensional situated documentation model proposed.
Joan & Ting (2025)	Sarawak (Kejaman, Belaga; n=123; 3 generations)	Social network analysis; Survey	Language maintenance; Social networks; Intergenerational transmission	Survey instrument; network analysis	Low-density uniplex networks insufficient to sustain Kejaman; Generation 3 consults non-kin in other languages; maintenance at critical risk.
Yusof & Alas (2021)	Brunei (Dusun at UBD Language Centre)	Mixed methods: experiment + focus group; n=students	Language teaching; Pedagogy; Revitalisation	Role-play technique; oral pre/post assessment	Role-play significantly improves oral Dusun competence ($p<0.05$); students report enhanced engagement, confidence, and memorisation.
Dani et al. (2019)	Sabah (Dusun, rural; n=120, age ~15)	Survey	Intergenerational transmission; Language revitalisation; Endangerment	ICT/digital media (proposed); survey instrument	Dusun domains severely limited; home transmission failing; three revitalisation proposals: status raising, school instruction, ICT utilisation.

Language Classification and Endangerment Assessment (RQ1)

Several included studies employ descriptive linguistic analysis as a basis for language classification and documentation, directly addressing RQ1. Lendik and Yuit (2021), in their analysis of Kenyah Long Wat, document distinctive descriptive forms as linguistic features under threat, revealing how the erosion of such features signals advanced endangerment. Their study underscores that documentation must attend not only to core lexical and phonological inventories but also to pragmatic and discourse-level features encoding cultural identity. Similarly, Norahim and Badawi (2024) surveyed approximately 51 indigenous languages of Sarawak, demonstrating that most remain substantially undescribed. Their work identifies classification gaps and calls for a multidimensional documentation model integrating descriptive linguistics with oral literature and community linguistic practices.

Melaia and Rigarb (2019) provide a preliminary documentation of the Punan language in Belaga, distinguishing it genealogically and culturally from the Penan. Their fieldwork-based data, drawn from eight longhouse communities, identify four linguistic variations within Punan, contributing to its classification as a distinct endangered language requiring urgent systematic attention. Burkhardt and Burkhardt (2019) adopt a collaborative approach to classification through orthography development for Berawan, a Bornean language with fewer than 4,000 speakers. Their work demonstrates how orthographic standardisation reflects and reinforces the classification of varieties as related but distinct, facilitating both documentation and revitalisation simultaneously. Florey (2010), in a broad Austronesian comparative perspective, situates Bornean languages within the wider genealogical and typological landscape, providing methodological frameworks applicable to classification and capacity-building across the region.

Beyond formal linguistic classification, several included studies employ sociolinguistic vitality frameworks as systematic instruments for assessing endangerment levels, complementing structural

classification with social and functional dimensions. Mohamed and Hashim (2012) apply UNESCO's nine criteria for language vitality to the Sihan community in Sarawak, assessing intergenerational transmission, absolute speaker numbers, domain use, availability of materials, and community attitudes. Their findings reveal that Sihan fails to meet all nine criteria, placing it in a severely threatened category. This framework provides a replicable classification instrument applicable to other under-documented Bornean languages.

Coluzzi (2010) conducts a survey-based sociolinguistic study among Iban and Murut (Lun Bawang) communities in Brunei's Temburong district, comparing language use and attitudes across age cohorts. The findings reveal a significant generational divide, with younger speakers demonstrating an accelerated shift towards Malay, and highlight differential endangerment between the two communities, with Iban facing a more acute decline. Joan and Ting (2025) employ social network analysis among 123 Kejaman speakers in Sarawak to assess transmission patterns across three generations. Their data demonstrate that Kejaman belongs to a low-density, uniplex social network community providing insufficient contact density to sustain intergenerational language transmission, offering a quantitative complement to sociolinguistic vitality frameworks.

An emerging convergence in the included literature is the intentional integration of language classification with revitalisation-oriented goals, challenging the traditional disciplinary boundary between descriptive linguistics and applied language work. Smith and Smith (2017) provide a panoramic comparative assessment of East Malaysian language endangerment, observing that of the 138 living languages across Malaysia, a substantial proportion are endangered or are actively shifting towards Malay and English. Critically, their review demonstrates that indigenous revitalisation efforts, including the development of writing systems, vocabulary documentation, and early childhood education programmes, rely on and operationalise prior structural classification of the target language. In this sense, classification is not a precondition that precedes preservation but a process that unfolds iteratively alongside it, refined through community engagement and pedagogical practice.

