

Research Article

Cite this article: Pilar, J. G. A. (2026). Technology Role in Lexicon Preservation and Developing Teaching Materials in MTB-MLE of the Kawayan Dialect in Southern Negros, Philippines. *Educational Process: International Journal*, 24, e2026096.
<https://doi.org/10.22521/edupij.2026.24.96>

Received September 13, 2025

Accepted November 11, 2025

Keywords: Kawayan lexicon, MTB-MLE, pedagogy, SDG 10, technology

Author for correspondence:

John Gerald A. Pilar

✉ john.pilar@chmsu.edu.ph

✉ Carlos Hilado Memorial State University, Philippines

Technology Role in Lexicon Preservation and Developing Teaching Materials in MTB-MLE of the Kawayan Dialect in Southern Negros, Philippines

John Gerald A. Pilar 

Abstract

Background/purpose. This study presents a technology-mediated pedagogical model for endangered-language education that aligns directly with SDG 10 by reducing educational inequality. This research integrates digital archival science and computational linguistics to create an empirically validated digital Kawayan lexicon. The resource serves as the foundational data source for instructional design and curriculum development, enabling the creation of authentic teaching materials rigorously aligned with the MTB-MLE Framework. Hence, this study demonstrates a scalable, data-driven approach to sustaining educational transmission for future generations.

Materials/Methods. Using a qualitative textual analysis approach and digital tools (e.g., pivot tables), lexical data from face-to-face interviews with nine native speakers from Southern Negros Occidental, Philippines, were analyzed. Participants were purposively selected according to ethical standards. The word retrieval process was timestamp-coded input. Then, inter-coders ensured the dataset's validity and reliability, creating an organized and understandable lexicon for analysis.

Results. This study highlights essential findings for establishing the Kawayan lexicon's foundations, which define its form and meaning. It also validates and demonstrates the Kawayan lexicon for linguistic categorization, serving as a pedagogical resource for MTB-MLE alignment. This analysis examines specific linguistic features, such as affixation and reduplication, while identifying crucial documentation gaps that will inform future archival and instructional design development.

Conclusion. In line with SDG 10, the study concludes by validating a technology-mediated pedagogical model for revitalizing the Kawayan dialect. The integrated digital lexicon serves as the empirical base for designing authentic MTB-MLE instructional materials. With this approach, the Kawayan dialect is successfully transformed into a dynamic medium for educational transmission. This study allows replication by future researchers.



OPEN ACCESS

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

1. Introduction

A pedagogical recalibration is required that incorporates language and dialect protocols and uses data archiving to document and preserve endangered languages, given the growing influence of the digital lexicon (Adamou et al., 2025; Ajani et al., 2024; Chaika et al., 2024). Ultimately, this approach fosters comprehensive language proficiency by aligning domain-specific vocabulary and instructional language education with the objectives of Mother Tongue-Based Multilingual Education (MTB-MLE) (Sarma, 2025; Tenorio, 2024). It examines its possibilities, challenges, and ways to prevent potential issues. Thus, this involves training diverse individuals with varying levels of knowledge and expertise (Ancheta, 2024; Dobrin & Berson, 1999).

Language is considered an important dimension of social and economic development, and it preceded it in the United Nations (UN) planning process (Taylor-Leech & Benson, 2017). However, immersion learning and teaching are made very challenging by unsuitable curricula and inadequate school facilities, especially when the teacher is also learning a foreign language (Brock-Utne, 2012). Mother tongue-based bilingual programs utilize the learner's first language, known as L1, to teach beginning reading and writing skills alongside academic content (Benson, 2004, cited in Da Silva et al., 2024).

Still, the UN's state-centered and logical focus overlooks the role of language in agency, capability, and reducing inequality (SDG10). It argues that mainstream Language Policy and Planning (LPP) limits the effectiveness of the Sustainable Development Goals (SDGs). (Norhana et al., 2024). There should be more emphasis on language-based solutions for sustainability, which should receive more attention, because actual sustainable development, characterized by long-term economic growth and human well-being, depends on community involvement and careful consideration of language and literacy, which are still important but often disregarded (Fettes, 2023; Trudell, 2009).

The development of the lexicon in this research serves as an initial exploration and documentation to enhance the Kawayan vocabulary. This Kawayan lexicon can support native speakers' day-to-day conversations across various scenarios. This case underscores the significance of the lexicon in facilitating effective vocabulary development and providing sufficient vocabulary resources for language learning, particularly in high-frequency contexts (Castellano-Risco, 2024). This confirms that it should conduct high-frequency word lists to check various versions using different counting methods. It can then rank the list items based on these counting methods (Yousefi & Biria, 2018; Dang et al., 2022).

This can be achieved through innovative methodologies for documenting emerging languages, characterized by sophisticated technological capabilities for recording, transcribing, annotating, managing, and analyzing language records within the pedagogical model of language education (Bird & Simons, 2003; Chun et al., 2016). In addition to supporting conservation efforts, this development may also lead to spontaneous phonetic convergence, influencing memory, speech, perception, social cognition, and language change by documenting complex linguistic, musical, and cultural features for heritage preservation and revitalization. This would provide educational sciences with important insights into language learning, communication, and cognitive development across various classroom settings (Coles-Harris, 2017; Grant, 2012).

In the Philippines, Indigenous languages outnumber non-indigenous ones, with many linguistic regions still needing to be explored. It is mentioned that there is a lack of a straightforward method to distinguish between languages and dialects (Radford, 2016), while such differentiation often depends on mutual intelligibility, cultural perceptions, and political factors, which significantly correlate with one another and can be predicted from objective measures like lexical similarity and phonological correspondence equally well (Lewis, 2024). The terms "language" and "dialect" are deeply rooted in our understanding. However, linguists generally agree that a clear distinction cannot

be made to categorize two speech varieties as separate or two dialects of the same language (Tang & Van Heuven, 2009; Wichmann, 2020). The Kawayan lexicon, on the other hand, is an essential linguistic and educational tool that enhances regional curricula, fosters inclusive, literacy-focused pedagogy, and promotes cultural identity in mother-tongue-based multilingual education, such as the development of more localized learning materials (Barruga, 2025; Ibno, 2025).

In this comparative study of eleven Bikol-Sorsogon language varieties in the Philippines, a phono-lexical approach with a 200-item basic wordlist was used to calculate lexical and phonetic similarities. In this context, phono-lexicostatistics is a systematic and helpful tool for examining language relationships in under-documented areas, such as the Bikol-Sorsogon varieties. Due to the lack of existing linguistic literature on these varieties, this research is an important starting point for further studies (De La Torre & Gonong, 2020). This growth is crucial for compiling a comprehensive list that accurately represents the cultural and linguistic diversity of the Philippines and its neighboring regions, such as Borneo and Celebes. Such efforts aim to facilitate comparative studies and reconstruct Philippine cultural history (Zorc, 1974). Digital advancements now provide linguists with easier access to recordings, enabling more precise analyses and the responsible repatriation of materials (Coles-Harris, 2017). In this study, the Kawayan lexicon has been systematically determined and examined using digital tools to enhance and promote vocabulary in the target language, thereby contributing to language literacy development among learners (Bocanegra et al., 2025).

Particularly, this study focuses on the Kawayan lexicon as a lexicostatistical database of the Kawayan dialect, collecting words and information related to the wordlist (Coxhead et al., 2019) used at a particular time, as well as the glosses assigned and accepted within the speech community and in academic language (Csomay & Prades, 2018). It also focuses on categorizing words from the Kawayan dialect spoken by locals. This aims to tackle language preservation and documentation challenges, especially for endangered dialects. These words will serve as teaching materials in MTB-MLE and help preserve local languages (Bocanegra et al., 2025). It highlights the crucial role of government, schools, and community support in preserving these languages. Although local languages have been incorporated into the curriculum, significant challenges remain. These include the need for teacher training, a shortage of high-quality teaching materials, and a lack of strong policy support (Subiantoro et al., 2023).

The research also highlights the impact of globalization and technology on language development. This corresponds with UNESCO's goals for educational sustainability (Crystal, 2001). This investigation aims to document the dialect interactions in Southern Negros. It explores the historical, community, and social aspects of linguistic diversity. This project aims to support language revival and enhance local language education (Atifnigar et al., 2025) by creating a lexicostatistical database for the Kawayan dialect, which will serve as a resource and teaching material for teaching the mother tongue (Nwachukwu & Joseph, 2024). This effort aims to support language revival and enhance local language education by developing a lexicostatistical database for the Kawayan dialect (Frederick et al., 2025). Indeed, the local dialect will serve as an additional means of communication. It even allows for the use of another language in teaching settings. It also demonstrates how educators encourage flexible language teaching, departing from the rigid structures of dual language bilingual classrooms and moving beyond the traditional view of separate languages and literacies (García & Wei, 2014).

Moreover, with the introduction of the MTB-MLE policy, which the government requires to improve language and literacy skills among children, some obstacles arise (Adriano et al., 2021). These include the lack of easily accessible print resources for MTB-MLE and the challenge of translating specialized terms into local languages (Ponce & Lucas, 2024). It also underscores the importance of involving local stakeholders in policy planning to bridge this gap, emphasizing that their experiences and wisdom are crucial to addressing the advantages of MTB-MLE and the challenges

hindering its full realization (Cabansag, 2016). On the other hand, the benefits identified include the opportunity to strengthen cultural ties, enhance understanding through instruction in the native language, the crucial involvement of parents in the MTB-MLE initiative, and the creation of more culturally relevant educational materials (Tan & Jehom, 2024). It highlights how mother tongues in education remain vulnerable, despite official language and policies seeming to support them (Tupas, 2015).

Building on previous studies, the primary goal of this study is to examine the Kawayan dialect's lexicon using code glosses and the alternative words that accompany them, with a particular focus on how dialect chaining influences glosses. This analysis aims to enable teachers and schools to choose the most suitable program designs and pedagogical approaches for their specific educational contexts. In particular, using a more grassroots approach, this study explores and documents the lexicon of the Kawayan dialect to create instructional materials for Mother Tongue-Based Multilingual Education.

2. Literature Review

2.1. The Role of Technology in Developing Local Lexicons as Educational Resources

This technological dialect documentation and preservation entails a set of actions that are more or less commonly recognized, such as recording, generating metadata, annotating media files, backing up files, and archiving materials for safekeeping (Cox, 2022; Schultze-Berndt, 2015). It presents the habits for tracking metadata (Thieberger, 2016) and archiving resources, such as word processing software for generating and tracing metadata, spreadsheets, or pen and paper (Thieberger, 2020). In this case, language development, such as the Kawayan dialect, involves studying how linguistic features vary analytically among distinct speaker groups or within the same speaker in different contexts (De La Torre & Gonong, 2020). Notably, it emphasizes that developing lexical competence, which encompasses vocabulary growth, word formation, and semantic understanding, is crucial in enhancing communication and professional skills (Ancheta, Jr. & Ochoco, 2024), particularly through innovative, technology-driven educational practices (Zhang, 2024). As education increasingly shifts toward digital and distance learning, integrating lexical competence training into these modern platforms is essential for maintaining linguistic proficiency and meeting contemporary educational needs (Lobanova et al., 2023).

The importance of the teacher as the direct channel between the curriculum and the student is emphasized in training proficiency in mother tongue instruction, particularly in multilingual education programs that prioritize mother tongue instruction. Nevertheless, it found the teacher identification process challenging due to a lack of professionally qualified teachers who also shared the learners' mother tongue (Kosonen, 2017, cited in Herzog-Punzenberger et al., 2017).

Upon closer examination, the interplay between technology and lexicon in Visayan languages highlights how globalization and media exposure drive lexical innovation and adaptation. It is indeed the morphologically integrated, often with Visayan affixation patterns, to fit native grammatical structures (e.g., modifying verbs or adjectives). The technology-related vocabulary has extensively expanded, which reflects not only the influence of colonial and post-colonial linguistic contact (Erzsébet et al., 2024) but also the growing need to express technological concepts that may lack an indigenous counterpart by integrating affixes and code-mixing strategies frequently appearing in intrasentential formations to enrich Visayan expression while also serving pedagogical and communicative functions (Dobrin & Berson, 1999).

This blend enhances communicative efficiency and cultural continuity for under-documented dialects, such as the Kawayan dialect. Incorporating technology-related vocabulary into lexicographic efforts could support linguistic revival and teaching (e.g., through Mother Tongue-Based Multilingual

Education, MTB-MLE) and ensure the dialect evolves to meet contemporary needs (Heraña & Azarias, 2024). This prioritizes eliciting modern term usage among speakers, especially for digital tools, social media phenomena, or everyday consumer technology (Shaghghi et al., 2021), to enrich the dialect corpus and maintain its relevance (Panaguiton & Mantawil, 2024).

