



Reflexive Fieldwork in the Baram–Tinjar Region: Documenting Minority Languages

Zulfadzlee Zulkiflee

Malay Language and Linguistic, Faculty of Arts and Social Sciences, University Brunei Darussalam,
Jalan Tungku Link Gadong BE 1410, Brunei Darussalam.

*Corresponding Author e-mail: 24zulfadzlee@gmail.com

Received: June 2026; Revised: July 2026; Published: July 2026

Abstract

The Baram–Tinjar region of northern Sarawak is among Borneo's most linguistically diverse areas, yet it remains severely under-documented in comparative linguistics, despite sporadic scholarly attention since the mid-nineteenth century. This study explores how reflexive, ethnographically grounded fieldwork practices influence the quality, reliability, and comparative value of phonological and lexical data collected from severely underdocumented minority languages. The data presented from doctoral fieldwork conducted among minority languages in north Sarawak, including Narum, Kayan, Kiput, Miriek, Berawan, Bakong, and Tireng, documenting phonological systems, lexical inventories, and narrative data through ethnographic engagement. Framed as a comparative-historical linguistic study, the article examines fieldwork methodology, researcher positionality, relationship-building, and the social conditions of data collection as variables that directly affect data quality and historical-linguistic findings. Findings show that sustained, relationship-based fieldwork (rather than single-visit elicitation) yields more naturalistic and reliable comparative data, and that researcher reflexivity functions as a methodological control, not merely an ethical disclosure. Building on these findings, the study discusses how the resulting documentation can also support community-oriented revitalisation outputs, including pedagogical materials and local archives, as applications of documentation rather than the study's primary research aim. Overall, the study demonstrates that integrating comparative linguistics with reflexive fieldwork yields more defensible historical-linguistic data and generates practical resources for the documented communities.

Keywords: Ayoreo; Zamucoan; Definiteness; Discursive Anchorage; Articles

How to Cite: Zulkiflee, Z. (2026). Reflexive Fieldwork in the Baram–Tinjar Region: Documenting Minority Languages. *International Journal of Linguistics and Indigenous Culture*, 4(2), 134-148. <https://doi.org/10.36312/ijlic.v4i2.5587>



<https://doi.org/10.36312/ijlic.v4i2.5587>

Copyright© 2026, Zulkiflee.

This is an open-access article under the [CC-BY-SA](#) License.



INTRODUCTION

The Baram–Tinjar region of northern Sarawak represents one of Borneo's most linguistically complex and under-documented areas. This region, situated along the Malaysia–Indonesia border, is characterised by extraordinary linguistic diversity, with multiple language families coexisting in close contact zones. Despite its significance for comparative Austronesian linguistics and historical language reconstruction, the region has received sporadic scholarly attention since the mid-nineteenth century, and many of its minority languages remain inadequately described from comparative-historical perspectives. This article offers an ethnographically based account of

doctoral fieldwork conducted among minority-language communities in northern Sarawak. It examines the processes, challenges, and insights that arise from sustained engagement with the Narum, Kayan, Kiput, Miriek, Berawan, Bakong, and Tireng languages.

Many minority languages worldwide, such as Ainu in Japan and various languages in sub-Saharan Africa, face similar threats of extinction (Whaley, 2003). This situation differs from the Baram–Tinjar context in three ways that directly affect scholars' attention. First, many of these languages are supported by existing institutional documentation infrastructures, including national revitalisation funding, established descriptive grammars, or longitudinal linguistic survey programs; such resources are currently absent for languages like Kiput, Kayan and Tireng. Second, the Baram–Tinjar languages occupy a position of critical discussion within the historical classification of Austronesian languages (Blust, 2013; Kroeger, 1998). Given that comparative data from this language group are essential for examining competing subgrouping hypotheses, the extinction of these languages would not only eliminate linguistic systems but also prevent the acquisition of evidence vital to reconstructing the broader Malayo-Polynesian family. Third, the opportunity to document this language is rapidly declining. Most fluent speakers belong to the older generation, and unlike more well-resourced contexts for endangered languages, there is a significant lack of audio recordings, transcriptions, or phonological records. Hence, it is significant to examine this language group to analyse its phonological and lexical structures, capture variation across speakers, and preserve critically endangered linguistic knowledge before it is irretrievably extinct.