Sercombe (2010) demonstrate this integrated approach through a modest lexicographic project for the Eastern Penan. The project documents the language's lexical inventory, with the explicit dual purpose of supporting first-language literacy development among lower primary school children and facilitating second- and foreign-language acquisition. By situating lexicographic work within a literacy and educational framework, Sercombe (2010) illustrates how classification outputs, in this case, a structured lexical record, directly serve the practical transmission needs of the speech community. The study thus reinforces the principle that classification is most socially meaningful when it is embedded in, rather than abstracted from, community-oriented preservation contexts.

Burkhardt and Burkhardt (2019) extend this logic further by outlining the development of a unified orthography for Berawan, an endangered Bornean language spoken by fewer than 4,000 members across four varieties in Sarawak: Long Jegan, Batu Belah, Long Teru, and Long Terawan. Applying and adapting Smalley's (1959, 1965) criteria for orthography development within a collaborative community framework involving three Berawan communities, the authors demonstrate how the process of making classificatory decisions, specifically determining how to represent phonological distinctions and dialectal variation in a writing system, simultaneously constitutes a revitalisation intervention by producing a well-documented, usable cultural artefact. Their study also proposes two new factors for orthographic implementation: the need for adequate testing to inform decision-making and ensure orthographic effectiveness, and the importance of allowing sufficient time for community members to learn and adopt the developed system. Together, these cases substantiate McLellan and Jones's (2015) observation that the most effective and sustainable revitalisation outcomes in Borneo emerge when linguistic analysis and community participation are treated not as sequential phases but as mutually constitutive dimensions of a single, integrated process.

Patterns of Language Endangerment

The decline of intergenerational transmission emerges as the most extensive and structurally decisive driver of language endangerment across the included studies, with convergent evidence documented across Brunei, Sarawak, and Sabah. Coluzzi (2010) presents one of the most systematic surveys of this phenomenon, comparing language use and attitudes among the Iban and Murut (Lun Bawang) communities in Brunei's Temburong district based on 2008 fieldwork. His findings reveal a distinct generational divide: younger speakers in both communities demonstrate significantly lower rates of heritage language use than the older generation, with the shift consistently and uniformly toward Malay, the national language. Notably, Iban is found to be in a more acute state of endangerment than Murut, indicating that even within the same geopolitical and social context, differential rates of transmission decline operate according to community-specific variables, including social network density and domain of use.

This pattern of transmission failure is demonstrated and extended by Dani et al. (2019), whose survey of 120 Dusun students aged approximately 15 years in a rural area of Sabah documents a critical weakening of the home as a site of heritage-language use. Despite the implementation of preschool and post-school reinforcement programmes, the majority of respondents report that the trend in existing Dusun language domains is highly limited, and that the utilisation of Dusun in new communicative domains and media remains minimal. The study underscores a critical finding: changes in social needs have triggered linguistic changes that extend well beyond simple vocabulary substitution, reshaping the broader vitality of the Dusun linguistic system. This suggests that intergenerational transmission breakdown is not merely a quantitative reduction in speaker numbers but a qualitative restructuring of the communicative ecology in which the language once functioned.

Joan and Ting (2025) provide the most structurally refined analysis of transmission dynamics among the included studies, employing social network analysis to examine intergenerational language use patterns among 123 Kejaman speakers across three generations in two longhouses in Belaga, Sarawak. Their findings reveal that across all three generations, exchange networks ($M = 131.7$) predominate over interactive networks ($M = 110.3$), indicating a community structure centred on kin-based consultation rather than broader social interaction. Although kinship ties still play a role in how the Kejaman people use their language, particularly in family-related matters, cultural practices, and taboos, speakers from the third generation are increasingly turning to non-kin for advice on a wider range of topics. Taken together, these findings substantiate Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS), situating most documented communities at advanced disruption stages, where casual, day-to-day use has substantially eroded.

Beyond demographic transmission dynamics, the included studies consistently demonstrate that intersecting socioeconomic forces and institutional language policy arrangements structurally condition language endangerment in Borneo. McLellan and Jones (2015) situate this dimension most explicitly, documenting that across Brunei and Sarawak. In the broader Bornean region, indigenous and minority languages are systematically marginalised within national education systems, receiving limited or no instructional time. At the same time, standardised varieties of Malay, including Bahasa Indonesia, Bahasa Malaysia, and Bahasa Melayu in Brunei, are supported as languages of instruction and receive exclusive institutional resources. This structural asymmetry contributes to uneven intergenerational positioning for minority language speakers, as their heritage languages receive limited institutional recognition and are less closely aligned to education and employment associated with dominant and national languages.