2.2. Lexicon Preservation and Documentation

The subfield of computational linguistics related to technology has emerged with the expected availability of affordable equipment and software for generating, storing, and producing digital texts specifically for preserving the underdocumented Kawayan dialect (Erzsébet et al., 2024). The current technologies used both for providing documentary materials, usually recognized to contain recorded language activities in various categories and settings, and enough annotation to make them understandable, and at the same time, enabling the materials to be utilized and produced in language acquisition, retrieval, revitalization, analysis, and even preservation (Berez-Kroeker et al., 2023). These materials are intended as resource data for MTB-MLE, like digital teaching aids, modules, and lessons for the target learners.

Many words in a language are used only in some language areas, and some terms are familiar only to speakers in specific regions. The challenge lies in identifying the phonetic characteristics of language adaptations. The familiar words often have varying pronunciations and meanings in local speech, reflecting regional differences that the phonological adaptation seeks to preserve in the sound system (Rajagukguk et al., 2022). Additionally, derivational integration and inflectional word-formation were mapped morphologically. Word-formation processes also attest to affixation and clipping. Additionally, inflections for gender, number, and possessive assignments are shown (Menson-Makalingkang & Echavez, 2021). In this view, the Kawayan dialect focuses on word formation analysis.

This technological dialect documentation and preservation entails a set of actions that are more or less commonly recognized, such as recording, generating metadata, annotating media files, backing up files, and archiving materials for safekeeping (Cox, 2022). It also advocates the use of appropriate technologies to record, manage, and annotate linguistic data, thereby creating accurate records of the studied languages and providing data upon which to make generalizations (Thieberger & Berez, 2011). In this case, language development, such as the Kawayan dialect, involves studying how linguistic features vary systematically among different speaker groups or within the same speaker from different perspectives (Chambers, 2017; De La Torre & Gonong, 2020). In this view, technological dialect documentation supports MTB-MLE by utilizing digital tools to preserve and study local languages, such as the Kawayan dialect, thereby facilitating effective mother-tongue-based learning through the local contextualization of vocabulary (Purcia & Castante, 2023).

2.3. Developing Teaching Materials in MTB-MLE

There have been initiatives to study local dialects in the Philippines for inclusion in the MTB-MLE in language learning, including the under-researched Kawayan dialect from Negros Island, Philippines. The main aim of this research is to catalogue the vocabulary of the Kawayan dialect, which, despite being an active part of daily communication in Southern Negros Occidental, lacks comprehensive documentation that is essential for preserving linguistic diversity, providing detailed records that include audio-visual materials, metadata, annotations, and contextual information necessary for effective language revitalization and research as resource data in developing teaching materials in language learning (Cox, 2022; Ivushkina, 2024). Communities using the dialect in areas such as Cauayan, Hinoba-an, Ilog, Candoni, Kabankalan City, and Sipalay City, as well as the highlands between Negros Oriental and Negros Occidental, are critical to this study, highlighting the dialect's undocumented status but ongoing significance (Zorc, 1977).

In the Philippine educational setting, Mother Tongue-Based Multilingual Education (MTB-MLE) is taught in early education, particularly from kindergarten to Grade 3, and is formally adopted and practiced. However, it often coexists with deep-seated ideological and structural barriers that hinder its transformative potential (Lumanlan et al., 2024). This paradox underscores that using mother tongues in instruction does not automatically foster positive attitudes toward these languages. Instead, MTB-MLE challenges educators and institutions to critically engage with entrenched linguistic ideologies that have historically devalued local languages in favor of dominant ones (Madrurnio et al., 2016). The success of MTB-MLE thus depends not only on pedagogical reform but also on a broader restructuring of educational and social systems that perpetuate linguistic hierarchies and marginalization (Tupas, 2015).

It assumes that teachers' beliefs about language significantly affect how they use the MTBMLE policy. Some teachers adapted to the policy, while others negotiated its terms, and some exhibited subtle resistance or even challenged their language ideologies (Bernstein et al., 2023) as they faced teaching challenges that made implementation difficult. It is in the hands of the teachers who play an active role in shaping policy rather than just being passive recipients (Velasco, 2025), through identifying language policy and planning as a key responsibility of governments at the national, regional, and local levels, and relevant organizations in all areas of public policy, such as teaching MTB-MLE (Metila et al., 2025). The discussion on how language policy and planning are developed using a five-stage cycle taken from public policy studies. These steps are the emergence of language-related issues, agenda-setting, policy formulation and adoption, implementation, and evaluation (Adriano et al., 2021).

2.4. Advancing Education through Language, MTB-MLE, and Resource Design

Language is central to education (Alqahtani, 2015; Tsujii, 2021). It is the primary way learners gain knowledge, share ideas, and connect with the world (Cabansag, 2016). When education is offered in a language that students understand best, they comprehend more, participate more actively, and achieve better academic outcomes (Alqahtani, 2015; Haneda, 2014; Nwachukwu & Joseph, 2024). Modern educational frameworks emphasize the importance of linguistic inclusivity and the use of local languages to bridge the gap between home and school learning (Adnyana et al., 2025; Woolard, 2020). By appreciating learners' linguistic and cultural backgrounds, educators support cognitive development (Bautista, 2025) and help build a strong sense of identity and belonging (Chaika et al., 2024; Ibno, 2025; Rajagukguk et al., 2022). These elements are essential for creating fair and effective learning environments (Nambiar et al., 2020).

Through valuing linguistic diversity (Atifnigar et al., 2025; Bernstein et al., 2023) and addressing educational disparities among speakers of indigenous or minority languages (Trudell, 2009), MTB-MLE promotes inclusivity and access to high-quality education in the Philippines and other multilingual countries (Heraña & Azarias, 2024). Therefore, analyzing language teaching with digital lexicons (Tan & Jehom, 2024), such as the Kawayan dialect, within the framework of Mother Tongue-Based Multilingual Education (MTB-MLE), which begins with instruction in the learner's native tongue and progressively introduces other languages (Tenorio, 2024). This method acknowledges that before moving on to other languages (Ponce & Lucas, 2024), early literacy and learning are most effectively developed in a familiar linguistic context (Subiantoro et al., 2023). To facilitate easier transitions to multilingual proficiency (Metila et al., 2025), MTB-MLE prepares students to build a solid foundation in writing, reading, and critical thinking (Hossain, 2024; Lumanlan et al., 2024).

Well-crafted, culturally appropriate teaching materials that take into account the linguistic realities of learners (Barruga, 2025; Domogen, 2023) are essential to the success of MTB-MLE implementation (Hutson et al., 2024). To ensure accuracy, authenticity, and accessibility, educators, linguists, and community members collaborate when creating such materials and building facilities

(Ancheta, Jr. & Ochoco, 2024). From interactive language apps to digital storybooks, technological tools (Chun et al., 2016) are also essential for creating, recording, and sharing localized learning materials (Periñán-Pascual, 2017; Zhang, 2024). The overall objective of improving education through language inclusion and innovation is supported when resource design (Berez-Kroeker et al., 2023) aligns with learners' cultural and linguistic contexts (Ajani et al., 2024), as this enhances engagement and comprehension (Sarma, 2025).

2.5. Synthesis of the Study's Educational Scaffold

By promoting language-inclusive and culturally grounded learning, the convergence of language documentation, MTB-MLE, and educational resource design enhances the study's educational opportunity. This focuses on developing accurate linguistic resources that preserve linguistic diversity and inform instruction through language documentation, particularly with the aid of digital tools that record and analyze regional dialects, such as the Kawayan language.

In this context, these recorded languages serve as essential teaching tools in Mother Tongue-Based Multilingual Education (MTB-MLE), enabling learners to develop their literacy and cognitive abilities in their native language before progressing to other languages. For speakers of indigenous or minority languages, this position ensures that education remains equitable, meaningful, and accessible.

In the long run, this study proposes that language documentation and MTB-MLE principles are translated into practical classroom applications through well-crafted and culturally sensitive educational materials created in collaboration between educators, linguists, and local communities. Furthermore, this paper encapsulates the work of these three strands in implementation, advancing language education through linguistic and cultural empowerment, while fostering language preservation and enhancing learner identity, comprehension, and participation.

3. Methodology

3.1. Research Design

This study employed a qualitative text analysis approach, which involves a labor-intensive, more diverse and intricate process, especially when dealing with data such as transcripts of face-to-face interviews and other interactions between the researcher and informants (native speakers). In this case, the researcher restricted the use of smaller sample sizes to analyze text data, assign code labels, and progressively develop findings (Guetterman et al., 2018). The study assumed that reality was constructed through language, emphasizing the need for detached observation to prevent researcher bias from skewing the truth (Czarniawska-Joerges, 1996). This study also focused on analyzing the lexicon of the Kawayan dialect through linguistic modeling, semantics, and code glosses as part of the documentation and preservation process, serving as a resource for future utilization in creating teaching materials.

In this, qualitative textual analysis is often more valid. This study employed this method to examine messages across various media. The transcripts from native speakers in this study served as the primary source for analysis. Computer-aided tools improved the researcher's ability to study texts more thoroughly in terms of scope and size. By focusing on texts beyond the researcher's ability to manage, the primary benefits of computer-assisted textual analysis stem from its ability to handle and retrieve large amounts of data (Smith, 2017), just as the Kawayan lexicon does. Moreover, the approach of textual analysis facilitates a degree of reflexivity. This aspect proves instrumental in addressing possible challenges associated with textual analysis. Incorporating reflexivity enhances the methodological rigor and overall quality of qualitative research (Darawsheh, 2014).

3.2. Protocol of the study

This qualitative study utilized transcriptions from interviews and audio recordings as the primary source of raw data. The researcher must be aware of how their position and views may influence the data collected and should address these issues directly during data collection. To ensure a thorough approach, the researcher sought assistance from an expert in the field. The interview protocol design not only supports the research claims but also influences the generalizability or transferability of the findings (Rinehart & Earl, 2016). In this study, face-to-face interviews guided with semi-structured questionnaires and diary-keeping were utilized (Willig & Rogers, 2017).

The interview protocol for this study reflected its field experiences and considered the participants' context, culture, and beliefs within the study area, specifically the southern part of Negros Occidental, Philippines. The elements of the protocol include: (1) conducting a reverse interview to establish trust and confidence with the informant; (2) utilizing modern technology for back-listening to recorded interviews; (3) creating a comfortable environment for the informant before the interview begins; (4) addressing gender dynamics by allowing a female research assistant to lead interviews with female interviewees; (5) ensuring the informant has the final say in the conversation to maintain confidence in their responses. An effective interview protocol was designed to reflect the researcher's field experiences while carefully considering the participants' context, culture, and beliefs within the study setting, thereby enhancing the relevance and sensitivity of the data collection process by the researcher meticulously noting every detail of the interview through digital recording and sound quality of the instrument being utilized (Jamshed, 2014).

3.3. Informants of the study

The researcher employed a purposive sampling approach to select informants who were local/native speakers from Candoni, Cauayan, Hinoba-an, Ilog, Kabankalan City, and Sipalay City (CCHIKS) in the Southern Region of Negros Occidental, Philippines. Due to pandemic concerns, various conditions and limitations influenced the informants' participation in the interview. The researcher focused on three specific areas: Cauayan, Hinoba-an, and Sipalay City. A total of nine informants participated in the study. Among them were three females from Cauayan, one from Sipalay City, and two males and three females from Hinoba-an, aged 73 to 81 years old. The females were housewives, and the males were retired farmers who willingly shared their personal stories and folktales. The study's main objective was to uncover the undocumented lexicon of the Kawayan dialect, creating a dataset based on the frequency of lexicon extracted from the interview responses, noting the sounds uttered by the informants to capture the point at which written transcriptions begin to encode the language.

Furthermore, the researcher applied the selection criteria outlined by De La Torre & Gonong (2020) and made adjustments to refine the selected informants. During this process, the researcher rigorously adhered to and considered the selection criteria, seeking input from field experts regarding the researcher's informants, who should meet the following criteria: (1) A native or local speaker with over 50 years of residency in the area; (2) Ideally aged over 70; (3) Proficient in their native or local dialect; (4) Capable of recounting folktales in addition to personal life narratives.

3.4. Validity and reliability

In this study, an Excel Pivot table was employed and monitored to track timestamps (Miller, 2014) for retrieving words from the informants (native speakers). Subsequently, the data underwent cross-examination by inter-coders, who assessed and verified its validity and reliability as part of the procedure to ensure data accuracy (Cheung & Tai, 2023). These inter-coders were proficient in linguistic analysis and had previously published papers in reputable, peer-reviewed journals.

This study conducted thorough and consistent validity and reliability tests, which are essential. It is said that construct validity refers to the established operational methods for the concepts being studied, while reliability pertains to the study processes that yield consistent results (Heale & Twycross, 2015). In this study, the researcher continuously assessed and scrutinized the data through regular discussions with designated inter-coders, who were well-versed in operational measures and data analysis techniques (Fairclough, 2014). Furthermore, the process enlisted competent inter-coders to ensure validity through inter-coding and backchecking, including an audit trail. It was indeed an extensive deliberation held among the inter-coders and the researcher to produce reliable outcomes and clarify the parameters for proceeding with the analysis stage. Thus, these inter-coders continued to examine the essential concepts required to achieve a majority agreement, aiming to redefine consensus among them (Madrurnio et al., 2016). This part involved discussing and resolving the agreement between the researcher and inter-coders to create the final Kawayan lexicon as a dataset for analysis.