Contemporary language documentation practice increasingly emphasises the role of reflexivity in fieldwork methodology, recognising that researcher positionality, community relationships, and contextual factors fundamentally shape the data collected and its interpretation (Everett & Everett, 2016; Haspelmath & Sims, 2010). A significant number of the published literature on Baram–Tinjar languages has proceeded without thorough reflection on the fieldwork experience itself or the social contexts that generate linguistic data. This article addresses that gap by foregrounding fieldwork practice as a critical site where linguistic, historical, and sociocultural data converge. Through discussions of data collection protocols, phonological documentation, lexical analysis, and narrative collection, the study demonstrates how reflexive documentation contributes not only to historical-comparative study but also to meaningful engagement with endangered language communities.

The significance of this article extends beyond the specific languages documented. The Baram–Tinjar region exemplifies the broader challenge facing language documentarians working in Southeast Asian contact: how to identify and describe linguistic boundaries, trace historical relationships, and collect comparable data from speakers whose multilingualism and language shift often make traditional fieldwork categories inadequate. Furthermore, as linguistic diversity in the region continues to decline due to pressure from regional lingua francas and formal education systems, sustained documentation becomes crucial (Dalby, 2003). This article contributes to the methodological literature on language documentation by examining the practical and ethical dimensions of fieldwork in endangered language

settings, while simultaneously advancing knowledge of Baram-Tinjar linguistic structure and historical relationships.

Background

The Baram-Tinjar Linguistic Landscape

The Baram-Tinjar region reaches approximately 5,000 square kilometres of lowland and upland rainforest in Miri Division, Sarawak (Lian, 1987). Linguistically, it represents a convergence point for multiple Austronesian language families. The dominant language families in the region include the Kayan-Kenyah languages (Austronesian, Malayo-Polynesian), the Berawan languages (sometimes classified as Central Sarawak), and various isolated or less well-established languages such as Kiput, Miriek, Bakong, and Tireng. Ethnically, the region is inhabited primarily by Kayan, Kenyah, Berawan, Penan, and smaller populations identifying with languages such as Kiput and Miriek (Blust, 2013).

Previous linguistic documentation of the region began with early colonial-era observations by naturalists and administrators in the late nineteenth century (Brooke, 1866). The most systematic scholarly contributions emerged after the middle of the twentieth century, particularly through the work of Rodney Needham, whose ethnographic studies of Penan and other communities incorporated linguistic observations (Needham, 1965). Subsequent work by Claudette Saguin, Peter Cain, and others expanded the documentation of Kayan, Berawan, and related languages (Saguin, 1980; Cain & Karam, 2010). However, significant gaps remain, particularly regarding phonological detail, systematic lexical comparison, and the documentation of languages with fewer fluent speakers, such as Kiput, Miriek, Narum, and Bakong.

The sociolinguistic context of the region has shifted dramatically since earlier documentation efforts. Formal education, predominantly conducted in Malay and English, has accelerated language shift among younger generations. Many communities that were monolingual or maintained strong L1 transmission fifty years ago now experience partial or functional language loss, with multilingualism and code-switching becoming normative (Hua, 2010). This sociolinguistic reality affects fieldwork methodology and necessitates adaptations in how comparative data are collected and historical relationships are reconstructed.

Historical Linguistic Relationships and Classification

The classification of the Baram-Tinjar languages within the broader Austronesian family remains contested. Blust's (2013) influential classification positions most Baram-Tinjar languages within the Malayo-Polynesian branch, with further subdivisions reflecting both internal linguistic features and geographic presence. However, several languages in the region, particularly Kiput, Miriek, and Bakong, have been difficult to classify, and their precise relationships to other Austronesian groups remain unclear. Some scholars have proposed connections between Kiput and some languages in West Kalimantan, while others have suggested Miriek represents a distinct linguistic innovation centre. These classifications underscore the importance of continued documentation, as improved phonological and lexical data enable more robust historical linguistic reconstruction (Haspelmath, 2013).

Proto-forms and historical phonological patterns are crucial data for linguistic classification. Comparative research across the Kayan-Kenyah languages has identified regular sound correspondences consistent with both horizontal language transmission and vertical inheritance from a proto-language spoken before community dispersal (Blust, 1974). Proto-Kayan-Kenyah reconstructions have been proposed for several hundred lexical items and numerous grammatical features. However, the integration of border languages such as Kiput and Miriek into comparative frameworks has been restricted by limited phonological documentation and the lack of systematic cognate sets. The fieldwork reported in this article partially addresses this limitation by providing previously unavailable phonological and lexical data for comparative analysis.

