

Future Pathways in Malay Philological Inquiry

Mardiah Mawar Kembaren¹, M.Husnan Lubis²

^{1,2} Universitas Sumatera Utara, Indonesia

Abstract

This article examines future directions for Malay philological inquiry by addressing the gap between manuscript digitisation and the production of linguistically meaningful knowledge. It aims to propose a critical digital Malay philology model that moves beyond preservation toward corpus-based, Jawi-sensitive, and ethically grounded interpretation. Using a multilayered mixed-method design, the study synthesises philological interpretation, quantitative corpus mapping, and triangulation of recent manuscript digitisation data, Jawi literacy research, and Malay linguistic scholarship. The findings show that DREAMSEA's 2024 documentation of more than 571,584 manuscript images from over 8,570 manuscripts across 168 owners and 57 regions provides a substantial empirical foundation for corpus-oriented philology. The 2025 inclusion of eighteen manuscripts, including eleven Malay manuscripts, in the Royal Asiatic Society Digital Library further confirms the transnational circulation of Malay textual heritage. However, digitised images generate philological knowledge only when transformed through transliteration, orthographic annotation, genre classification, semantic tagging, and discourse interpretation. The study also finds that Jawi spelling variation affects readability, searchability, and historical interpretation. The proposed model contributes by integrating classical philology, corpus linguistics, historical semantics, sociolinguistics, digital humanities, and community collaboration to reframe Malay manuscripts as living archives of language, law, religion, diplomacy, and civilisation.

ARTICLE HISTORY

Received : 30 August 2025

Revised : 2 November 2025

Accepted : 20 December 2025

KEYWORDS

Corpus Linguistics; Digital Philology; Jawi Manuscripts; Malay Philology; Manuscript Digitisation.

PUBLISHER'S NOTE

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license



CORRESPONDING AUTHOR

Mardiah Mawar Kembaren, Universitas Sumatera Utara, Indonesia. Email: mardiah@usu.ac.id

Introduction

Malay philology is now at a decisive epistemic crossroads. For more than a century, Malay philology has largely been concerned with the search, collation, editing, and interpretation of Jawi sources, including manuscripts written in various forms of Jawi or transliterated into Romanised Malay, within manuscript traditions shared across the wider Malay-Indonesian archipelago. However, between 2020 and 2025, at least three intersecting developments have reshaped the direction of the field: the massive digitisation of manuscripts, the renewed scholarly interest in endangered scripts, and the imperative to read manuscripts not only as literary remains but also as records of law, theology, medicine, diplomacy, commerce, ecological knowledge, political imagination, and civilisational memory (Amanda et al., 2025; Anjani & Siregar, 2023; Siregar et al., 2023). For example, the British Library digitised its Malay manuscripts, and around 120 of them are now freely available online, allowing global access to materials that could previously be consulted only in elite institutional reading rooms (Gallop, 2021). Likewise, in 2025, the Royal Asiatic Society contributed eighteen manuscripts to its Digital Library, eleven of which were Malay manuscripts, indicating that Malay manuscript studies now operate within a global and transnational digital infrastructure (Royal Asiatic Society, 2025).

The theoretical foundation of this study lies at the intersection of classical philology, historical linguistics, corpus linguistics, and critical digital humanities. In its classical sense, philology is concerned with textual transmission, variant readings, scribal practice, script forms, and semantic continuity (Fajriah et al., 2025; Gallop, 2021; Sari Tri Anjani & Iskandarsyah Siregar, 2023). Yet contemporary Malay philological inquiry can no longer be confined to textual recovery and editorial reconstruction alone. Malay philology now needs to be understood as a science of textual mobilities: how words, orthographies, genres, legal-theological vocabularies, and cultural concepts move across space, time, script, institution, and community (Burhanudin, 2022; Daneshgar, 2022, 2023). This perspective is consistent with recent digital philology scholarship, which does not treat manuscripts merely as static artefacts to be preserved, but as structured data environments through which orthographic variation, textual variants, and historical meaning can be examined through purposefully curated digital systems (Syahrullah et al., 2024).