3.5. Data Gathering Procedure

This study followed the fieldwork protocol outlined by Zorc (1977) in the same area, which the present researcher also followed for nearly 11 months of data collection. Some modifications were made to the procedure during the data collection phases. Phase 1 involved the initial mapping of informants, with assistance from Local Government Units (LGUs) and their Senior Citizen Offices. Phase 2 entailed obtaining informed consent from the informants, which was translated into their native language. Phase 3 consisted of face-to-face interviews. The researcher collected data on specific speech variations, such as historical narratives and folklore selected by participants, as well as autobiographical or personal anecdotes, to assemble the dataset for the final analysis.

The researcher's adjustments in data collection and fieldwork were documented as follows: (1) presenting the study's purpose and objectives to the Local Government Unit (LGU); (2) Collaborating with local assistants, institutions, or organizations in the area; (3) Identifying potential informants typically aged 70 and above; (4) Mapping out the locations of the selected informants; (5) Seeking consensus and consent from informants and their immediate family members by explaining the research's purpose; (6) Employing two data collection settings: autobiographical accounts and oral traditions passed down from ancestors; (7) Revisiting identified informants to validate and confirm their narratives, both autobiographical and folklore, using electronic devices to record the informants' voice narratives. This procedure employed rigorous criteria to ensure the dataset's accuracy and authenticity, including responsiveness, reflexivity, purposive sampling, thick description, triangulation, transparency, and transferability (Tol, 2014).

In the data-gathering process, the researcher sought assistance from field experts to serve as inter-coders, consultants, and audit-trail auditors. This study focused on recording and documenting data to compile a Kawayan lexicon based on information provided by native speakers/informants. The research explored the nuanced meanings of words by presenting usage examples in various contexts and analyzing informant responses to develop a comprehensive wordlist/lexicon of the Kawayan dialect. Moreover, the study utilized different coders, including computer experts, who used software to categorize conversations into distinct words or statements.

This study centered on the robustness of the analytical perspectives employed to support the interpretations (Golafshani, 2015). It also required significant time and rigor to determine and justify the conditions, outcomes, and cases used in its application (O'Connor & Joffe, 2020). Therefore, this study included sample excerpts to clarify analytical explanations and arguments. Using the Excel Pivot Table to extract data from raw transcripts, 2,254 responses were obtained from the nine informants, leading an expert intercoder in the field to select a contextualized 278-lexicon. Figure 1 illustrates that the pivot table generates raw data from the Kawayan lexicon.



Figure 1. The pivot table contains computational raw data from the Kawayan lexicon.

3.6. Data Analysis Procedure

The approach employed in data analysis, known as phono-lexicostatistics (Guzmán Naranjo & Mertner, 2023), distinguished itself from similar studies by combining phonetic and lexical identification (Pakendorf et al., 2021). This approach utilized specific criteria derived from Nettle (1998). The researcher modified and refined the approach previously used by other researchers. These adaptations included the creation of the 278 lexicon consisting of words meeting two criteria: (1) presence in at least two out of nine informants, and (2) validation through consensus among inter-coders and intra-coders on the classification of content words as nouns, verbs, adjectives, or adverbs. Even in familiar tasks, such as learning a list of common words, the demands of the 278-lexicon task can prompt the use of strategies, and the development of strategy use during a memory task reflects the acquisition of cognitive skills (Bardel et al., 2013; Laine et al., 2024).

The inter-coders utilized the syntactic rules defined by Zorc (1977) and core grammatical categories (Corradini, 2009) to curate the wordlist, ensuring data reliability for interpretation. The category matrix from De La Torre & Gonong (2020), which included phonemic inventories and features, was repeatedly replayed from the digital recorder, with timestamps assigned, to monitor informant utterances. Then, printed copies containing words or sentence structures aligned with descriptions of nouns, adjectives, adverbs, and verbs in the dialect were distributed to the inter-coders, along with linguistic researchers, among them were the colleagues of the present researcher, for discussion to reach a consensus on content word categorization (Tegegne, 2015). Finally, another set of external experts further validated the interpretation and analysis of the finalized lexical lexicon to ensure study accuracy and precision, including segmental sounds (Brown & Sonderegger, 2025; Roettger & Baer-Henney, 2019).

As mentioned, Zorc (1977) defines voluntary texts as those that were recorded and transcribed from native speakers (informants) and processed using computer-aided tools to generate a word concordance and an Excel Pivot Table for data analysis. The dataset consisted of unstructured text data; therefore, it was suitable for demonstrating Text Data Analytics using Computational Linguistics. This analysis protocol described the complexity and volume of enormous unstructured datasets (Dehmer & Emmert-Streib, 2018; Periñán-Pascual, 2017).

Specifically, the analysis protocol of this study included (1) transcribing the conversation (interview), (2) excluding interview words to maintain the interviewee's responses as raw data, (3)

storing the raw data in a database for future use, and (4) reaching a consensus among inter-coders. The researchers filtered and analyzed the raw data to derive the final set of words for analysis. Lastly, the wordlist underwent rigorous consultation to validate, ensure reliability, and assess readiness for analysis. Any discrepancies in the data descriptions were discussed with inter-coders to establish the dataset boundaries and finalize the Kawayan dialect lexicon for analysis. The unit of analysis in this study was the word level for linguistic descriptions and code glosses, whereas semantic meanings were at the phrasal and clausal levels. Figure 2 illustrates that the Excel pivot table uses the timestamps and coding from the raw dataset of the Kawayan lexicon.

	A	B	C	D	E
1	Timestamp	Code	Code1	Text	Text1
2	0:35	01-001	01-001	Diri	Diri
138	2:35	01-022	01-022	to	to
139	2:35	01-022	01-022	ah	ah
140	2:35	01-022	01-022	Kung	Kung
141	2:35	01-022	01-022	kis-a	kis-a
142	2:35	01-022	01-022	pananom	pananom
143	2:35	01-022	01-022	sang	sang
144	2:35	01-022	01-022	lubi	lubi
145	2:35	01-022	01-022	pananom	pananom
146	2:35	01-022	01-022	sang	sang
147	2:35	01-022	01-022	kung	kung
148	2:35	01-022	01-022	ano	ano
149	2:35	01-022	01-022	lang	lang
150	2:35	01-022	01-022	kag	kag
151	2:35	01-022	01-022	mga	mga

Figure 2. The Excel pivot table of the raw dataset of the Kawayan lexicon, showing the timestamps

3.7. Ethical Considerations

Ethical considerations are necessary to ensure courtesy, respect, confidentiality, and secrecy of information gathered to establish and safeguard the conduct of this research, the researchers strictly observed the following: (1) Anonymity of the informants throughout the research conduct; (2) Confidentiality of the data provided by the informants; and (3) Adherence to the Data Privacy Act and its implementing rules on access to the personal information of the informants (Petrova et al., 2016). To ensure this, letters of communication containing the purpose of the study were directed to the relevant authorities. Upon approval, consent forms were given to the informants (Chassang, 2017) as witnessed by their immediate family members, encompassing the specific information regarding the purpose and the extent of processing the data needed, as well as the importance of participation of the informants, all of which were translated into the native tongue. The analyzed data were limited to what was necessary for the study's purpose. Only information relevant to the study's objectives was included in the analysis. Unnecessary conversations were removed from the informants' files, and their actual names were kept confidential. Furthermore, the information obtained from the informants was used only for the research study, and prior to finalization, the informants and their immediate family were informed of the pertinent findings. In other words, the information gathered was used exclusively for this study, and preliminary findings were shared with participants and their families prior to the completion of the research.

4. Results

The research focuses on describing the 278 Kawayan lexicon. Each word consists of its form (a set of phonological features) and meaning (a set of semantic units), and lexical research addresses both aspects (Zorc, 1977). During fieldwork, glosses (translations in a contact language) are provided to prompt forms that accurately convey those meanings. When compiling a lexicon, it is essential to

consider the glosses to be used carefully, the forms expected to be elicited by those glosses, and the nature of those forms. Table 1 presents the 278 Kawayan lexicon (excerpt) that can be more easily linked to the Bisayan family, specifically the West Bisayan subgroup. This wordlist is significant for shedding light on the genetic relationship between the Kawayan dialect and the broader Philippine language group.

Code glosses provide detailed explanations and capture the meanings of elements within the narratives. At times, these glosses define or clarify a word, phrase, or idiom, aiding in the reinterpretation of knowledge related to ideation (Nausa, 2019). Furthermore, 'alternate words' are included, semantically associated with the word that best fits the context or is closely related to the language or word being discussed (Brochhagen & Boleda, 2022). Incorporating code glosses and alternative words in the Kawayan lexicon in this study provides a valuable resource for future reference and MTB-MLE teaching materials, enabling researchers to delve deeper into exploration and investigation.

It delves into the presentation and dissemination of this study to its target audience and academic community, which are both insightful and significant. While exploring the Kawayan dialect represents a pioneering effort, albeit challenging, it is a noteworthy milestone in documenting a dialect that warrants inclusion in the roster of Philippine dialects. The present researcher encourages us to delve deeper into this study. Table 1 presents an excerpt from the 278 Kawayan lexicon for analysis.

Table 1. An Excerpt of 278 Kawayan Lexicon

Kawayan Words	Code-Glosses	Alternate
1 alagyan	passage	way
2 alihid	side	corner
3 amomonggo	legend creature	Ape-like Humanoid
4 anggid-anggid	same	look-alike
5 angkunon	take	grab
6 anibong	kind of leaf	
7 aroy-aroy	weak	sick
8 asawahay	married	husband and wife
9 asawahon	wife	spouse
10 babaylan	faith healer	Shaman

Prior research has focused on well-known lexicostatistical inventories and lists by Gudschinsky (215 glosses) and Swadesh (100 and 200 glosses), which are acknowledged for their non-cultural, communal language meanings. Researchers in their work have regularly used these lists/inventories, which are arranged alphabetically in both the contact or business language and the linguist's native or local language. The lists are further sorted based on suggested etymologies for each gloss, after being semantically grouped into categories such as bodily parts, pronouns, and natural experiences.

In lists used in the New Guinea region, Laycock (2006) discovered a structural grouping system that arranged items by word classes (noun, verb, adjective). However, lists utilized in the Philippines have yet to show this arrangement (Zorc, 1974), like phonology with distinguishing features, the componential (or feature) method characterizes words semantically by factoring out the most basic components in their potential universality, for instance, in terms of a fixed set of semantic characteristics. Words may be synonymous in terms of their semantic components and descriptions (Clark, 2020). These elements can illustrate how synonyms, antonyms, and other related words interact with one another (Ludewig et al., 2023).

Table 1 shows that the classification of vocabulary demands depends on the narrative type provided by native speakers or informants. For instance, informants who share their biographies will have a different category of vocabulary lists than those who tell folkloric narratives, as some word families may need to meet the same threshold in the analysis. In this view, the ratio of affixes in a word may vary, highlighting the significance of identifying and comprehending the root word to facilitate code-switching and substitute words. This study explores two potential scenarios for Kawayan word formation: affixation and reduplication, which assume that all Kawayan words utilize affixes. However, the study also considers the presence of reduplicated root words, such as 'dagyaw-dagyaw' (help one another), 'saot-saot' (dancing), and 'tipon-tipon' (getting together). This suggests a separate vocabulary list might be needed to categorize these reduplicated forms. Another issue is the lexical demand and corpus analysis, which focus on the need for more data to analyze the Kawayan corpus's lexical profile. Although the current dataset is limited, the study underscores the importance of examining variations within the data to gain a deeper understanding of word formation.

In this view, the study acknowledges that the Kawayan dialect heavily relies on affixes to convey meaning and identify alternate forms of words. eduplication, particularly with hyphenation, is considered an obligatory process in this dialect (Urbanczyk, 2017). Further research is needed to understand the whole interplay between these linguistic features. While this investigation offers valuable insights, it is essential to note that it does not provide a definitive solution for all cases. There are words in the wordlist without code glosses or alternate words, despite being present in informant utterances. These words, included for future research, highlight the need to explore their meanings further. There are predominant problems of meaning and translation evident in the sample analysis of lexical denotation and part of speech of property-denoting expressions, demonstrating that these expressions present a variety of views on the meaning of translation equivalents (Tursunovich et al., 2024).

Moreover, subject-specific vocabulary underscores the need for both quantitative and qualitative reliable criteria to identify such terms (Jablonkai, 2020). The variations and developmental use of lexical collections produce empirical evidence regarding potential differences in frequency, structure, and function across qualitative and quantitative methodological approaches (Borgstede & Scholz, 2021). This research also highlights the unique challenges of documenting and exploring the Kawayan dialect. As an undocumented dialect of the Philippines, Kawayan exhibits uncommon words rarely found even in closely related Bisayan languages.