Literature Review

The literature on linguistic fieldwork, language documentation, and the challenges of working with endangered and minority languages in contact zones creates a substantial body of work that informs the current study. Fieldwork in endangered language settings requires not only technical linguistic competence but also a reflexive awareness of the social, ethical, and methodological conditions under which data are gathered. Recent research has drawn increasing attention to the situatedness of linguistic fieldwork and to how researchers' identities and positions shape the data they collect. Siegl (2021) explores the conditions of "time-bounded fieldwork" conducted outside the speech community, arguing that constraints on immersive fieldwork introduce specific methodological and epistemological challenges that researchers must acknowledge. His study highlights the risks of conducting documentation without extended community presence and the importance of transparency regarding the conditions and limitations of fieldwork contexts. This issue is especially pressing for dying or highly endangered speech communities, where the number of speakers is low, and the variation among those remaining speakers can be significant.

A significant number of published research on the Baram-Tinjar languages has proceeded with limited reflection on the fieldwork experience itself or the social contexts that generate linguistic data. Needham's (1965) ethnographic description of Penan and Belait communities, for instance, reports socio-economic and linguistic observations but offers limited discussion of how the researcher's own position, relationships with informants, or the conditions of data collection may have shaped what was recorded. Similarly, Saguin's (1980) and Cain & Karam's (2010) documentation of Kayan and Berawan language structure presents linguistic data and analysis with minimal methodological discussion of how that data was elicited, from whom, or under what fieldwork conditions, leaving readers and researchers unable to assess the reliability or representativeness of the reported research.

A growing body of work treats the researcher's positionality not as background context but as a determinant of data quality. Everett and Everett (2016) and Siegl (2021) both argue that the nature of fieldwork, whether it involves extended immersion in a community or only short, limited visits, affects the epistemological boundaries of the data collected. However, they differ in their focus; Everett and Everett (2016) approach this primarily as a philosophical issue regarding what fieldwork can truly understand. In contrast, Siegl (2021) views it as a practical

limitation, emphasising that researchers must clearly communicate what aspects their fieldwork design cannot capture. Siegl (2018) extends this further into the specific case of moribund speech communities, arguing that when only a handful of elderly speakers remain, the researcher's role becomes simultaneously linguistic, cultural, and relational, a role that standard elicitation-based methodology does not account for. These sources demonstrate that reflexivity is not just a desirable ethical practice; it is a necessary methodological requirement. Without it, as seen in much of the earlier Baram-Tinjar literature discussed above, readers cannot adequately assess the reliability of the presented data.

Kroeger (1998) demonstrates that many of Sarawak's minority languages occupy contested classificatory positions because the comparative data available are too limited to resolve genetic affiliation, a finding consistent with Blust's (2013) broader observation that several Baram-Tinjar languages, particularly Kiput, Miriek, and Bakong, have resisted certain subgrouping within Malayo-Polynesian. Dhakal (2014), working on a geographically distant minority language, shows that classificatory and structural questions cannot be separated from the sociolinguistic ecology in which a language survives: patterns of multilingualism, intergenerational transmission, and community attitude directly affect what comparative data can be elicited. These sources indicate that the classificatory uncertainty surrounding Baram-Tinjar languages is not simply a data-collection issue to be solved by more fieldwork, but a structural feature of working with small, multilingual, rapidly shifting speech communities.

Grenoble et al. (2008), Akumbu (2019), and Nathan (2012) collectively reframe the linguist's role in endangered-language settings as extending well beyond data collection. Grenoble et al. (2008) argue that documentation disconnected from training and materials development fails the broader goal of language maintenance; Akumbu (2019) provides an empirical case for this state, showing that community co-participation in grammar-writing improves both descriptive accuracy and the probability that documentation will actually be used; and Nathan (2012) extends the argument to archiving practices, challenging that archive design oriented towards future academia rather than present-day speaker communities repeats the same extractive logic that community-based approaches aim to improve. Nathan (2012) highlights that the consistency across the three independent factors of literature methodology, practice, and infrastructure indicates that community-oriented documentation should be viewed as a coherent paradigm shift in the field rather than as three separate recommendations. This is the framework that the commitments to data sharing and archiving in the present study are designed to support.

Penfield (2013) and Himmelmann (1998, 2006) agree that linguistic structure cannot be fully understood in isolation from the social and cultural context that produces it, but they differ in scope. Himmelmann's (1998, 2006) influential distinction between documentary and descriptive linguistics treats context primarily as a resource for richer corpora, whereas Penfield (2013) advances further in arguing that ethnographic method changes *what counts as data* in the first place, community relationships, use-context, and cultural meaning become part of the linguistic record rather than background to it. This distinction is important for the current study because it clarifies that the ethnographic aspect of the fieldwork presented here is not

just an additional detail. Instead, it is considered, following Penfield's (2013) stronger argument, to be essential to the phonological and lexical data itself. Generally, the literature reviewed establishes a robust scholarly framework for understanding the methodological, ethical, and applied dimensions of language documentation in endangered minority language settings.