The expansion of manuscript digitisation has provided the material infrastructure for this new direction. DREAMSEA's manuscript preservation work, including the transfer of eight years of digitisation outputs to the National Library of Indonesia in 2024, demonstrates that Southeast Asian manuscript preservation is entering an institutional and transnational phase (DREAMSEA, 2024a; Shofiyanti & Rahman, 2026). DREAMSEA's focus on rare and endangered manuscripts is important because it protects fragile collections while also expanding scholarly and public accessibility to manuscript heritage (DREAMSEA, 2024a). However, access alone does not automatically produce linguistic knowledge. Digital images of manuscripts must be converted into annotated corpora through transliteration, orthographic annotation, genre classification, semantic tagging, and discourse interpretation before they can support systematic analysis of spelling variation, lexical change, textual transmission, semantic fields, and intertextual networks (Korompot et al., 2024; Pangaribuan et al., 2025; Rachmawati et al., 2025).

This is the central problem addressed in this article: the availability of digital copies has not yet produced a sufficiently new theory of Malay philology. Many digitisation and preservation initiatives have successfully expanded access to manuscript materials, but linguistic analysis particularly historical semantics, discourse formation, orthographic variation, intertextual transmission, and sociolinguistic mobility remains insufficiently integrated into the study of digital Malay manuscripts (Iskandarsyah Siregar & Samsur Rijal Yahaya, 2023; Wahyuni et al., 2024, 2025). The large-scale documentation of manuscript images and metadata is therefore both an opportunity and a methodological challenge. Manuscript images can be stored, mapped, and displayed, but without transliteration, linguistic annotation, and contextual interpretation, they remain limited as research corpora. Consequently, the field requires a framework that moves beyond preservation toward the production of philological knowledge (Wardani et al., 2025).

A further methodological issue concerns Jawi. Recent research on Jawi spelling patterns shows that Malay functions as a synchronic digraphic language and that inconsistent spelling patterns can affect readability, searchability, and language processing (Salehuddin et al., 2025). This finding has direct implications for Malay philology because Jawi manuscripts are not merely literary or historical artefacts; they are also archival evidence of script ecology, orthographic instability, scribal practice, and language ideology. Treating Jawi simply as a source to be mechanically converted into Rumi risks erasing historically meaningful variation. Therefore, any future model of Malay philology must preserve original Jawi forms while also providing normalised transliteration

layers for computational searching and comparative linguistic analysis (Bania & Akob, 2025; Dewi et al., 2023; Hizbullah et al., 2019; Kalsum & Rachmadini, 2023).

Recent Malay linguistic scholarship also indicates that the field has become increasingly diverse, interdisciplinary, and methodologically active. Bibliometric studies of Malay linguistics reveal expanding themes, author networks, and methodological approaches, including studies of Malay communities, dialectal variation, identity formation, structural linguistics, and applied language research (Yaqin et al., 2024). However, this development has not yet been fully connected to the growing archive of digitised Malay manuscripts. The gap, therefore, is not the absence of manuscript preservation or Malay linguistic research, but the discontinuity between manuscript preservation, linguistic theory, corpus methodology, and sociopolitical interpretation. This article responds to that gap by proposing that digitised Malay manuscripts should be developed into searchable, annotated, and ethically contextualised linguistic corpora.

This study aims to establish future trajectories in Malay philological studies through a linguistically informed, digitally facilitated, and ethically grounded framework. More specifically, it seeks to investigate Malay manuscripts as historical-linguistic corpora; identify the limitations of traditional models of textual editing; connect philological work with corpus linguistics, discourse analysis, semiotics, and legal-cultural history; and propose a forward-looking model of collaborative manuscript reading. The main research problems are: what kind of linguistic knowledge should Malay philology pursue in order to move beyond preservation toward knowledge production; how do methodological limitations prevent digitised manuscripts from becoming reliable research corpora; and how can community ownership, digital humanities, and critical linguistic paradigms be embedded synergistically in future Malay philological inquiry?

This article introduces the concept of “critical digital Malay philology” as its main contribution. Unlike earlier studies that emphasise manuscript preservation, access, catalogue development, textual editing, or literary-historical interpretation, this article argues that the next phase of Malay philology must integrate classical philological rigour with corpus annotation, Jawi-sensitive transliteration, historical semantics, sociolinguistics, discourse analysis, semiotics, digital humanities, and ethical community collaboration. Digitisation itself is not new; what is proposed here is a linguistically grounded, digitally supported, and socially accountable science of textual civilisation. In this model, Malay manuscripts are understood not only as preserved objects but as deeply historicised linguistic events shaped by scribes, patrons, religious networks, colonial encounters, institutional repositories, and twenty-first-century digital infrastructures.