This suggests that the ease of documentation might play a more significant role than previously presumed in understanding the Kawayan lexicon. However, other contributing factors likely warrant further investigation using comprehensive intralinguistic lexicon corpora (Durrant, 2025). It is also translated to the nearest equivalent in the Bisayan language, as well as the code glosses that sometimes emerge in a text that needs to be translated, which can constantly obstruct translators' understanding in the physical world of translation. However, little is known about how they are translated into the target language, a challenging and understudied aspect that warrants further

investigation. In the meantime, it is supported by descriptive and in-depth empirical data analyzed from the translated text.

While a reasonable number of wordlists is generally assumed to yield acceptable analyses, it suggests that a specific number might be advantageous. Understanding the local culture can foster a sense of community and belonging, making the lexicon valuable teaching materials for local language instruction and learning (Nambiar et al., 2020). Evidence from around the world demonstrates that small groups in favored locations did not merely rely on time to separate their languages from those of their neighbors. This prevailing belief is that every clan should have a unique dialect (Woolard, 2020). Even in the variant forms proper to their local languages, driving relentless diversification (Blenkinsop & Kuchta, 2024). Conversely, the wordlists provide up-to-date information and linguistic resources (e.g., folklore) that enhance language learning. It is suggested that input resources for language learning have a lower lexical challenge than real-world materials (Domogen, 2023).

Thus, teaching a mother tongue through the Kawayan lexicon can be beneficial, particularly for young learners. It also focuses on specific and specialized vocabulary that can enhance language learning outcomes once general language proficiency is established (Alqahtani, 2015). The following sample extracts showcase the semantic meanings and lexical-grammatical features of the Kawayan lexicon, highlighting its diverse structures and forms.

Extract 1

...Anu ni kay mangasawa na kag ang among anak kwan kabilog pulo.... (This is [about] to marry and then our son, hmmm... they are ten (10) siblings....)

The lexical word "mangasawa" is defined as "to marry a man (husband)." It refers to a woman who is desired and has 10 siblings. Looking at the semantic meaning of this extract and its nearest English translation, it may be determined that mangasawa comes from a root word- asawa, which means in Hiligaynon- wife. This is also correct in the Kawayan dialect, where the term asawa means wife. Unlike the term "mag-AsaWA," it refers to a couple. Still, the root word is asawa. In Tagalog, "asawa" refers to both husband and wife, regardless of gender. Generally speaking, asawa means a couple.

Extract 2

Tiya ko nakapangasawa sang siling nila bala nga Bukidnon. Native diri sa Maricalum. ([My] aunt married, they said here that is it, a (Bukidnon- a native tribe). Native here in Maricalum.)

In dissecting the lexical word "nakapangasawa," it is defined as "married to a native man." Being married to a man signals the word tiya (aunt). Pragmatically, Tiya married a native (local) man in Maricalum-[a place in Sipalay City, Negros Occidental, Philippines].

Extract 3

Ang amay ko nag asawahay sila diri.... ([My] father, [they] married here....)

In extract 3, the lexical word "asawahay" manifests a married couple. Extract 1 observes that the definition of asawa refers to a couple in general; likewise, in Tagalog, asawa is intended for both genders.

Indeed, it signifies a universal concept across many Philippine languages, highlighting the close connections between the Bisayan languages and the entire archipelago's linguistic diversity (McFarland, 2004). Canoy (2023) notes that this interconnectedness is evident in the numerous related indigenous languages found in the Philippines. Notably, certain words share striking similarities, exhibiting identical vocabulary across these languages, illustrating the country's deep linguistic ties (Masorong, 2021).

However, some problems arise in lexicographic works for short-term language documentation projects. These issues would be addressed, including the selection of a dialect, collaboration with the speech community, and the difficulty of producing a valuable piece of work that meets the scientific standards of lexicography despite limited staff, financial, and time resources (Vamarasi, 2014). A closer look at the sample extracts 1, 2, and 3 shows how lexico-grammatical features are demonstrated. Table 2 shows the sample lexico-grammatical features.

Table 2. Sample Lexico-Grammatical Features

Lexical Item	Root Word	Syllabication	Affixation
mangasawa	asawa	manGAsawa*	<i>mang-</i>
nakapangasawa	asawa	nakapanGAsawa*	<i>naka- + pang-</i>
asawahay	asawa	asaWAhay*	<i>-hay</i>

**Uppercase syllable means stressed sound, otherwise unstressed for not in an uppercase syllable*

While this research cannot definitively quantify the frequency of lexical features, particularly affixation, in the Kawayan lexicon due to the limited data available, this presents a valuable opportunity for future research. This study focuses on the form, which was critically observed by employing one of the words in the sentence, modified by adding a single affix. In contrast, all the other words in the sentence remained unchanged, including the lexical roots and affixes. Contradicting the hierarchical syntactic construct and consistent with the lexico-semantic one (Kalenkovich et al., 2022).

Still, the challenge of constructing a comprehensive wordlist for an undocumented dialect like Kawayan highlights the need for further exploration. As Wangdi & Shimray (2022) underscored, the frequency and coverage of the wordlist are crucial for in-depth linguistic analysis. Identifying the most frequent wordlist for variations in lexical needs across sub-areas in language acquisition requires a higher lexical requirement, as determined by applying a set of criteria for widespread use and relevance (Ludewig et al., 2023). However, the development process and criteria for word selection in this context require further clarification.

In preparation for the raw dataset in the analysis, computational linguistics in Text Analytics using machine tool like Excel Pivot Table utilizes in today's technology-driven society, which it is almost unimaginable to imagine the degree to which computational materials, the scope of primary and secondary storage, and software technologies are evolved to the remarkable development made in integrated of technologies and related academic fields, particularly computational linguistics (Wan Jusoh et al., 2024).

In this study, not all terms in the 278-lexicon have annotations, resulting in lower frequency counts. This might be due to challenges in categorizing informants' narratives, which focus on folklore or personal stories. This discrepancy necessitates a thorough examination of the lexicon to gain a deeper understanding of the Kawayan dialect (Masaru, 1988). The study leverages computational methods, such as text analytics in Excel Pivot Tables, to analyze text, highlighting that raw data only gains meaning through interpretation, as Halder (2023) noted. Consequently, the nature of the data is independent of the analytical approach used.

This study introduces the 278 lexicon to be scrutinized from the Kawayan dialect and aims to provide a foundation for future research into its complex linguistic structure. Highlighting the importance of a transparent methodology, including sampling, data sources, and analysis, and integrating both quantitative and qualitative approaches, suggests areas for deeper expert exploration (Amini Farsani et al., 2021; Gopalakrishnan et al., 2024).

Figure 3 presents and explores the complexities of computational linguistics theory and the role of technology in the documentation and preservation of processing using the Excel pivot table of text analytics as the machine tool for analyzing large datasets as a basis for resource data materials in MTB-MLE implementation. Figure 3 illustrates the integration of technology-mediated lifestyle for Kawayan dialect preservation with MTB-MLE pedagogical approaches.

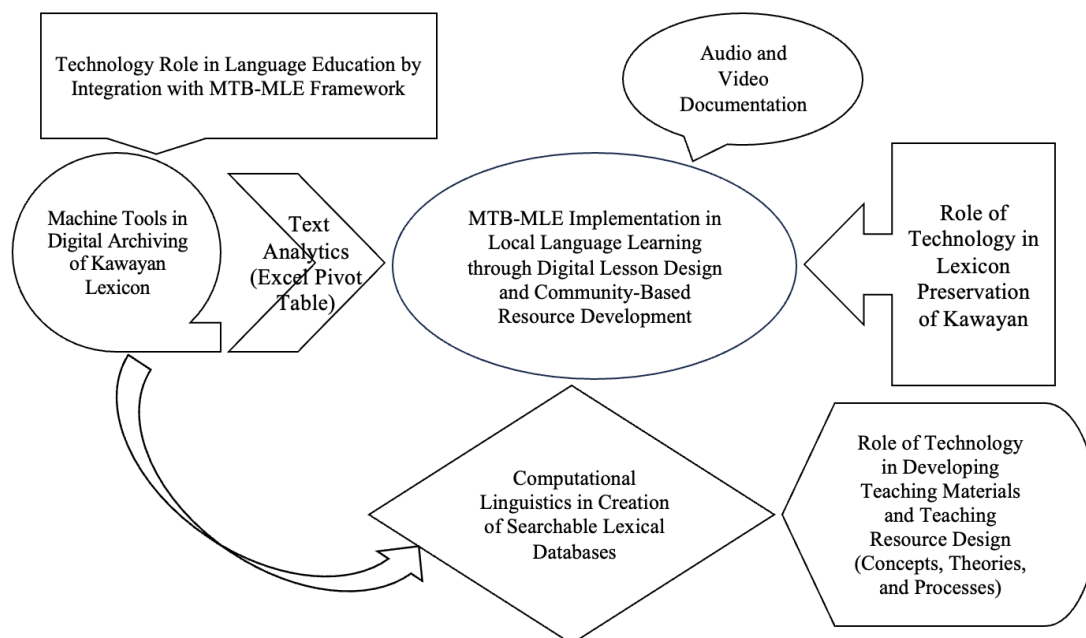


Figure 3. Technology-mediated lifecycle for Kawayan dialect preservation towards MTB-MLE pedagogical integration

The diagram presents a comprehensive educational technology model with a focus on revitalizing and pedagogically applying the Kawayan dialect. This framework posits that the primary enabling mechanism for a comprehensive approach to curriculum development and language instruction is the integration of technology. The initial phases, which include computational linguistics for creating searchable databases and digital archival science for preserving the lexicon, are essentially processes of designing and curating educational resources to ensure the accuracy and accessibility of linguistic data. By identifying the most common vocabulary and grammatical structures essential for efficient instruction and language acquisition, text analytics provides the empirical basis for data-driven instructional design.

The accompanying digital resources, strategically used for pedagogical integration, are enhanced by multimodal documentation (audio/video). This critical step involves utilizing technology to systematically produce authentic teaching materials that align with the Mother Tongue-Based Multilingual Education (MTB-MLE) Framework. Consequently, the entire model serves as a model for converting a static heritage language into a dynamic, technologically mediated local language through digital lesson design and community-based resource development.

5. Discussion

5.1. Foundations in Form and Meaning, Guided by Lexical Theory

Lexical research, as theorized by, treats phonological features and their meaning, defined by sets of semantic units. In this view, the fieldwork underpinning this study reveals that code glosses, which are translations in a contact language like the Kawayan dialect, play a pivotal role in eliciting the correct forms from speakers. These code glosses and translations of familiar words are resources for creating MTB-MLE teaching materials for learners, including pronunciations and meanings.

Rajagukguk et al. (2022) noted that familiar words often have varying pronunciations and meanings in local speech, reflecting regional differences that the phonological adaptation aims to preserve in the sound system.

Taking careful account of the glosses used, the elicited forms, and their anticipated phonological and semantic characteristics is fundamental when compiling a lexicon. This meticulous approach ensures that each lexical entry accurately captures form and meaning, providing a reliable foundation for systematic documentation and linguistic analysis (Gardner, 2007). This dual concern ensures that each entry in the lexicon reflects accurate pronunciation and appropriate meaning, forming a robust basis for systematic documentation and analysis (Nicklin et al., 2025). Thus, this case highlights the importance of the Kawayan lexicon for effective vocabulary building and providing sufficient vocabulary resources for language teaching and learning, particularly in high-frequency contexts (Castellano-Risco, 2024).

5.2. Kawayan Wordlist as Genetic and Pedagogical Resource

The research presents a curated wordlist featuring 278 entries in the Kawayan dialect, which have been identified as closely aligned with the West Bisayan subgroup, a significant finding in mapping lexical genealogy within Philippine Austronesian languages. This combination of methods enhances the list's utility for genetic linguistic studies and practical teaching applications, such as creating digital teaching materials for mother tongue lessons. It provides a starting point for future researchers and is a well-supported resource for exploring comparative and contextual vocabulary, enabling them to teach Kawayan vocabulary systematically to learners. Bocanegra et al. (2025) noted that determining and thoroughly examining the language using digital tools to enhance and promote the vocabulary of the target language contributes to language literacy development among learners.

Moreover, it provides detailed interpretations within narrative settings and primary forms, such as code glosses. Alternative word entries give related options or suitable variations that deepen understanding and increase the adaptability of vocabulary entries. This approach adds depth and practicality. It provides future researchers with practical tools for analysis and comparison, thereby enhancing the interpretive layer (Bender et al., 2012).

The significance and scholarly merit of the lexicographic effort mark a groundbreaking milestone in the documentation of the Kawayan dialect of Southern Negros. As a pioneering endeavor, it foregrounds a previously underdocumented variety and encourages deeper academic attention toward dialects still absent from formal linguistic records. Furthermore, the careful presentation of word glosses, form variation, and alternative lexical items enhances interpretability and strengthens cross-disciplinary relevance, thereby bridging historical linguistics, language pedagogy, and dialectology (Hazen, 2014).

Ultimately, this work provides a comprehensive database and a systematic model. It encourages us to explore documentation and sharing practices for the Kawayan dialect in teaching the learners and future generations.