METHOD

This study employed an ethnographic approach to language documentation, combining traditional comparative linguistic methods with reflexive fieldwork practice grounded in community-based research principles. Reflexivity has become an important expectation in ethnographic linguistics (Everett & Everett, 2016; Himmelmann, 1998). However, it is usually implemented through disclosure, where researchers describe their positionality and the conditions of their fieldwork as contextual information, separate from the linguistic analysis itself. This study departs from that convention by treating fieldwork conditions specifically, the stage of the researcher-informant relationship at which data were collected, as a variable that is tracked alongside the linguistic data and how that data is exerted in ensuing comparative analysis. Data collected during early-stage encounters with lower trust levels were identified as more likely to reflect formal or elicitation-oriented speech. In contrast, data from later-stage encounters, once informants became comfortable with the recording equipment and the researcher's presence, were considered more representative of natural speech. As a result, this latter data was given greater importance in phonological and lexical comparisons. In this sense, reflexivity functions here as a methodological control on data reliability, not only as an ethical or narrative framing device, which is the specific respect in which this study extends, rather than simply restates, established ethnographic-linguistic practice.

The research design prioritised sustained engagement with language communities, enabling the development of relationships with language informants and a deepened understanding of sociolinguistic contexts shaping language use and transmission. Ethical clearance was obtained from the institutional review board at Universiti Brunei Darussalam and Curtin University, Malaysia, before commencing fieldwork. Language informants participated voluntarily and received a token of appreciation for their time. Documentation of all work, including recordings and transcripts, was conducted with the explicit consent of informants for archival preservation, academic use, and community access. Language informants were identified through community networks and recommendations from local leaders. Primary informants were selected based on several criteria: native fluency in the target language, willingness to participate in sustained documentation activities, and ability to communicate about linguistic structures in the regional lingua franca (Malay or English). Community selection prioritised sites where language was still used for intergenerational transmission, where local leaders endorsed the research, and where accessibility and security permitted sustained engagement. This resulted in a focus on main settlements rather than dispersed household clusters, which introduced certain limitations: speakers consulted may have been somewhat more exposed to regional lingua francas than some community members. However, this trade-off enabled in-

depth work with communities and facilitated compliance with safety protocols necessary for foreign researchers working in remote areas.

Phonological data were collected through a combination of elicited word lists, minimal pair identification tasks, and natural speech recordings. Initial sessions employed Swadesh 200-word lists (Swadesh, 1971) in modified form, supplemented by topically organised vocabulary sets targeting domains such as kinship, flora, fauna, agriculture, and material culture. For each target lexeme, multiple elicitations from different informants were recorded to document speaker variation. Phonetic transcription was conducted using the International Phonetic Alphabet, following standard linguistic practice for documenting endangered languages (Maddieson, 2013).

Lexical documentation encompassed systematic word lists organised by semantic domains. Elicitation was conducted using visual stimuli (photographs and drawings) and conceptual frameworks, yielding definitions and usage examples. When informants' metalinguistic awareness and bilingual competence permitted, explicit discussion of etymological relationships, borrowing history, and cognate relationships across languages was documented. This proved particularly valuable for identifying borrowing patterns across languages and for recognising shared cognates that potentially indicate historical relationships (Dryer & Haspelmath, 2013).

Narrative data were collected through ethnographic interviews, personal stories, oral histories, and traditional narratives. Informants were invited to share stories, which were recorded with permission for preservation and academic use. Narrative collection served multiple functions: it provided data on connected speech and natural phonological variation beyond word-list contexts; it generated broader linguistic data, including morphosyntactic structures; and it contributed to cultural documentation reflecting community knowledge and values. Narrative transcription and analysis involved consultant collaboration, with linguistic transcriptions verified through playback and explanation by original speakers.

RESULTS AND DISCUSSION

Methodological Insights from Reflexive Fieldwork

This fieldwork experience generated several methodological insights regarding language documentation in endangered, multilingual contact zones. First, sustained, context-sensitive engagement with communities fundamentally shapes the quality and comprehensiveness of linguistic data obtained. Initial field sessions frequently yielded data reflecting the consultant's awareness of the researcher's interests and the conventions of elicitation-based documentation. As relationships deepened and informants became comfortable with recording equipment and analytical approaches, data increasingly reflected natural language use patterns. This progression suggests that proper language documentation in endangered contexts requires investment of time substantially greater than often assumed in grant proposals and project timelines (Himmelman, 1998).