Smith and Smith (2017) supported this analysis from a broader East Malaysian perspective, noting that many indigenous parents have internalised the logic of national language dominance to the extent that they proactively speak Malay to their children in the home domain, calculating that this gives their children a stronger start in the government school system. This finding is particularly significant because it reveals how macro-level policy choices produce micro-level behavioural adaptations that accelerate the shift even in the absence of explicit assimilation pressure. The decision by parents to abandon heritage language transmission in favour of perceived educational advantage demonstrates how language policy operates not only through formal institutions but through the rationalised choices of individual speakers navigating socioeconomic realities. Dani et al. (2019) add further discussion to this picture by documenting that, among Dusun communities in Sabah, social pressures associated with employment, education, and geographic mobility have triggered linguistic changes that extend beyond vocabulary additions, producing a comprehensive restructuring of the language's vitality profile. Their findings show that even where school-based reinforcement programmes exist, they are insufficient to counteract the broader socioeconomic pull towards Malay and English as languages of upward mobility. Lendik and Yuit (2021) complement these empirical accounts with a critical theoretical framing, arguing that the institutions of the modern nation state and the national language constitute fundamental structural threats to ethnolinguistic minorities, as their establishment entails the risk of either exclusion or assimilation.

Preservation and Revitalisation Strategies (RQ2)

Documentation constitutes the most extensively and consistently represented preservation strategy across the 11 included studies, functioning as both an end in itself, capturing irreplaceable linguistic and cultural data, and as the foundational precondition for subsequent revitalisation work, directly addressing RQ2. Norahim and Badawi (2024) provide a thorough overview of documentation practices for approximately 51 indigenous languages in Sarawak, Borneo. Their research shows that only a few languages have been extensively described, and crucial oral literature, key to cultural heritage and linguistic data, has not been systematically documented. To fill this gap, they propose a multidimensional documentation model that

goes beyond structural linguistic description to include linguistic practices, oral traditions, and the sociocultural contexts of speaker communities. This model aims to create "comprehensive, everlasting records of linguistic practices," contributing to the geo- and historical linguistics of the region while offering communities long-term language resources.

Burkhardt and Burkhardt (2019) illustrate how documentation can be productively embedded within a collaborative community process rather than conducted as an external academic exercise. Their development of a unified orthography for Berawan involved systematic, collaborative sessions with representatives from three community varieties, generating documentation of phonological inventories, dialectal variation, and orthographic preferences as by-products of the community engagement process itself. The study's adaptation of Smalley's orthographic criteria to the Berawan context demonstrates that rigorous documentation methodology should not be separated from community participation; rather, structured collaborative documentation can produce both academically valid linguistic records and community-owned written language artefacts simultaneously.

Sercombe (2010) demonstrates a further dimension of documentation practice through his account of a lexicographic project for the Eastern Penan. The project aims to complement the development of first-language literacy among lower primary school children while simultaneously fostering the broader maintenance of Eastern Penan and facilitating second-language acquisition. This study illustrates that documentation outputs are most impactful when designed from the outset with community utility in mind: a lexicographic resource that directly supports literacy instruction and first language development represents a fundamentally different preservation intervention than a documentation archive produced primarily for academic circulation. Lendik and Yuit (2021) reinforce this point by demonstrating that effective documentation must attend to the full range of a language's expressive resources, including discourse-level and pragmatic features, which encode cultural identity and social relations. Their preliminary study of Kenyah Long Wat demonstrates that the erosion of such features, driven by the dominance of national language ideologies, constitutes a form of language endangerment that purely structural documentation approaches may fail to capture.

Findings consistently support that community participation is crucial to effective language revitalisation. McLellan and Jones (2015) conduct a comparative policy analysis across Brunei and Sarawak, demonstrating that community-based, bottom-up initiatives consistently generate more sustainable and culturally responsive outcomes than top-down governmental policies. Their analysis identifies several key revitalisation strategies, including community-led language classes, intergenerational transmission programmes, the development of locally grounded teaching materials, the incorporation of language use into cultural practices, and grassroots advocacy efforts. Their analysis of five revitalisation strategies shows that the most durable results come from those embedded in community agency and local participation. Smith and Smith (2017) provide supported evidence from Sabah and Sarawak, where indigenous communities have independently developed writing systems, linguistic reference materials, and early childhood education programmes, with success closely linked to community ownership of the revitalisation process.