5.3. Affixation, Reduplication, and Documentation Gaps

In the Kawayan dialect, affixation and reduplication, often accompanied by hyphenation, play pivotal roles in conveying meaning and generating alternative forms, with processes such as root-level duplication being obligatory in certain lexical items. However, notable gaps persist in the documentation, like some elicited words lack code glosses or alternate entries despite appearing in informant data, underscoring the need for further semantic and lexical analysis (Khaya, 2025).

Additionally, challenges arise in capturing nuanced meaning and grammatical class variations, as supported by a previous study, which viewed language as a minor variable that can be readily

overcome through standard classroom instruction on the reasons for educational failure. However, language has begun to emerge as a significant variable in producing gains in educational efficiency. Hence, they propose that children are taught to read first in the local language (via the local language) and are taught other key subjects via the local language as well, and especially within property-denoting expressions that highlight the complex interplay between lexical denotation, semantic range, and part-of-speech categorization included in teaching local language (Henkin-Roitfarb et al., 2025).

5.4. The Necessity of Robust Methodological Approaches

Quantitative and qualitative frameworks are essential to reliably document subject-specific or dialect-sensitive vocabulary (Jablonkai, 2020). This research emphasis extends to the Kawayan lexicon, where systematic methods such as tracking frequency, structural patterns, and functional distribution would bolster empirical understanding (Gao et al., 2024). The lexicon's uniqueness, underscored by words not found in closely related Bisayan languages, invites more in-depth corpus-based and intralinguistic investigation to uncover underlying patterns (Masorong, 2021). Likewise, the role of code glosses becomes critical in translation fidelity and interpretation; yet, the cognitive processes that translators engage in remain underexplored and merit descriptive study (Alzain, 2024).

In this study, the Kawayan lexicon is a dialect-sensitive vocabulary, with uniqueness accentuated by words that are not found in any other Bisayan languages. Therefore, continuous teacher training plays a crucial role in developing curricula and teaching materials, as well as in adopting digital technology to target learners, especially in a mother-tongue-first multilingual education programme (Kosonen, 2017, cited in Herzog-Punzenberger et al., 2017).

5.5. Lexical Resources, Cultural Identity, and Pedagogy

Lexicons or wordlists rooted in local culture foster a more profound sense of community and belonging, making them invaluable pedagogical tools (Anyichie et al., 2023). Additionally, linguistic diversification is often actively maintained within tight-knit communities, with elders reinforcing local variants as markers of identity, rather than being shaped passively over time (Evans & Levinson, 2009). Beyond cultural resonance, carefully curated wordlist/lexicon like Kawayan, especially those that integrate cultural artifacts like folklore, offer language learners more accessible inputs and materials, thereby reducing lexical difficulty and supporting gradual proficiency building in extracting Kawayan lexicon to create modules and lessons focusing on vocabulary in teaching the mother tongue, especially kindergarten up to grade 3 (Hutson et al., 2024).

Consequently, when general proficiency is achieved, focusing instruction on specific and subject-related vocabulary, facilitated by a well-structured wordlist, enhances learning efficiency and outcome effectiveness, explicitly focusing on Kawayan lexicon as learners' first language (L1) (Fengyu, 2023). In a previous study by Benson (2004), as cited in Da Silva et al. (2024), it was noted that mother-tongue-based bilingual programs utilize the learner's first language, known as L1, to teach beginning reading and writing skills alongside academic content. This study is also supported by previous findings indicating that the approach provides educational skills and offers important insights into language learning, communication, and cognitive development across various classroom settings (Coles-Harris, 2017; Grant, 2012).

5.6. Affixation Dynamics & Wordlist Limitations

Although this investigation cannot definitively quantify the frequency of affixation and other lexical features within the Kawayan lexicon due to the limited data available, it opens up vital avenues for future research into morphological dynamics. For instance, Kalenkovich et al. (2022) illustrate how adding a single affix to a word in a sentence, while all other lexical elements remain unchanged, can

illuminate the interaction between hierarchical syntax and lexico-semantic structure. Although not directly focused on Kawayan, this methodological insight highlights the value of minimal contrast in examining form-function relationships. Meanwhile, foundational work in vocabulary studies emphasizes that frequency and coverage are critical metrics for constructing a meaningful lexicon (Gardner & Davies, 2014; Tong et al., 2025). In the context of Kawayan, however, it notes that the criteria for selecting entries remain under-specified, suggesting a need to define clearer inclusion standards calibrated by sub-area variation and relevance to language acquisition (Mtallo & Kimambo, 2024). This argument is also supported by studies on related educational vocabulary (Lei & Reynolds, 2022).

This foundational vocabulary from Kawayan's lexicon provides detailed, scientific resources for this study. This is also supported by previous studies by Cox (2022) and Ivushkina (2024), which detailed records that include audio-visual materials, metadata, annotations, and contextual information necessary for effective language revitalization and research, serving as a valuable resource for developing teaching materials in language learning.

5.7. Computational Methods and Interpretation Challenges

The utility of computational methods, such as Excel pivot tables, which this study utilizes to analyze and contextualize wordlist data in language documentation, is exemplified by the Kawayan dialect. This highlights a critical intersection of traditional fieldwork and modern linguistic technology tailored to the digital platform of teaching and learning the mother tongue. A previous study by Tsujii (2021) outlines the unprecedented scale and integration of computational tools across lexicon analysis, storage, and archival processes. This also affirms that the raw linguistic data only gains analytic value when interpreted through robust methodological frameworks (Iso et al., 2020).

Similarly, narrative-based gaps in annotation, such as unglossed terms drawn from folklore, underscore the limitations of frequency counts alone and the need for qualitative scrutiny, as noted by Safari (2018). Moreover, Ngulube (2023) and Oluka (2025) argue that data becomes meaningful only through interpretive analysis, regardless of analytical tools employed. This study presents the Kawayan lexicon, derived from interpretive analysis, utilizing a pivot table of text analytics as a tool for extracting and documenting resource materials for language learning, such as teaching the mother tongue. This notion is also supported by Berez-Kroeker et al. (2023), who stated that the current technologies enable the creation and use of documentary materials that record and annotate language activities, making them valuable resources for teaching, learning, and preserving languages in MTB-MLE contexts.

Finally, the study presents a foundation of 278 Kawayan wordlist, which offers a promising basis for future exploration. However, the researcher prioritizes a transparent methodology, including sampling, data selection, and mixed-methods analysis, to advance the understanding of this underdocumented variety (Riazi & Farsani, 2024).

5.8. Lexicon's Educational Inclusion

In contemporary efforts to document and preserve underrepresented Philippine dialects, the systematic compilation of lexicons is indispensable for educational inclusion, particularly when rapid digitalization and the dominance of foreign languages threaten linguistic diversity (Lamorinas et al., 2024). The pioneering Kawayan dialect study exemplifies this, leveraging an Excel pivot table to effectively analyze informants' lexical data and enhance its validity through cross-checking, thereby providing a robust foundation for pedagogy, revitalization, and future research. This approach focuses on MTB-MLE as the foundation for preserving and learning the Kawayan lexicon digitally, utilizing text analytics through an Excel pivot table. This digitalized documentation provides resource materials for crafting modules and lessons in teaching the mother tongue. This is also supported by

UNESCO (2025), which has long championed the cause of mother-tongue-based multilingual education, emphasizing that learning that begins in one's home language enhances learning outcomes and preserves cultural heritage. This emphasis on migration, technology, and recognition of multilingualism's benefits has reshaped the linguistic landscape, guiding policies that value multilingualism in education.

Moreover, the study by Ahmed (2025), which advocates for linguistic human rights and diversity, requires prioritizing mother-tongue-based multilingual education and establishing spaces that protect and promote endangered languages. By incorporating vocabulary (derived from the Kawayan dialect) based on personal narratives and local folklore of native speakers (informants), this approach not only aids learning and lexical acquisition (Hossain, 2024) but also highlights the importance of interpreting complex or context-dependent terms (Mohamed et al., 2024).

Indeed, this study of the Kawayan dialect highlights how semantic and narrative nuances must inform assessments of wordlist utility in understanding the dialect's semantic and grammatical structures (Gennari et al., 2007). This approach asserts that technology-assisted lexicographic methods (Jackson, 2022) and qualitative scrutiny (Bryda & Costa, 2023) provide a promising approach to safeguarding local languages, such as the Kawayan lexicon, in educational settings. This study is supported by a previous study on using translanguaging pedagogy that attempts to improve sustainable L1-mediated learning for primary graders grounded in teachers' opinions, including the availability of resources, teachers' high-quality training, and collaborative work, which can leverage mother-tongue-based schooling and society's culture in education (Ouattara et al., 2025). This is also highlighted in the previous study, which found that the wordlist has been systematically analyzed using digital tools to enhance the target language's vocabulary and support learners' language literacy development, utilizing the digital lexicon of the Kawayan dialect.

6. Conclusion

Learning vocabulary systematically from lists, such as the Kawayan lexicon, is essential in vocabulary programs, particularly through mother-tongue-based education. It is essential to provide students with fair learning opportunities and involve them actively in working with vocabulary suited to their level, such as the word list from the Kawayan lexicon study. However, learning challenges, such as difficulties contextualizing the words in reading materials, may arise. Students might also focus more on spelling and definitions than on practical usage in speech and writing, especially when they are unfamiliar with the Kawayan dialect.

Speakers of the Kawayan dialect have unique accents and a rich, contextually variable vocabulary. This dialect's distinctiveness shines in its word variety. Research on this dialect employs a mixed-method approach, utilizing Excel pivot tables for word organization and expert collaboration for data accuracy. This study represents a significant step toward understanding the vocabulary of the Kawayan dialect in Southern Negros, Philippines. It highlights its unique word features and the need for more research into its morphosemantic roots.

The current research presents the Kawayan wordlist, which includes code glosses and alternate terms. Many experts find this technique important for mastering new vocabulary. They argue that learners should engage with words in various meaningful contexts to effectively understand their complex nature.

In learning a new dialect like Kawayan, decontextualized methods are used, in which words are taught outside their usual context and lack real communicative value. However, it is also recommended to use contextualized methods, which are important for vocabulary learning. This approach enables students to use the language authentically in real-world situations, exposing them to a variety of words in the Kawayan dialect.

Future researchers face the challenge of seizing opportunities to delve deeper and establish connections with their existing knowledge of other Bisayan languages, such as Hiligaynon and Waray-Waray. While significant progress has been made in compiling and exploring the Kawayan lexicon, a gap remains regarding the undocumented dialects in the Philippines and addressing the scarcity of mother-tongue instructional programs.

7. Suggestion

The study's findings provide several suggestions for integrating contextualized vocabulary instruction. To address the limitations of decontextualized learning, educators should prioritize integrating Kawayan vocabulary into meaningful, real-life contexts, such as storytelling, dialogue, and local cultural narratives. Contextualized instruction ensures that learners move beyond rote memorization to practical application, which is essential for achieving communicative competence in the dialect.

In terms of Developing Multilevel Lexical Resources Tailored to Learners, the lexical richness and variability of the Kawayan dialect, combined with the creation of leveled vocabulary lists aligned with learners' language proficiency, are crucial. This Kawayan lexicon should be segmented by difficulty and thematic relevance, accompanied by code glosses, example sentences, and cultural notes to aid comprehension and retention across diverse learner levels.

In the age of Utilizing Digital Tools for Lexicon Analysis and Pedagogy, tools such as Excel pivot tables for word organization can be expanded into educational applications. Teachers and researchers can develop interactive digital glossaries and vocabulary tracking systems to enhance engagement, accuracy, and accessibility for learners and fellow researchers working with the Kawayan dialect.

It is essential to Bridge Lexical Research with Regional Bisayan Languages to explore comparative studies between the Kawayan lexicon and related Bisayan languages (e.g., Hiligaynon, Cebuano, Waray-Waray) to understand morphosemantic patterns and shared linguistic features. Such cross-linguistic insights can support pedagogical strategies and deepen understanding of the dialect's evolution and structural uniqueness.

Examining Support Teacher Training and Materials Development in MTB-MLE, it is proposed that the gap in mother-tongue-based instruction be reduced, and that targeted teacher training focus on effectively utilizing dialect-based vocabulary resources, such as the Kawayan lexicon/wordlist. This includes training in contextual language teaching, pronunciation guidance, and the creation of culturally relevant instructional materials rooted in the local linguistic environment.

Declarations

Author Contributions. Pilar, John Gerald A.: Conceptualization, literature review, data analysis, interpretation, review editing, and writing. The author has read and approved the published version of the article. The author uses ChatGPT in this work to improve the manuscript's syntax. The author states that he took full responsibility for the content of the published article and reviewed and edited the final output as necessary.

Conflicts of Interest. The author declares no conflict of interest.

Funding. Carlos Hilado Memorial State University, Negros Occidental, Philippines, through the Research and Development Services (RDS) research fund.

Ethical Approval. Carlos Hilado Memorial State University's Research, Extension, and Intellectual Property Committee (REICO) approved the study.

Data Availability Statement. The raw data supporting this study's findings are available from the author upon reasonable written request.