This observation is consistent with broader theoretical discussions of reflexive fieldwork methodology. Siegl (2021) has emphasised that temporal constraints on fieldwork apply specific epistemological limitations in the data gathered, particularly when the researcher cannot be embedded within the speech community for extended

periods. In the Baram-Tinjar context, where communities are geographically dispersed and accessible only during specific periods, fieldwork visits across multiple seasons proved essential for deepening informant relationships and obtaining increasingly naturalistic data. Early sessions were characterised by a degree of formal register on the part of informants, who oriented their speech toward what they perceived as the researcher's expectations. Over time, as informants became familiar with the fieldwork process, they increasingly produced data that reflected everyday rather than formal linguistic registers, a shift of considerable importance for phonological and lexical documentation. Siegl's (2018) reflections on fieldwork within endangered speech communities further highlight that researchers must navigate the dual challenge of eliciting comprehensive data from speakers with limited remaining fluency while maintaining respectful, non-exploitative relationships that acknowledge the emotional and cultural weight of documenting an endangered language. Both dimensions of this challenge were encountered throughout the Baram-Tinjar fieldwork, reinforcing the case for reflexive, relationship-centred approaches to documentation.

Second, the multilingualism of informants, rather than representing a methodological problem to be minimised, offers opportunities for comparative work if properly managed. Many informants used Malay, English, or neighbouring languages fluently or extensively, creating the potential for confusion about language boundaries. However, explicit meta-linguistic discussion of language boundaries, encouraging informants to maintain focus on target languages, and comparing multiple informants' productions of identical items proved effective. Furthermore, informants' multilingual competence enabled them to recognise cognates across languages and to discuss etymology and borrowing patterns, contributing substantially to comparative analysis. The documentation of code-switching and language contact phenomena itself provided valuable sociolinguistic and historical information (Haspelmath, 2013).

Third, flexibility in field methods proved essential. Schedules had to adapt to seasonal variations in community activities, agricultural calendars, and social events. Certain narrative data emerged naturally during social gatherings rather than formal elicitation sessions. Phonological data collection methods required adjustment when informants found particular elicitation tasks difficult or uncomfortable. Rather than rigidly adhering to predetermined protocols, effective fieldwork involved responsiveness to community rhythms and consultant preferences while maintaining analytical focus. This flexibility did not compromise data quality; instead, it often improved both the quality of data obtained and the ethical dimensions of research engagement (Everett & Everett, 2016).

Historical Linguistic Relationships and Implications

The linguistic data documented through this fieldwork advance understanding of the historical relationships among the Baram-Tinjar languages and their place within broader Austronesian diversification. Evidence from phonological correspondences, cognate sets, and morphosyntactic features consistently supports a primary division between Kayan-Kenyah languages and Berawan, with secondary clustering of Kiput and Miriek closer to Kayan-Kenyah than to Berawan. Bakong and Tireng exhibit intermediate characteristics; a preliminary analysis suggests a possible

closer relationship to the Kayan-Kenyah languages, though limited data from these languages warrant cautious interpretation. The geographic distribution of contemporary languages reflects complex settlement patterns and possibly post-settlement secondary dispersal and linguistic divergence (Blust, 2013).

These findings carry important implications for the broader project of Austronesian historical linguistics in Borneo. Kroeger's (1998) foundational work on language classification in Sarawak established that many of the region's minority languages occupy contested classificatory positions, with insufficient data to resolve their genetic affiliations with confidence. The phonological and lexical evidence gathered in the present study partially addresses this gap for several Baram-Tinjar languages, providing new data points for comparative analysis. Specifically, the regular sound correspondences identified between Kiput and languages within the Kayan-Kenyah grouping support a closer genetic affiliation than earlier research based on limited data had proposed. The preliminary proto-form reconstructions, while incomplete, establish a comparative baseline that future researchers can refine as additional data become available. It is important to highlight, as Kroeger (1998) also points out, that classification decisions in linguistically complex regions must remain provisional when based on restricted comparative datasets. The data reported here should therefore be understood as a contribution to an ongoing classificatory conversation rather than a definitive resolution of the historical relationships among these languages.

Proto-form reconstruction remains incomplete, constrained by available cognate data and by the challenge of establishing regular sound correspondences for peripheral languages represented by fewer fluent speakers and more limited comparative data. However, the 127 proto-items confidently reconstructed provide a foundation for further comparative work and for exploring wider Austronesian relationships. Some reconstructed proto-forms show patterns consistent with proposed connections to other Austronesian language groups in Borneo and peninsular Malaysia, suggesting that future reconstruction efforts integrating additional comparative data could illuminate broader Southeast Asian language prehistory.