This study is limited to archipelagic manuscript traditions that are relevant to language history, knowledge transmission, law, religion, diplomacy, and cultural memory at the intersection between manuscript culture and digital remediation. It hypothesises that digitised Malay manuscripts will gain greater academic value when they are transformed into linguistically searchable and contextually annotated corpora; that philological analysis of Jawi-encoded texts can reveal patterns of intellectual and linguistic mobility that conventional literary-historical approaches may overlook; and that ethical collaboration with custodians of manuscript heritage can produce more sustainable and scientifically defensible forms of philological knowledge than institutionalised digitisation practices alone.

Method

This study employed a multilayered mixed-method design that combines qualitative philological interpretation, quantitative corpus mapping, and triangulation (Guenduez et al., 2025; Maeir, 2025; Tang & Li, 2024). The design was selected because the study aims not only to interpret Malay manuscripts as textual and cultural artefacts but also to examine how digitised manuscript materials can be transformed into searchable, annotated, and linguistically meaningful corpora. The qualitative component focused on textual transmission, Jawi orthographic variation, semantic change, genre circulation, and discourse formation, while the quantitative component was used to map numerical evidence from recent manuscript digitisation, Jawi literacy research, and Malay linguistic bibliometric studies (DREAMSEA, 2024a; Royal Asiatic Society, 2025; Salehuddin et al., 2025; Yaqin et al., 2024).

The data sources consisted of primary and secondary documentary materials. The primary data included digitised Malay manuscripts, manuscript catalogues, metadata records, repository descriptions, and institutional reports from credible manuscript-holding bodies, particularly DREAMSEA, the National Library of Indonesia, and the Royal Asiatic Society Digital Library. DREAMSEA's digitisation programme was used as one of the main empirical anchors because it documents large-scale manuscript preservation and access initiatives in Southeast Asia (DREAMSEA, 2024b, 2024a). The Royal Asiatic Society Digital Library was also used because its 2025 additions demonstrate the continued circulation of Malay manuscript heritage within international archival networks (Royal Asiatic Society, 2025). Secondary data included recent scholarly studies on Malay philology, Jawi literacy, digital philology, manuscript preservation, corpus linguistics, historical semantics, sociolinguistics, and digital humanities.

The selection of sources was conducted purposively based on four criteria. First, the source had to be directly related to Malay manuscripts, Jawi writing, manuscript digitisation, corpus-based philology, or Malay linguistic scholarship. Second, the source had to come from a credible academic, institutional, or repository-based publication. Third, the source had to provide verifiable information, such as metadata, publication details, DOI, repository records, or institutional documentation. Fourth, the source had to contribute to the construction of a future-oriented model of Malay philological inquiry. Sources that were not directly related to Malay philology, lacked clear metadata, or did not contribute to the conceptual and methodological focus of the study were excluded.

Data analysis was conducted in four stages. First, relevant manuscript and repository data were classified according to script, genre, approximate period, region, subject matter, and textual function. Second, selected textual and documentary information was organised into analytical categories, including Jawi spelling variation, lexical items, semantic fields, discourse markers, legal-religious terminology, formulaic expressions, and intertextual traces. Third, qualitative interpretation was applied to examine how these categories reflect historical semantics, sociolinguistic mobility, genre circulation, and cultural meaning. Fourth, quantitative mapping was used to interpret numerical data from digitisation records, Jawi literacy findings, and bibliometric evidence, including manuscript totals, metadata entries, regional distribution, manuscript ownership, participant-based Jawi readability findings, and publication trends.

To ensure credibility and analytical consistency, this study applied three forms of triangulation. Data triangulation was conducted by comparing manuscript data, repository records, institutional reports, and previous scholarly studies. Methodological triangulation was applied by

integrating philological reading, corpus-oriented classification, quantitative mapping, and discourse interpretation. Theoretical triangulation was used by combining perspectives from classical philology, historical linguistics, sociolinguistics, corpus linguistics, semiotics, and critical digital humanities. Through these procedures, the study offers a replicable and ethically accountable framework for developing critical digital Malay philology as a future direction for Malay manuscript studies.