Acknowledgments. The researcher wishes to acknowledge Carlos Hilado Memorial State University, under the administration of University President Dr. Norberto P. Mangulabnan, as well as the informants, research assistants, and Local Government Units of Cauayan, Sipalay, and Hinobaan.

References

- Adamou, E., Guillaume, S., & Michaud, A. (2025). The Pangloss Collection: Opening up research data on endangered and underdocumented languages. *Language*, 101(1), e38–e59. <https://doi.org/10.1353/lan.2025.a954237>
- Adnyana, I. B. A., Sitawati, A. A. R., Ghafar, Z. N., & Hazaymeh, O. M.-A. A. (2025). Mother tongue matters: A critical study of indigenous language integration in formal education systems. *Journal of Language, Literature, Social and Cultural Studies*, 3(2), 113–127. <https://doi.org/10.58881/jllscs.v3i2.346>
- Adriano, Ma. N. I., Franco, N. T., & Estrella, E. A. (2021). Language-in-education policies and stakeholders' perception of the current MTB-MLE policy in an ASEAN country. *The Australian Journal of Language and Literacy*, 44(1), 84–99. <https://doi.org/10.1007/BF03652067>
- Ahmed, A. (2025). Distractive language education policies and the endangerment of Indigenous languages in Bangladesh. *Current Issues in Language Planning*, 26(3), 292–307. <https://doi.org/10.1080/14664208.2024.2370670>
- Ajani, Y. A., Oladokun, B. D., Olarongbe, S. A., Amaechi, M. N., Rabi, N., & Bashorun, M. T. (2024). Revitalizing Indigenous Knowledge Systems via Digital Media Technologies for Sustainability of Indigenous Languages. *Preservation, Digital Technology & Culture*, 53(1), 35–44. <https://doi.org/10.1515/pdte-2023-0051>
- Alqahtani, M. (2015). The importance of vocabulary in language learning and how it can be taught. *International Journal of Teaching and Education*, III(3), 21–34. <https://doi.org/10.20472/TE.2015.3.3.002>
- Alzain, M. (2024). Code-Switching in Translation: Linguistic Analysis of Multilingual Texts and their Translations. *International Journal of Linguistics, Literature and Translation*, 7(10), 119–134. <https://doi.org/10.32996/ijllt.2024.7.10.12>
- Amini Farsani, M., Jamali, H. R., Beikmohammadi, M., Ghorbani, B. D., & Soleimani, L. (2021). Methodological orientations, academic citations, and scientific collaboration in applied linguistics: What do research synthesis and bibliometrics indicate? *System*, 100, 102547. <https://doi.org/10.1016/j.system.2021.102547>
- Ancheta, J., O. O., & Ochoco, C. A. E. (2024). Enhancing Language Instruction: A Study on MTB-MLE Implementation in Urban Education. *Diversitas Journal*, 9(3). <https://doi.org/10.48017/dj.v9i3.2996>
- Ancheta, O. (2024). *ScholarOne - Enhancing Language Instruction: A Study on MTB-MLE Implementation in Urban Education*. Preprints. <https://doi.org/10.31124/advance.172355671.13051672/v1>
- Anyichie, A. C., Butler, D. L., Perry, N. E., & Nashon, S. M. (2023). Examining Classroom Contexts in Support of Culturally Diverse Learners' Engagement: An Integration of Self-Regulated Learning and Culturally Responsive Pedagogical Practices. *Frontline Learning Research*, 11(1), 1–39. <https://doi.org/10.14786/flr.v11i1.1115>

- Atifnigar, H., Dawlatzai, A. K., & Habib, S. (2025). Effective Strategies for Preserving Mother Tongues: Approaches to Linguistic and Cultural Sustainability. *European Journal of Contemporary Education and E-Learning*, 3(2), 3–14. [https://doi.org/10.59324/ejceel.2025.3\(2\).01](https://doi.org/10.59324/ejceel.2025.3(2).01)
- Bardel, C., Lindqvist, C., & Laufer, B. (Eds.). (2013). *L2 vocabulary acquisition, knowledge, and use: New perspectives on assessment and corpus analysis* (1st ed.). EuroSla.
- Barruga, B. M. (2025). Classroom implementation of the mother tongue-based multilingual education policy by Masbatenyo public elementary teachers: A case study. *Language Policy*, 24(1), 21–50. <https://doi.org/10.1007/s10993-024-09691-w>
- Bautista, R. (2025). Community-based Language Development of Northern Catanduanes, Philippines. *Diversitas Journal*, 10(special_1). https://doi.org/10.48017/dj.v10ispecial_1.3176
- Bender, E. M., Schikowski, R., & Bickel, B. (2012). *Deriving a lexicon for a precision grammar from language documentation resources: A case study of Chintang*. <https://doi.org/10.5167/UZH-77707>
- Berez-Kroecker, A. L., Gabber, S., & Slayton, A. (2023). Recent Advances in Technologies for Resource Creation and Mobilization in Language Documentation. *Annual Review of Linguistics*, 9(1), 195–214. <https://doi.org/10.1146/annurev-linguistics-031220-120504>
- Bernstein, K. A., Anderson, K. T., Close, K., & Rodriguez Martinez, S. (2023). Teacher beliefs about multilingual learners: How language ideologies shape teachers' hypothetical policymaking. *International Multilingual Research Journal*, 17(3), 191–219. <https://doi.org/10.1080/19313152.2023.2182094>
- Bird, S., & Simons, G. (2003). Seven Dimensions of Portability for Language Documentation and Description. *Language*, 79(3), 557–582. <https://doi.org/10.1353/lan.2003.0149>
- Blenkinsop, S., & Kuchta, E. C. (2024). *Ecologizing Education: Nature-Centered Teaching for Cultural Change* (1st ed.). Cornell University Press. <https://doi.org/10.7591/cornell/9781501774713.001.0001>
- Bocanegra, A., Sedillo, Ma. E., & Futralan, Ma. C. (2025). Efficacy of Digital Tools for Students' Vocabulary Expansion. *Journal of Interdisciplinary Perspectives*, 3(7). <https://doi.org/10.69569/jip.2025.346>
- Borgstede, M., & Scholz, M. (2021). Quantitative and Qualitative Approaches to Generalization and Replication—A Representationalist View. *Frontiers in Psychology*, 12, 605191. <https://doi.org/10.3389/fpsyg.2021.605191>
- Brochhagen, T., & Boleda, G. (2022). When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition*, 226, 105179. <https://doi.org/10.1016/j.cognition.2022.105179>
- Brock-Utne, B. (2012). Language and inequality: Global challenges to education. *Compare: A Journal of Comparative and International Education*, 42(5), 773–793. <https://doi.org/10.1080/03057925.2012.706453>
- Brown, J., & Sonderegger, M. (2025). A sociophonetic study of creaky voice across language, gender and age in Canadian English-French bilinguals. *Journal of Phonetics*, 112, 101431. <https://doi.org/10.1016/j.wocn.2025.101431>
- Bryda, G., & Costa, A. P. (2023). Qualitative Research in Digital Era: Innovations, Methodologies and Collaborations. *Social Sciences*, 12(10), 570. <https://doi.org/10.3390/socsci12100570>

- Cabansag, J. N. (2016). The Implementation of Mother Tongue-Based Multilingual Education: Seeing It from the Stakeholders' Perspective. *International Journal of English Linguistics*, 6(5), 43. <https://doi.org/10.5539/ijel.v6n5p43>
- Canoy, F. T. (2023). A Contrastive Analysis of English and Bisayan Languages: Cases of Pronouns in Focus. *Asian Journal of Education and Social Studies*, 46(3), 53–58. <https://doi.org/10.9734/ajess/2023/v46i31006>
- Castellano-Risco, I. (2024). An Introduction to Vocabulary Studies: Knowledge and Frequency. In I. Castellano-Risco, *Lexical Demands in CLIL Primary-School Textbooks* (pp. 9–21). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-75160-8_2
- Chaika, O., Sharmanova, N., & Makaruk, O. (2024). Revitalising Endangered Languages: Challenges, Successes, and Cultural Implications. *Futurity of Social Sciences*, 38–61. <https://doi.org/10.57125/FS.2024.06.20.03>
- Chambers, J. (2017). William Labov: An Appreciation. *Annual Review of Linguistics*, 3(1), 1–23. <https://doi.org/10.1146/annurev-linguistics-051216-040225>
- Chassang, G. (2017). The impact of the EU general data protection regulation on scientific research. *Ecancermedicalscience*, 11. <https://doi.org/10.3332/ecancer.2017.709>
- Cheung, K. K. C., & Tai, K. W. H. (2023). The use of intercoder reliability in qualitative interview data analysis in science education. *Research in Science & Technological Education*, 41(3), 1155–1175. <https://doi.org/10.1080/02635143.2021.1993179>
- Chun, D., Kern, R., & Smith, B. (2016). Technology in Language Use, Language Teaching, and Language Learning. *The Modern Language Journal*, 100(S1), 64–80. <https://doi.org/10.1111/modl.12302>
- Clark, B. (2020). The Evolvability of Words: On the Nature of Lexical Items in Minimalism. *Frontiers in Psychology*, 10, 3071. <https://doi.org/10.3389/fpsyg.2019.03071>
- Coles-Harris, E. H. (2017). Perspectives on the motivations for phonetic convergence. *Language and Linguistics Compass*, 11(12), e12268. <https://doi.org/10.1111/lnc3.12268>
- Corradini, R. (2009). A Peircian Approach to Hiligaynon Causatives. *Oceanic Linguistics*, 48(2), 337–345. <https://doi.org/10.1353/ol.0.0046>
- Cox, C. (2022). Managing Data in a Language Documentation Corpus. In A. L. Berez-Kroeker, B. McDonnell, E. Koller, & L. B. Collister (Eds.), *The Open Handbook of Linguistic Data Management* (pp. 277–286). The MIT Press. <https://doi.org/10.7551/mitpress/12200.003.0027>
- Coxhead, A., McLaughlin, E., & Reid, A. (2019). The development and application of a specialised word list: The case of Fabrication. *Journal of Vocational Education & Training*, 71(2), 175–200. <https://doi.org/10.1080/13636820.2018.1471094>
- Crystal, D. (2001). *Language and the Internet* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139164771>
- Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes*, 33, 100–118. <https://doi.org/10.1016/j.jeap.2018.02.003>
- Czarniawska-Joerges, B. (1996). Realities and relationships. Soundings in social construction. *Scandinavian Journal of Management*, 12(4), 468–470. [https://doi.org/10.1016/0956-5221\(97\)81362-0](https://doi.org/10.1016/0956-5221(97)81362-0)

- Da Silva, C., Pereira, F., & Amorim, J. P. (2024). The integration of indigenous knowledge in school: A systematic review. *Compare: A Journal of Comparative and International Education*, 54(7), 1210–1228. <https://doi.org/10.1080/03057925.2023.2184200>
- Dang, T. N. Y., Webb, S., & Coxhead, A. (2022). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, 26(4), 617–641. <https://doi.org/10.1177/1362168820911189>
- Darawsheh, W. (2014). Reflexivity in research: Promoting rigour, reliability and validity in qualitative research. *International Journal of Therapy and Rehabilitation*, 21(12), 560–568. <https://doi.org/10.12968/ijtr.2014.21.12.560>
- De La Torre, R. G., & Gonong, G. O. (2020). A phono-lexicostatistical analysis of Bikol-Sorsogon varieties. *The Normal Lights*, 14(2). <https://doi.org/10.56278/tnl.v14i2.1655>
- Dehmer, M., & Emmert-Streib, F. (Eds.). (2018). *Frontiers in data science*. CRC Press.
- Dobrin, L. M., & Berson, J. (1999). Speakers and language documentation. In P. K. Austin & J. Sallabank (Eds.), *The Cambridge Handbook of Endangered Languages* (1st ed., pp. 187–211). Cambridge University Press. <https://doi.org/10.1017/CBO9780511975981.010>
- Domogen, J. C. (2023). Authenticity in Language Learning: A Study of Language Materials in Public High Schools. *OKARA: Jurnal Bahasa Dan Sastra*, 17(2), 193–211. <https://doi.org/10.19105/ojbs.v17i2.8577>
- Durrant, P. (2025). What can a corpus tell us about school writing? Findings, challenges, and future directions. *Applied Corpus Linguistics*, 5(2), 100134. <https://doi.org/10.1016/j.acorp.2025.100134>
- Erzsébet, T.-C., Marta, B., Francesco, G., Femmy, A., Mirjam, B., Erik, B., Vera, C., Rita, G., Peter, G., Péter, K., Walter, S., Bartłomiej, S., Maria, V.-R., & Ulrike, W. (2024). *Research Data Management for Arts and Humanities: Intergrating Voices of the Community*. <https://doi.org/10.47952/gro-publ-218>
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429–448. <https://doi.org/10.1017/S0140525X0999094X>
- Fairclough, N. (2014). *Critical Language Awareness* (0 ed.). Routledge. <https://doi.org/10.4324/9781315845661>
- Fengyu, Z. (2023). The Impact of Vocabulary Learning Methods on Students' Vocabulary Application Skills. *English Language Teaching and Linguistics Studies*, 5(4), p206. <https://doi.org/10.22158/eltls.v5n4p206>
- Fettes, M. (2023). Language and the Sustainable Development Goals: Challenges to Language Policy and Planning. In L. J. McEntee-Atalianis & H. Tonkin (Eds.), *Language and Sustainable Development* (Vol. 32, pp. 11–26). Springer International Publishing. https://doi.org/10.1007/978-3-031-24918-1_2
- Frederick, M., Roeschley, A., & Zavalina, O. L. (2025). Beyond language archives: Proposing the archival community informatics framework as an interdisciplinary link to revitalization lexicography. *Information Research an International Electronic Journal*, 30(iConf), 1169–1179. <https://doi.org/10.47989/ir30iConf46942>