Language Documentation and Minority Language Protection

Beyond its contributions to comparative linguistics, this documentation work is directly relevant to minority language protection and revitalisation efforts. Several languages represented in this study have fewer than 500 fluent speakers remaining, with limited intergenerational transmission and accelerating language shift toward Malay (Lewis et al., 2016). Language documentation creates permanent records to counter total language loss and provides essential resources that communities can utilise in revitalisation or maintenance efforts. Several communities expressed interest in using documentation data for school-based teaching materials and for cultural transmission to younger generations. The establishment of archive copies in community-accessible locations, rather than in institutions, is an important practice for ensuring that communities retain access to documentation of their own linguistic heritage.

Grenoble et al. (2008) argue that linguists have the responsibility to ensure that documentation outputs are not deposited in academic archives but are actively

transferred to communities in usable forms, accompanied by training that enables community members to conduct further documentation and to use existing records for revitalisation. The findings of this study advance this responsibility by involving community informants as collaborative partners rather than passive data sources and by committing to producing community-accessible outputs alongside academic publications. Nathan's (2012) work on language archives further reinforces this approach, arguing that archiving practices must be reimagined with the language-speaking community as the primary intended beneficiary, rather than the future academic researcher. For the Baram-Tinjar communities documented in this study, the establishment of accessible local archives represents an especially urgent priority, given that the institutional infrastructure for language maintenance in this region remains underdeveloped compared to other parts of Malaysia.

Language documentation also provides baseline data against which language change can be measured. As younger generations increasingly shift toward Malay, future documentation efforts will reveal how linguistic systems transform under conditions of language shift (Zulkiflee & Chuchu, 2025a). Comparative data from this study, when contrasted with future documentation, will enable analysis of directionality and mechanisms of language change in community settings. This historical perspective on language change contributes valuable data to theories of language acquisition, multilingual negotiation, and language contact processes (Thomason, 2001).

Challenges and Gaps

Despite the achievements of this fieldwork, significant challenges and research gaps remain. Informants with deepest fluency and knowledge of traditional linguistic forms were often elderly, with limited availability for extended documentation sessions. While documenting elderly speakers helps preserve traditional language forms, it may underrepresent innovations and variations characteristic of younger speakers. Intentional documentation efforts focusing on younger speakers with varying proficiency levels would complement this work and provide a more complete picture of contemporary language systems across age groups. Similarly, documentation with geographic variations within language communities and across socioeconomic and educational backgrounds would strengthen the comparative base (Himmelman, 2006).

The challenge of speaker availability is compounded by broader structural constraints that shape fieldwork in remote areas of northern Sarawak. Transportation infrastructure in the Baram-Tinjar region remains limited, and community accessibility varies significantly by season, restricting the frequency and duration of fieldwork visits. These logistical realities are not unique to the present study; Siegl (2021) has documented comparable constraints in other remote language documentation projects and argued that researchers must develop methodologies suited to the conditions of restricted access rather than assuming ideal fieldwork circumstances. In the Baram-Tinjar context, the findings revealed a combination of geographic remoteness, limited speaker numbers, and the advanced age of the most fluent informants, creating a particularly urgent documentation emergency. The opportunity for comprehensive phonological, lexical, and narrative documentation is

narrowing rapidly, and the gaps that remain in the present study represent scholarly priorities for future fieldwork initiatives.

Linguistic domains are not thoroughly documented in this fieldwork, including systematic grammatical description and discourse-level features of language organisation. The documented phonological systems are primarily segmental inventories and basic features; a more detailed description of allophonic variation, phonotactic constraints, and phonological processes in connected speech would strengthen the phonological analysis. The project's scope and focus limited grammatical documentation; a comprehensive description of syntax, morphology, and semantic structure would require substantial additional fieldwork.

Implications

This study generates a set of interrelated implications that extends beyond descriptive documentation and contributes to the theoretical, methodological, and applied domains of linguistic research in the Baram–Tinjar region. At the linguistic level, the documentation of phonological, lexical, and narrative data provides a robust empirical foundation for refining comparative Austronesian analysis, particularly for under-described languages such as Bakong, Kiput, Narum, and Tireng (Omar & Norahim, 2020). The findings demonstrate the continued relevance of the comparative method in reconstructing proto-forms and clarifying subgrouping relationships, while also highlighting the need to incorporate real-time language shift dynamics into historical-linguistic modelling. As communities increasingly aim towards Malay and other dominant languages, processes of contact-induced change and levelling emerge as critical areas for future investigation.