Yusof and Alas (2021) provide experimental evidence for the effectiveness of community-oriented revitalisation within formal education, demonstrating that role-play significantly improves students' oral competence in Dusun and that students report heightened engagement and motivation. Florey (2010) further demonstrates that community members serving as co-researchers in documentation projects strengthen both the quality of documentation outputs and the social prestige of the language within the community. Collectively, these findings reinforce the idea that the most effective revitalisation outcomes emerge from models that empower communities as agents rather than position them as passive beneficiaries.

Technology-mediated language preservation is represented across several included studies. Dani et al. (2019) articulate the most explicit technological proposal in the corpus, recommending the utilisation of information and communication technology (ICT) and digital media as one of three strategic interventions for revitalising the Dusun language in Sabah, alongside raising the language's societal status and implementing systematic school instruction. By proposing ICT as a vehicle for revitalisation, Dani et al. (2019) recognise that language presence in digital domains is no longer optional but structurally significant for intergenerational relevance. Yusof and Alas (2021) integrate multimedia tools within role-play-based Dusun instruction, demonstrating that technologically supported pedagogy can enhance engagement and learning outcomes. Norahim and Badawi (2024) identify digital archiving and online platforms as essential

components of a multidimensional documentation model, noting that repositories extend the longevity of documentation outputs to community and global-access networks.

Despite these contributions, the included studies collectively indicate that technological solutions in Borneo are evolving yet unevenly distributed. No included study reports a fully operational digital language preservation platform for any Bornean endangered language; instead, technology features as a proposed or supplementary component within broader documentation and revitalisation frameworks. Implementation faces constraints related to infrastructure access, digital literacy, funding sustainability, and the availability of language-specific technical expertise in remote communities. Consequently, technology should be conceptualised as an enabling platform that supports, rather than replaces, face-to-face transmission, embodied cultural practices, and everyday conversational use within speech communities.

DISCUSSION

The synthesis of the 11 reviewed studies shows that language classification and preservation in Borneo are deeply interconnected and mutually reinforcing processes. Addressing RQ1, the findings demonstrate that classification in the included corpus is conducted through a range of methodologies, from formal descriptive linguistics and orthography development to sociolinguistic vitality assessment and social network analysis. The most robust classification outcomes emerge from studies that triangulate the structural, functional, and social dimensions of language. UNESCO vitality frameworks (Mohamed & Hashim, 2012; Coluzzi, 2010) and social network analysis (Joan & Ting, 2025) offer particularly systematic and replicable instruments for classification of endangerment levels applicable across Borneo's diverse linguistic contexts.

Addressing RQ2, the findings show that documentation, community-based revitalisation, and technological interventions form the three pillars of preservation practice in Borneo, with community participation emerging consistently as the key factor distinguishing sustainable from short-lived initiatives (McLellan & Jones, 2015). The interaction between classification and preservation is iterative: improved classification enhances targeted documentation strategies, while preservation activities generate new empirical data that refine classification. This dynamic relationship underscores the need for integrated approaches combining comparative linguistics, sociolinguistics, and applied documentation within a unified analytical framework. Recent research focusing on Brunei provides more context-specific evidence on language maintenance dynamics. Zulkiflee and Chuchu (2025a) demonstrate that the Belait language, despite its critically endangered status, continues to be maintained through selective intergenerational transmission within tightly knit family networks. However, this transmission is neither uniform nor stable, as younger speakers increasingly prioritise dominant languages such as Malay and English in education and employment domains. The study highlights that language maintenance is sustained not through frequency of use, but through culturally embedded practices, identity alignment, and community attitudes toward linguistic heritage.

Regional disparities are evident across the included studies. Sarawak shows relatively stronger documentation and revitalisation infrastructure, with active community literacy programmes, orthography development, and NGO-backed projects (Burkhardt & Burkhardt, 2019; Norahim & Badawi, 2024; McLellan & Jones, 2015). In contrast, Brunei-based minorities face more acute shift dynamics driven by Malay dominance and urbanisation (Coluzzi, 2010; Sharbawi & Hassan, 2024; Ho, 2026), while Sabah presents mixed profiles with documentation coverage remaining uneven across interior language communities (Dani et al., 2019; Smith & Smith, 2017). These contrasts have direct policy implications, suggesting that a coordinated regional strategy is needed rather than nation-state-specific approaches that fail to account for trans-border linguistic communities.