- Gao, Z., Li, X., & Liao, H. (2024). Teacher support and its impact on ESL student engagement in blended learning: The mediating effects of L2 grit and intended effort. *Acta Psychologica*, 248, 104428. <https://doi.org/10.1016/j.actpsy.2024.104428>
- García, O., & Wei, L. (2014). Language, Bilingualism and Education. In O. García & L. Wei, *Translanguaging* (pp. 46–62). Palgrave Macmillan UK. https://doi.org/10.1057/9781137385765_4
- Gardner, D. (2007). Validating the Construct of Word in Applied Corpus-based Vocabulary Research: A Critical Survey. *Applied Linguistics*, 28(2), 241–265. <https://doi.org/10.1093/applin/amm010>
- Gardner, D., & Davies, M. (2014). A New Academic Vocabulary List. *Applied Linguistics*, 35(3), 305–327. <https://doi.org/10.1093/applin/amt015>
- Gazzola, M., Grin, F., Cardinal, L., & Heugh, K. (2023). Language policy and planning From theory to practice. In M. Gazzola, F. Grin, L. Cardinal, & K. Heugh, *The Routledge Handbook of Language Policy and Planning* (1st ed., pp. 1–31). Routledge. <https://doi.org/10.4324/9780429448843-1>
- Gennari, S. P., MacDonald, M. C., Postle, B. R., & Seidenberg, M. S. (2007). Context-dependent interpretation of words: Evidence for interactive neural processes. *NeuroImage*, 35(3), 1278–1286. <https://doi.org/10.1016/j.neuroimage.2007.01.015>
- Golafshani, N. (2015). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2003.1870>
- Gopalakrishnan, A., Mathieu, C. S., & LaScotte, D. K. (2024). Investigating Mixed Methods Research in Applied Linguistics: Methodological Avoidance and Possible Barriers in the Field. *Studies in Applied Linguistics and TESOL*, 24(1). <https://doi.org/10.52214/salt.v24i1.10190>
- Grant, C. (2012). *Strengthening the Vitality and Viability of Endangered Music Genres: The Potential of Language Maintenance to Inform Approaches to music Sustainability*. Griffith University. <https://doi.org/10.25904/1912/3512>
- Guetterman, T. C., Chang, T., DeJonckheere, M., Basu, T., Scruggs, E., & Vydiswaran, V. V. (2018). Augmenting Qualitative Text Analysis with Natural Language Processing: Methodological Study. *Journal of Medical Internet Research*, 20(6), 231. <https://doi.org/10.2196/jmir.9702>
- Guzmán Naranjo, M., & Mertner, M. (2023). Estimating areal effects in typology: A case study of African phoneme inventories. *Linguistic Typology*, 27(2), 455–480. <https://doi.org/10.1515/lingty-2022-0037>
- Halder, A. (2023). *How to Use Excel Pivot Tables for Summarizing and Presenting Data*. SAGE Publications Ltd. <https://doi.org/10.4135/9781529670332>
- Haneda, M. (2014). From academic language to academic communication: Building on English learners' resources. *Linguistics and Education*, 26, 126–135. <https://doi.org/10.1016/j.linged.2014.01.004>
- Hazen, K. (2014). Dialectology. In K. Hazen, *Linguistics*. Oxford University Press. <https://doi.org/10.1093/obo/9780199772810-0163>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Henkin-Roitfarb, R., Uziel, S., & Ishaq, R. (2025). Development of Verb Inflectional Complexity in Palestinian Arabic. *Journal of Speech, Language, and Hearing Research*, 68(3S), 1402–1423. https://doi.org/10.1044/2024_JSLHR-23-00722

- Heraña, J. C., & Azarias, R. A. (2024). Teaching strategies in mother-based multi-lingual education context: Inputs to the development of contextualized model. *International Journal of Research Studies in Education*, 13(5). <https://doi.org/10.5861/ijrse.2024.24035>
- Herzog-Punzenberger, B., Pichon, E. M. M., & Siarova, H. (2017). *Multilingual education in the light of diversity: Lessons learned: analytical report*. Publications Office. <https://data.europa.eu/doi/10.2766/71255>
- Hossain, K. I. (2024). Literature-based language learning: Challenges, and opportunities for English learners. *Ampersand*, 13, 100201. <https://doi.org/10.1016/j.amper.2024.100201>
- Hutson, J., Ellsworth, P., & Ellsworth, M. (2024). Preserving Linguistic Diversity in the Digital Age: A Scalable Model for Cultural Heritage Continuity. *Journal of Contemporary Language Research*, 3(1), 10–19. <https://doi.org/10.58803/jclr.v3i1.96>
- Ibno, N. (2025). Promoting Literacy Development in Early Elementary Years: The Case of Maimbung District. *Journal of Education and Academic Settings*, 2(1), 1–24. <https://doi.org/10.62596/dsnmr651>
- Iso, H., Uehara, Y., Ishigaki, T., Noji, H., Aramaki, E., Kobayashi, I., Miyao, Y., Okazaki, N., & Takamura, H. (2020). Learning to Select, Track, and Generate for Data-to-Text. *Journal of Natural Language Processing*, 27(3), 599–626. <https://doi.org/10.5715/jnlp.27.599>
- Ivushkina, T. (2024). Linguistic Diversity in the Philippines. *Linguistics & Polyglot Studies*, 10(3), 111–121. <https://doi.org/10.24833/2410-2423-2024-3-40-111-121>
- Jablonkai, R. R. (2020). Leveraging Professional Wordlists for Productive Vocabulary Knowledge. *ESP Today*, 8(1), 2–24. <https://doi.org/10.18485/esptoday.2020.8.1.1>
- Jackson, H. (Ed.). (2022). *The Bloomsbury Handbook of Lexicography* (1st ed.). Bloomsbury Publishing Plc. <https://doi.org/10.5040/9781350181731>
- Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy*, 5(4), 87. <https://doi.org/10.4103/0976-0105.141942>
- Kalenkovich, E., Shestakova, A., & Kazanina, N. (2022). Frequency tagging of syntactic structure or lexical properties; a registered MEG study. *Cortex*, 146, 24–38. <https://doi.org/10.1016/j.cortex.2021.09.012>
- Khaya, Y. (2025). Code-Switching and the Variable Grammar Hypothesis: Evidence from Moroccan Arabic–French Bilinguals. *Journal of Critical Studies in Language and Literature*, 6(4), 19–38. <https://doi.org/10.46809/jcsll.v6i4.366>
- Laine, M., Fellman, D., Eräste, T., Ritakallio, L., & Salmi, J. (2024). Strategy use and its evolvement in word list learning: A replication study. *Royal Society Open Science*, 11(2), 230651. <https://doi.org/10.1098/rsos.230651>
- Lamorinas, D. D., Bullado, A. J. B., & Lutian, M. J. M. (2024). Preserving initiatives for the Filipino language: Analysis of language protection through personal efforts among Filipino students. *Environment and Social Psychology*, 9(12). <https://doi.org/10.59429/esp.v9i12.3230>
- Laycock, H. (2006). *Words without Objects: Semantics, Ontology, and Logic for Non-Singularity* (1st ed.). Oxford University Press Oxford. <https://doi.org/10.1093/0199281718.001.0001>
- Lei, Y., & Reynolds, B. L. (2022). Learning English vocabulary from word cards: A research synthesis. *Frontiers in Psychology*, 13, 984211. <https://doi.org/10.3389/fpsyg.2022.984211>
- Lewis, A. A. (2024). Linguistic Features of French Guiana Creole: Analyzing Phonology, Morphology, and Syntax Compared to Standard French and Other Creoles. *International Journal of Research*

and Innovation in Social Science, VIII(VII), 3379–3393.
<https://doi.org/10.47772/IJRISS.2024.807258>

- Lobanova, S., Kovalchuk, O., Melnychuk, O., Zabiiaka, I., Alyieva, A., & Kanonik, N. (2023). Formation of Lexical Competence in Applicants for Education at Distance Learning (Experience of Foreign Scientists). *World Journal of English Language*, 13(4), 49.
<https://doi.org/10.5430/wjel.v13n4p49>
- Ludewig, U., Hübner, N., & Schroeder, S. (2023). Vocabulary, text coverage, word frequency and the lexical threshold in elementary school reading comprehension. *Reading and Writing*, 36(9), 2409–2431. <https://doi.org/10.1007/s11145-022-10385-0>
- Lumanlan, J., De La Torre, R., Asistio, F. V., De Guia, L., Do, H., Ramos, E. T., & Acuña, Ma. J. (2024). Perspectives on Mother Tongue-Based Multilingual Education Implementation in the Philippines. *Journal of Interdisciplinary Perspectives*, 2(12).
<https://doi.org/10.69569/jip.2024.0540>
- Madrunio, M. R., Martin, I. P., & Plata, S. M. (2016). English Language Education in the Philippines: Policies, Problems, and Prospects. In R. Kirkpatrick (Ed.), *English Language Education Policy in Asia* (Vol. 11, pp. 245–264). Springer International Publishing. https://doi.org/10.1007/978-3-319-22464-0_11
- Masaru, M. (1988). *THE HANUNOO-MANGYAN: Society, Religion and Law among a Mountain People of Mindoro Island, Philippines*. 国立民族学博物館. <https://doi.org/10.15021/00003241>
- Masorong, S. K. (2021). Traces and Roots: Exploring Lexical Rapport of the Bisayan and Tausug Languages. *International Journal of Linguistics, Literature and Translation*, 4(2), 120–128.
<https://doi.org/10.32996/ijllt.2021.4.2.15>
- McFarland, C. D. (2004). The Philippine language situation. *World Englishes*, 23(1), 59–75.
<https://doi.org/10.1111/j.1467-971X.2004.00335.x>
- Menson-Makalingkang, A. B., & Echavez, N. Q. (2021). *Phonological Adaptation of Arabic Loanwords in Maguindanaon*: International Conference on Sustainable Innovation Track Humanities Education and Social Sciences (ICSIHES 2021), Bantul, Indonesia.
<https://doi.org/10.2991/assehr.k.211227.010>
- Metila, R. A., D’Agostino, T., & Iwasaki, E. (2025). Why the Philippines reversed its mother-tongue instruction policy: Tracing the erosion of political priority for first language-based education in the Philippines. *International Journal of Educational Development*, 118, 103408.
<https://doi.org/10.1016/j.ijedudev.2025.103408>
- Miller, A. (2014). Introduction to Using Excel® Pivot Tables and Pivot Charts to Increase Efficiency in Library Data Analysis and Illustration. *Journal of Library Administration*, 54(2), 94–106.
<https://doi.org/10.1080/01930826.2014.903365>
- Mohamed, Y. A., Khanan, A., Bashir, M., Mohamed, A. H. H. M., Adiel, M. A. E., & Elsadig, M. A. (2024). The Impact of Artificial Intelligence on Language Translation: A Review. *IEEE Access*, 12, 25553–25579. <https://doi.org/10.1109/ACCESS.2024.3366802>
- Mtallo, G., & Kimambo, G. (2024). Sociolinguistic Factors Influencing the Acquisition of English Vocabulary in Multilingual Contexts: Insights from Selected Secondary School Teachers in Dodoma Tanzania. پژوهش‌های نوین در مطالعات زبان انگلیسی, Online First.
<https://doi.org/10.30479/jmrels.2024.19603.2292>
- Nambiar, R. M. K., Ibrahim, N., Hashim, R. S., Yasin, R. M., Azman, H., Yusof, N. Mohd., Ramli, R., & Mustafa, R. (2020). Impact of Local Culture-Based Reading Materials on Students’ Skill