Methodologically, the research underscores the importance of reflexive and context-sensitive fieldwork. Field conditions in the Baram–Tinjar region, characterised by environmental constraints, logistical challenges, and multilingual social settings, necessitate adaptive strategies in data collection, verification, and researcher–community interaction (Damit & Zulkiflee, 2025). The study, therefore, contributes to language documentation practice by demonstrating that rigorous data collection must be complemented by ethical awareness, flexibility, and sustained engagement with speakers. This reinforces the argument that documentation is not a technical exercise but a socially embedded process shaped by trust, accessibility, and reciprocity. From an applied perspective, the documentation produced through this research holds significant potential for educational and revitalisation initiatives. The formation of structured lexical datasets, phonological descriptions, and narrative materials provides a foundation for developing pedagogical resources that can be integrated into community-based learning, school curricula, and higher education contexts (Chowdhury & Alzarrad, 2025). Such materials offer authentic, community-verified content that supports intergenerational transmission and strengthens linguistic identity. In this context, the study aligns with broader frameworks that advocate for documentation outputs that are simultaneously academically rigorous and socially meaningful.

The reflective dimension of this research further deepens its implications by situating language endangerment within lived socio-cultural realities. Field observations reveal that language shift is driven by both internal and external pressures, including reduced motivation among younger speakers, the perceived

economic value of dominant languages, increased mobility, and the influence of formal education systems. These factors operate within a broader ecology of contact, where smaller language communities face displacement pressures from larger regional groups. Consequently, language endangerment must be understood not only as a linguistic phenomenon but as part of wider processes of social transformation (Costa, 2013).

This reflexive engagement also highlights the ethical responsibilities of the researcher. Fieldwork experiences demonstrate that linguistic research carries responsibility towards the communities whose knowledge systems are being documented. As such, future work must prioritise collaborative approaches that involve native speakers, local educators, and community stakeholders as active participants rather than passive informants. Building local capacity in documentation and analysis is essential to ensuring that linguistic knowledge remains accessible, sustainable, and community-controlled. Looking forward, several strategic directions emerge. First, expanded comparative documentation across the Baram–Tinjar region will enable more precise reconstruction of historical relationships and help resolve classification uncertainties. Second, the development of community-accessible teaching materials such as primers, lexicons, and oral history archives should be prioritised to support revitalisation efforts. Third, institutional integration of minority language documentation into university curricula and training programmes can create pathways for sustained research and community engagement. Finally, the establishment of durable, locally accessible archival infrastructures is critical for documenting linguistic data for both academic and community use (Zulkiflee & Chuchu, 2025b). Fundamentally, these efforts should be theorised as temporally expansive practices that transcend the boundedness of research timelines and require enduring, ethically grounded engagements with community relationships (Zulkiflee, 2026).

Overall, the implications of this study converge on a model of linguistic research that is empirically grounded, methodologically adaptive, and socially responsive. By integrating documentation, reflection, and community-oriented practice, the research contributes to a more comprehensive understanding of language endangerment. It provides a foundation for future work that bridges the gap between theoretical inquiry and real-world impact.

CONCLUSION

This article presents comprehensive documentation of linguistic fieldwork conducted among the minority languages of the Baram–Tinjar region of northern Sarawak. Through engagement grounded in reflexive methodology, ethnographic sensitivity, and collaborative community relationships, the research generated detailed phonological, lexical, and narrative data contributing to comparative Austronesian linguistics and to minority language documentation. The work demonstrates that sustained, context-sensitive fieldwork remains essential for documenting linguistic diversity in endangered language settings. As languages throughout Southeast Asia face pressure from lingua francas and language shift accelerates, continued investment in rigorous documentation becomes increasingly crucial. This research contributes both to specific knowledge of the Baram–Tinjar

languages and to broader methodological discussions of language documentation practice. For the communities whose languages are documented here, this work represents recognition of linguistic value and cultural heritage, and it provides archival resources that may support future language maintenance and revitalisation efforts.

Funding

The author declared the study received no external funding.

Acknowledgment

Not applicable.

Conflict of interests

The authors declare no conflict of interest.