The findings underscore several directions for future research. First, more extensive documentation is needed for critically under-researched languages, particularly in Kalimantan and interior Sabah. Second, further investigation into the sociolinguistic dynamics of the shift can inform more effective revitalisation design. Zulkiflee and Chuchu (2025b), in their study of the Kiput communities in Kuala Tutoh and Benawa, demonstrate that language vitality is deeply intertwined with cultural practices, kinship structures, and spatial settlement patterns. Third, policy-oriented research should develop frameworks supporting minority language maintenance across Borneo's distinct political influences. Finally, the integration of computational methods with traditional linguistic analysis offers a promising approach for improving both classification accuracy and preservation scalability.

CONCLUSION

This systematic review demonstrates that endangered languages in Borneo are shaped by a complex interplay of genealogical relationships, contact-induced change, and rapidly shifting sociolinguistic ecologies, requiring integrated approaches that bridge linguistic classification and preservation practices. The synthesis of 11 Scopus-indexed studies reveals that descriptive, documentary, and sociolinguistic vitality frameworks are the primary instruments for language classification in the literature. At the same time, preservation efforts range from community-led revitalisation and intergenerational transmission initiatives to digital archiving and technology-enhanced documentation. The two research questions guiding this review have been addressed: RQ1 reveals a methodologically diverse but increasingly sociolinguistically oriented classification landscape in which vitality assessment frameworks complement formal structural analysis; RQ2 demonstrates that community ownership and participation are the strongest predictors of preservation effectiveness across the region. Regional disparities are evident, with Sarawak exhibiting relatively stronger documentation ecosystems than Brunei, Sabah, and Kalimantan, where linguistic data remain underdocumented. Language endangerment in Borneo is deeply embedded in broader socio-cultural, political, and economic transformations, including urbanisation, language policy orientation, and shifting identity practices. This synthesis suggests that the future of endangered language research in Borneo depends on the development of coordinated, interdisciplinary, and sustainable frameworks that bridge theoretical linguistics, applied documentation, and sociocultural engagement in a coherent, policy-relevant manner.

REFERENCES

- Abingosa, D., Bokingito Jr, P., Pasandalan, S. N., Alovera, J. R. G., & Otano, J. (2025, September). Digitising the Higaonon language: a mobile application for indigenous preservation in the Philippines. In *Informatics* (Vol. 12, No. 3, p. 90). MDPI.
- Adelaar, K. A., & Himmelmann, N. (2004). *The Austronesian languages of Asia and Madagascar*. Routledge.
- Alisoy, H. (2025). Multilingualism: Cognitive, social, and educational dimensions in a globalised world. *Journal of Azerbaijan Language and Education Studies*, 2(2), 28–38.
- Blust, R. (2013). *The Austronesian languages*. Asia-Pacific Linguistics, School of Culture, History and Language, College of Asia and the Pacific, The Australian National University.
- Burkhardt, J. L., & Burkhardt, J. M. (2019). Developing a unified orthography for Berawan: An endangered Bornean language. *Written Language & Literacy*, 22(2), 280–306.
- Coluzzi, P. (2010). Endangered languages in Borneo: a survey among the Iban and Murut (Lun Bawang) in Temburong, Brunei. *Oceanic Linguistics*, 119–143.
- Damit, A. R., & Zulkiflee, Z. (2025). Language shifts and transnational migration in the Belait community: A history of resilience and adaptation. In *Transnational Southeast Asia: Communities, Contestations and Cultures* (pp. 121–142). Singapore: Springer Nature Singapore.
- Dani, N. A., Kiram, N. M., Ombi, K., Suhailin, R., Sha'ri, S. N., & Paul, D. (2019). Endangered Intergenerational Language Transmission: Evidence from the Indigenous Dusun Society of Sabah, Malaysia. *Pertanika Journal of Social Sciences & Humanities*, 27(1).
- Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages* (Vol. 76). Multilingual matters.
- Florey, M. (Ed.). (2010). *Endangered languages of Austronesia*. Oxford University Press.
- Ghamrawi, N., Shal, T., Ghamrawi, N. A., Abu-Tineh, A., Alshaboul, Y., & Alazaizeh, M. A. (2025). A step-by-step approach to systematic reviews in educational research. *European Journal of Educational Research*, 14(2), 549–566.
- Gu, C. (2025). Scripting English in Jawi: English disguised in Arabic-based ‘Tulisan Jawi’ in Brunei’s linguistic landscape. *Asian Englishes*, 27(3), 591–624.
- Ho, D. G. (2026). The emerging influence of Mandarin: The catalyst for the critically endangered state of Chinese dialects in Brunei?. *The Chinese in Brunei*, 159–183.
- Joan, A., & Ting, S. H. (2025). Social networks and maintenance of an endangered language: the Kejaman of Malaysia. *Southeast Asia: A Multidisciplinary Journal*, 25(1), 42–57.
- Lendik, L. S., & Chan, M. Y. (2021). A preliminary study on the use of epithets in Kenyah Long Wat. *The Journal of Asian Linguistic Anthropology*, 3(1), 56–75.