- Development and Confidence in English. *Universal Journal of Educational Research*, 8(2), 445–453. <https://doi.org/10.13189/ujer.2020.080215>
- Nausa, R. (2019). Modality and Code Glosses to Transition from Academic Written to Oral Discourses. *Íkala*, 24(1), 51–67. <https://doi.org/10.17533/udea.ikala.v24n01a02>
- Nettle, D. (1998). Explaining Global Patterns of Language Diversity. *Journal of Anthropological Archaeology*, 17(4), 354–374. <https://doi.org/10.1006/jaar.1998.0328>
- Ngulube, P. (2023). Improving the quality of reporting findings using computer data analysis applications in educational research in context. *Heliyon*, 9(9), e19683. <https://doi.org/10.1016/j.heliyon.2023.e19683>
- Nicklin, C., Bailey, D., McLean, S., Kim, Y. A., Kang, H., & Vitta, J. P. (2025). Applying lexical sophistication models to wordlist development: A proof-of-concept study. *Research Methods in Applied Linguistics*, 4(1), 100175. <https://doi.org/10.1016/j.rmal.2024.100175>
- Norhana, M. A., Fatimah Mahmud, Lee, K. L., & Hussain, S. (2024). SYSTEMATIC LITERATURE REVIEW: LEAN ERGONOMICS ANALYSIS IN SMALL & MEDIUM ENTERPRISES (SMEs) MANUFACTURING SECTOR. *Journal of Technology and Operations Management*, 19(1), 89–108. <https://doi.org/10.32890/jtom2024.19.1.7>
- Nwachukwu, E. L., & Joseph, C. (2024). Bridging the gap between theory and implementation: National language policy and documentation of mother tongue. *Journal of Language and Pragmatics Studies*, 3(1), 22–30. <https://doi.org/10.58881/jlps.v3i1.45>
- O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19, 1609406919899220. <https://doi.org/10.1177/1609406919899220>
- Oluka, A. (2025). Phenomenological Research Strategy: Descriptive and Interpretive Approaches. *F1000Research*, 14, 725. <https://doi.org/10.12688/f1000research.166273.1>
- Ouattara, C. A. T., Tang, Y., Luo, S., Diallo, B. S., Okagbue, E. F., & Loum, O. J. A. (2025). Mother-tongue-mediated teaching and learning in primary education in Mali: Exploring the policy-action maze through a translanguaging pedagogy perspective. *International Multilingual Research Journal*, 1–16. <https://doi.org/10.1080/19313152.2025.2495438>
- Pakendorf, B., Dobrushina, N., & Khanina, O. (2021). A typology of small-scale multilingualism. *International Journal of Bilingualism*, 25(4), 835–859. <https://doi.org/10.1177/13670069211023137>
- Panaguítón, R. J., & Mantawil, B. (2024). *Technology and Mother Tongue: Use of Digital Materials in Mother Tongue Based-Multilingual Education*. <https://doi.org/10.5281/ZENODO.14061712>
- Periñán-Pascual, C. (2017). Bridging the gap within text-data analytics: A computer environment for data analysis in linguistic research. *Revista de Lenguas Para Fines Específicos*. <https://doi.org/10.20420/rlfe.2017.339>
- Petrova, E., Dewing, J., & Camilleri, M. (2016). Confidentiality in participatory research: Challenges from one study. *Nursing Ethics*, 23(4), 442–454. <https://doi.org/10.1177/0969733014564909>
- Ponce, A. R. C., & Lucas, R. I. G. (2024). Language attitudes in linguistically diverse contexts: Implications for mother tongue education programme in Central Mindanao, Philippines. *Journal of Multilingual and Multicultural Development*, 45(5), 1581–1595. <https://doi.org/10.1080/01434632.2021.2002341>

- Purcia, E. L., & Castante, N. S. (2023). Abrogating Mother-Tongue Multilingual Education in the Philippines: The Lens of MTB-MLE Teachers in Public Elementary Schools. *American Journal of Multidisciplinary Research and Innovation*, 2(4), 83–92. <https://doi.org/10.54536/ajmri.v2i4.1818>
- Radford, A. (2016). *Analysing English Sentences* (2nd ed.). Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/CBO9780511980312>
- Rajagukguk, R., Malems, T., & Ginting, M. (2022). Preserving Linguistic Diversity: Strategies for Language Preservation and Cultural Heritage Safeguarding. *Jurnal Ilmu Pendidikan Dan Humaniora*, 11(3), 221–236. <https://doi.org/10.35335/jiph.v11i3.25>
- Riazi, A. M., & Farsani, M. A. (2024). Mixed-methods research in applied linguistics: Charting the progress through the second decade of the twenty-first century. *Language Teaching*, 57(2), 143–182. <https://doi.org/10.1017/S0261444823000332>
- Rinehart, R. E., & Earl, K. (2016). Auto-, duo- and collaborative-ethnographies: “Caring” in an audit culture climate. *Qualitative Research Journal*, 16(3), 210–224. <https://doi.org/10.1108/QRJ-04-2016-0024>
- Roettger, T. B., & Baer-Henney, D. (2019). Toward a replication culture: Speech production research in the classroom. *Phonological Data and Analysis*, 1(4), 1–23. <https://doi.org/10.3765/pda.v1art4.13>
- Sales, A. J. G. (2022). Linguistic Borrowing of English Words and Utterances among Philippine’s Generation Z in Cebuano Visayan. *REiLA : Journal of Research and Innovation in Language*, 4(1), 41–53. <https://doi.org/10.31849/reila.v4i1.7559>
- Sarma, V. M. (2025). Intersections between heritage, multilingualism, and education: Language acquisition in India. *Frontiers in Human Neuroscience*, 19, 1538482. <https://doi.org/10.3389/fnhum.2025.1538482>
- Schultze-Berndt, E. (2015). 60. Language Documentation. In T. Kiss & A. Alexiadou (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 42/3* (pp. 2063–2095). DE GRUYTER. <https://doi.org/10.1515/9783110363685-020>
- Shaghaghi, N., Ghosh, S., & Kapoor, R. (2021). ClassRoute: An English to Punjabi Educational Video Translation Pipeline for Supporting Punjabi Mother-Tongue Education. *2021 IEEE Global Humanitarian Technology Conference (GHTC)*, 342–348. <https://doi.org/10.1109/GHTC53159.2021.9612485>
- Smith, J. A. (2017). Textual Analysis. In *The International Encyclopedia of Communication Research Methods(bll* (pp. 1–7). Wiley. <https://doi.org/10.1002/9781118901731.iecrm0248>
- Subiantoro, Ali, M. M., & Firdaus, F. A. (2023). REVITALIZATION OF LOCAL LANGUAGES THROUGH EDUCATIONAL CURRICULUM: STRATEGIES AND IMPLICATIONS. *IJoLaC: International Journal of Language and Culture*, 1(2), 61–67. <https://doi.org/10.63762/ijolac.v1i2.21>
- Tan, Y., & Jehom, W. J. (2024). The Function of Digital Technology in Minority Language Preservation: The Case of the Gyalrong Tibetan Language. *Preservation, Digital Technology & Culture*, 53(3), 165–177. <https://doi.org/10.1515/pdte-2024-0021>
- Tang, C., & Van Heuven, V. J. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119(5), 709–732. <https://doi.org/10.1016/j.lingua.2008.10.001>

- Taylor-Leech, K., & Benson, C. (2017). Language planning and development aid: The (in)visibility of language in development aid discourse. *Current Issues in Language Planning*, 18(4), 339–355. <https://doi.org/10.1080/14664208.2017.1360690>
- Tegegne, W. (2015). The Use of Dialects in Education and Its Impacts on Students' Learning and Achievements. *Education Journal*, 4(5), 263. <https://doi.org/10.11648/j.edu.20150405.22>
- Tenorio, A. D. (2024). When language gets into the equation: Mother tongue-based multilingual education (MTB-MLE) policy appropriation in elementary mathematics instruction. *Journal of Multilingual and Multicultural Development*, 45(6), 2064–2077. <https://doi.org/10.1080/01434632.2022.2039673>
- Thieberger, N. (2016). What Remains to be Done—Exposing Invisible Collections in the other 7,000 Languages and Why it is a DH Enterprise. *Digital Scholarship in the Humanities*, fqw006. <https://doi.org/10.1093/llc/fqw006>
- Thieberger, N. (2020). Technology in Support of Languages of The Pacific: Neo-Colonial or Post-Colonial? *ASIAN-EUROPEAN MUSIC RESEARCH JOURNAL*, 5, 17–24. <https://doi.org/10.30819/aemr.5-3>
- Thieberger, N., & Berez, A. L. (2011). *Linguistic Data Management*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199571888.013.0005>
- Tol, R. S. J. (2014). Quantifying the consensus on anthropogenic global warming in the literature: A re-analysis. *Energy Policy*, 73, 701–705. <https://doi.org/10.1016/j.enpol.2014.04.045>
- Tong, J., Che Abdul Rahman, A. N., & Hamat, A. (2025). A systematic literature review on word selection criteria in corpus-based wordlists development. *Cogent Arts & Humanities*, 12(1), 2432131. <https://doi.org/10.1080/23311983.2024.2432131>
- Trudell, B. (2009). Local-language literacy and sustainable development in Africa. *International Journal of Educational Development*, 29(1), 73–79. <https://doi.org/10.1016/j.ijedudev.2008.07.002>
- Tsujii, J. (2021). Natural Language Processing and Computational Linguistics. *Computational Linguistics*, 1–21. https://doi.org/10.1162/coli_a_00420
- Tupas, R. (2015). Inequalities of multilingualism: Challenges to mother tongue-based multilingual education. *Language and Education*, 29(2), 112–124. <https://doi.org/10.1080/09500782.2014.977295>
- Tursunovich, R., Abdurakhimovna, R., Bakhtiyarovna, Y., & Allovidinovna, I. (2024). Lexical Problems Encountered in The Translation Of Small Genres. In A. Kusuma Wardana (Ed.), *Proceedings of the UPY-ICCCM International Conference on Education and Social Science (UPINCESS 2024)* (Vol. 895, pp. 80–90). Atlantis Press SARL. https://doi.org/10.2991/978-2-38476-338-2_11
- UNESCO, 2025. (2025). *Languages matter: Global guidance on multilingual education*. UNESCO. <https://doi.org/10.54675/MLIO7101>
- Urbanczyk, S. (2017). Phonological and Morphological Aspects of Reduplication. In S. Urbanczyk, *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.80>
- Vamarasi, M. (2014). The creation of learner-centred dictionaries for endangered languages: A Rotuman example. *Journal of Multilingual and Multicultural Development*, 35(2), 167–178. <https://doi.org/10.1080/01434632.2013.825266>

- Velasco, Y. P. (2025). The Interplay Between Language Ideologies and Mother Tongue-Based Multilingual Education (MTBMLE) Policy Implementation in the Philippines. *The Asia-Pacific Education Researcher*, 34(2), 765–775. <https://doi.org/10.1007/s40299-024-00894-7>
- Wan Jusoh, W. N. A. S., Jobar, N. A., Md Yusoff, M. Z. N., & Mahat, H. (2024). Exploring corpus linguistics via computational tool analysis: Key finding review. *Indonesian Journal of Electrical Engineering and Computer Science*, 34(2), 1052. <https://doi.org/10.11591/ijeecs.v34.i2.pp1052-1062>
- Wangdi, T., & Shimray, R. (2022). Investigating the Significance of Coxhead's Academic Word List for Self-Access Learners. *Studies in Self-Access Learning Journal*, 13(3), 367–388. <https://doi.org/10.37237/130305>
- Wichmann, S. (2020). How to Distinguish Languages and Dialects. *Computational Linguistics*, 45(4), 823–831. https://doi.org/10.1162/coli_a_00366
- Willig, C., & Rogers, W. S. (2017). *The SAGE Handbook of Qualitative Research in Psychology*. SAGE Publications Ltd. <https://doi.org/10.4135/9781526405555>
- Woolard, K. A. (2020). Language Ideology. In J. Stanlaw (Ed.), *The International Encyclopedia of Linguistic Anthropology* (1st ed., pp. 1–21). Wiley. <https://doi.org/10.1002/9781118786093.iela0217>
- Yousefi, M. H., & Biria, R. (2018). The effectiveness of L2 vocabulary instruction: A meta-analysis. *Asian-Pacific Journal of Second and Foreign Language Education*, 3(1), 21. <https://doi.org/10.1186/s40862-018-0062-2>
- Zhang, Y. (2024). Enhancing Vocabulary Acquisition and Reading Comprehension in Intermediate Chinese: A Corpus-Driven Approach for International Students. *Journal of E-Learning Research*, 3(1), 45–60. <https://doi.org/10.33422/jelr.v3i1.755>
- Zorc, R. D. (1974). Towards a Definitive Philippine Wordlist: The Qualitative Use of Vocabulary in Identifying and Classifying Languages. *Oceanic Linguistics*, 13(1/2), 409. <https://doi.org/10.2307/3622750>
- Zorc, R. D. (1977). *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Department of Linguistics, Research School of Pacific Studies, Australian National University.

About the Contributor(s)

John Gerald Arbias Pilar, PhD, is an Associate Professor of Carlos Hilado Memorial State University, Negros Occidental. He serves as the research coordinator at Carlos Hilado Memorial State University-Fortune Towne campus. He is also an associate member of the National Research Council of the Philippines (NRCP). His research interests include sociolinguistics, local language documentation and description, forensic linguistics, and educational pedagogy. He has developed various intellectual property rights related to the documentation and analysis of local dialects.

Email: john.pilar@chmsu.edu.ph

ORCID: <http://orcid0009-0007-6819-6432>

Publisher's Note: The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.