REFERENCES

- Akumbu, P. W. (2019, March). Reflections on a community-based approach to writing grammars of endangered languages. In *conference Descriptive Grammars and Typology, University of Helsinki* (pp. 27-29).
- Blust, R. (1974). The Proto-North Sarawak vowel system and its reflexes. *Oceanic Linguistics*, 13(1-2), 1-37.
- Blust, R. (2013). *The Austronesian languages* (Revised edition). Australian National University.
- Brooke, J. (1866). *Ten years of Sarawak*. John Murray.
- Cain, B., & Karam, J. M. (2010). Documentation of the Kayan language: Field reports and linguistic data. *American Linguistic Association Proceedings*, 28(2), 156-178.
- Chowdhury, S., & Alzarrad, A. (2025). Advancing community-based education: Strategies, challenges, and future directions for scaling impact in higher education. *Trends in Higher Education*, 4(2), 21.
- Costa, J. (2013). Language endangerment and revitalisation as elements of regimes of truth: Shifting terminology to shift perspective. *Journal of Multilingual and Multicultural Development*, 34(4), 317-331.
- 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.
- Dalby, D. (2003). *The register of the world's languages and language families*. Linguasphere Press.
- Dhaka, D. (2014). A sociolinguistic study of the Baram language. *Himalayan Linguistics*.
- Dryer, M. S., & Haspelmath, M. (Eds.). (2013). *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology.
- Everett, D. L., & Everett, K. (2016). Philosophical and methodological foundations of linguistic fieldwork. *Journal of Language Contact*, 9(1), 23-47.

- Grenoble, L. A., Rice, K. D., & Richards, N. (2008). The Role of the Linguist in Language Maintenance and Revitalization: Documentation, Training and Materials Development. *Language and Poverty*, 183-201.
- Haspelmath, M. (2013). On system-dependent occurrence of grammatical forms. *Language*, 89(3), 554-569.
- Haspelmath, M., & Sims, A. D. (2010). The world atlas of language structures plus. Max Planck Institute for Evolutionary Anthropology.
- Himmelman, N. P. (1998). Documentary and descriptive linguistics. *Linguistics*, 36(1), 161-195.
- Himmelman, N. P. (2006). Language documentation: What is it and what is it good for? In J. Gippert, N. P. Himmelman, & U. Mosel (Eds.), *Essentials of language documentation* (pp. 1-30). Mouton de Gruyter.
- Hua, Z. (2010). Language policies and multilingualism in China. *Journal of Multilingual and Multicultural Development*, 31(1), 77-92.
- Kroeger, P. R. (1998). Language classification in Sarawak: a status report. *The Sarawak Museum Journal*, 53(74), 137-173.
- Lewis, M. P., Simons, G. F., & Fenning, C. D. (Eds.). (2016). *Ethnologue: Languages of the world* (19th ed.). SIL International.
- Lian, F. J. (1987). *Farmers' perceptions and economic change-the case of Kenyah farmers of the fourth division, Sarawak*. The Australian National University (Australia).
- Maddieson, I. (2013). *Patterns of sounds* (Second ed.). Cambridge University Press.
- Nathan, D. J. (2012). From Preserving Language Resources to Serving Language Speakers: New Prospects for Endangered Languages Archives. In *ADHO 2012-Hamburg*.
- Needham, R. (1965). Penan and Belait: Socio-economic aspects of some Bornean hunter-gatherers. *Journal of the Anthropological Institute*, 95(2), 313-347.
- Omar, A. H., & Norahim, N. The GLOCAL Proceedings Asia 2020.
- Penfield, S. D. (2013). Ethnographic contributions to the study of endangered languages. Ed. by Tania Granadillo and Heidi A. Orcutt-Gachiri. Tucson: University of Arizona Press, 2011. Pp. 248. ISBN 9780816526994. \$55 (Hb). *Language*, 89(2), 363-366.
- Saguin, C. (1980). *Language structure and cultural organization in Borneo*. Oxford University Press.
- Siegl, F. (2018). Reflections on linguistic fieldwork within moribund speech communities. *Études finno-ougriennes* (49-50).
- Siegl, F. (2021). Time-bounded Fieldwork “Ex-situ” or On Fieldwork Without Unlearning Another Language. *Studie z aplikované lingvistiky*, 12(1), 53-80.
- Swadesh, M. (1971). *The origin and diversification of language*. Aldine-Atherton.
- Thomason, S. G. (2001). *Language contact: An introduction*. Georgetown University Press.
- Woodbury, A. C. (2011). Language documentation. In R. Mesthrie (Ed.), *The Cambridge handbook of sociolinguistics* (pp. 159-186). Cambridge University Press.
- Whaley, L. (2003). The future of native languages. *Futures*, 35(9), 961-973.

- Zulkiflee, Z., & Chuchu, I. F. (2025a). 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.
- Zulkiflee, Z., & Chuchu, S. I. F. H. (2025b). Belait language maintenance in Brunei Darussalam. *International Journal of Linguistics and Literature*, 3(1), 77-92.
- Zulkiflee, Z. (2026). From Riverbanks to Academia: An Autoethnographic Journey Through Brunei's Endangered Belait Language. In *International Student Narratives in Globalized Higher Education* (pp. 337-364). IGI Global Scientific Publishing.