- McLellan, J., & Jones, G. (2015). Maintaining and revitalising the indigenous endangered languages of Borneo: Comparing 'top-down' and community-based policy initiatives and strategies. *Policy and planning for endangered languages*, 18–30.
- Melaia, Z., & Rigarb, A. (2019). Moribund Language Documentation and Preservation: A Preliminary Study on the Punan Language. *The GLOCAL 2019*, 239.
- Mohamed, N., & Hashim, N. H. (2012). Language vitality of the Sihan community in Sarawak, Malaysia. *KEMANUSIAAN: The Asian Journal of Humanities*, 19(1), 59–86.
- Moseley, C. (Ed.). (2010). *Atlas of the World's Languages in Danger*. UNESCO.
- Norahim, N., & Badawi, H. (2024). Documenting Under-resourced Languages of Sarawak, Borneo. *The Journal of Asian Linguistic Anthropology*, 6(2), 37–57.
- Omar, A. H., & Norahim, N. (2020). The GLOCAL Proceedings Asia 2020.
- Sercombe, P. G. (2009). To be or not to be: challenges facing Eastern Penan in Borneo. *Endangered Languages of Austronesia*, 191.
- Sharbawi, S., & Hassan, N. H. (2024). The Malays of Brunei: An investigation of their family language policies. *Journal of Multilingual and Multicultural Development*, 45(4), 884–905.
- Sharma, S., & Ram, Y. (2025). Revitalising Indigenous Languages: Analysing Key Factors and Strategies in the Context of NEP 2020. *International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)*, 2(2), 348–355.
- Smalley, W. A. (1959). How shall I write this language?. *The Bible Translator*, 10(2), 49–69.
- Smalley, W. A. (1965). *Orthographic studies: Articles on new writing systems*. London: United Bible Societies.
- Smith, A. D. (2017). The languages of Borneo: A comprehensive classification. *Unpublished doctoral dissertation. University of Hawai'i at Mānoa, Honolulu, HI*.
- Smith, J. A., & Smith, K. J. (2017). Indigenous language development in East Malaysia. *International Journal of the Sociology of Language*, 2017(244), 119–135.
- Tasie, C. C. (2025). Artificial Intelligence (AI): An Innovative Tool for the Revitalisation of the Ikwerre Language. *LALICO Journal of Languages, Literature, and Communication*, 3(2), 113–120.
- Umar, H. A., Shehu, J. A., & Ahmad, A. B. (2025). Dynamics of language evolution: exploring sociolinguistic factors, perceptions, and barriers from theoretical foundations. *Aminu Kano Academic Scholars Association Multidisciplinary Journal*, 1(3), 55–72.
- Yusof, N., & Alas, Y. (2021). Using Role-Play to Teach Minority Ethnic Languages: A Case Study at Universiti Brunei Darussalam. In *Globalisation, Education, and Reform in Brunei Darussalam* (pp. 177–196). Cham: Springer International Publishing.
- Zaini, A. I., Karsono, B., & Awang Sulong, A. H. (2018, December). From sacred to secular: Baruk architecture on secular buildings in Sarawak. In *IOP Conference Series: Earth and Environmental Science* (Vol. 213, No. 1, p. 012043). IOP Publishing.
- Zulkiflee, Z., & Chuchu, S. I. F. H. (2025a). Belait language maintenance in Brunei Darussalam. *International Journal of Linguistics and Literature*, 3(1), 77–92.
- Zulkiflee, Z., & Chuchu, I. F. (2025b). Sociocultural Perspective Among the Kiput of Baram: A Study of Kuala Tutoh and Benawa Communities. *TAMADDUN: Jurnal Ilmu Sosial, Seni, dan Humaniora*, 3(3), 132–146.

CURRICULUM VITAE

Complete Name : Zulfadzlee Zulkiflee
 Institution : Universiti Brunei Darussalam
 Education : Sedang menjalani PhD (Malay Language & Linguistics)
 Research Interests : Historical linguistics