Results and Discussion

Result

The findings indicate that Malay philological inquiry is undergoing a structural transition from manuscript preservation toward the development of linguistically analysable knowledge systems. Four main findings emerged from the qualitative, quantitative, and triangulated analysis: the rapid expansion of digitised manuscript repositories, the transnational circulation of Malay textual heritage, the methodological significance of Jawi orthography, and the growing but still weakly integrated body of Malay linguistic scholarship.

The first finding concerns the scale of manuscript digitisation. The DREAMSEA 2024 manuscript dataset documented more than 571,584 high-resolution manuscript images, 8,570 manuscripts, 8,470 metadata entries, 168 manuscript owners, and 57 regions covered during the digitisation process from 2018 to 2024. These figures show that Malay and Nusantara manuscript studies now have a substantial empirical foundation for corpus-oriented philological research. On average, the dataset contains approximately 66.70 images per manuscript and 51.01 manuscripts per owner. This scale indicates that Malay manuscript studies can be developed through broader comparison across script, genre, region, ownership, metadata, and textual function.

The second finding relates to the international circulation of Malay manuscript heritage. In 2025, the Royal Asiatic Society added eighteen manuscripts to its Digital Library, consisting of eleven Malay manuscripts and seven Javanese manuscripts. Malay manuscripts represented 61.10% of the new additions, while Javanese manuscripts represented 38.90%. This evidence shows that Malay philological materials are not confined to local or national repositories, but are distributed across transnational archival networks involving Southeast Asian, British, and other international institutions. The inclusion of Malay manuscripts in global digital repositories also demonstrates the need for comparative and cross-repository approaches in future Malay philological research.

The third finding concerns Jawi as a methodological variable in Malay philology. Recent Jawi literacy data show that Malay functions as a synchronic digraphic language written in both Rumi and Jawi, and that spelling variation affects reading accuracy and processing. In the Jawi spelling study, thirty native Malay participants were involved in three experiments using audio and visual prompts to examine Jawi spelling patterns. The findings showed that spelling conditions influenced correct responses and reaction time. In one experiment, the average correct response was 48.60%, indicating difficulty in recognising certain Jawi spelling patterns. The effect of spelling type on correct response was significant, $F(3, 87) = 8.78, p = .001$. The main effect of data type was also significant, $F(3, 87) = 5.97, p = .004, \eta^2 = .171$. Reading accuracy increased from 71.00% for vowel-letter spelling to 85.00% when vowel diacritics were used. These findings show that Jawi spelling variation affects readability, searchability, and historical interpretation.

The fourth finding is drawn from bibliometric evidence on Malay linguistic studies. Over a twenty-year period, 724 publications on Malay linguistics were identified, with the highest number

of publications appearing in 2023, reaching 91 publications. This peak represents approximately 12.60% of the total two-decade dataset and is around 2.50 times higher than the simple annual average of 36.20 publications. The bibliometric data indicate that Malay linguistics is an active and expanding field. However, the dominant themes remain Malay communities, dialectal variation, identity formation, structural linguistics, and language application, while the connection between Malay linguistics and digitised manuscript-based corpus philology remains underdeveloped.

Table 1. Empirical Evidence and Philological Implications

Empirical anchor	Numerical result	Main finding	Philological implication
DREAMSEA manuscript images	571,584+	Large-scale digitised visual archive	Requires image-to-text processing, transliteration, and corpus infrastructure
DREAMSEA manuscripts	8,570	Substantial manuscript base	Enables comparative analysis of genre, script, region, and textual function
DREAMSEA metadata entries	8,470	High metadata coverage	Requires linguistic enrichment for searchable philological analysis
DREAMSEA owners	168	Community-based manuscript ecology	Supports ethical and collaborative manuscript research
DREAMSEA regions	57	Broad regional distribution	Enables spatial philology and regional linguistic comparison
RAS 2025 additions	18 manuscripts	Continued international digitisation	Confirms the global archival circulation of Malay textual heritage
RAS Malay manuscripts	11/18 = 61.10%	Malay manuscripts dominate the new additions	Shows the centrality of Malay texts in maritime Southeast Asian collections
RAS Javanese manuscripts	7/18 = 38.90%	Comparative archipelagic archive	Supports comparative Malay-Javanese and regional manuscript studies
Jawi literacy data	30 participants; 3 experiments	Jawi spelling affects readability and processing	Confirms the need for Jawi-sensitive transliteration and annotation
Malay linguistics bibliometrics	724 publications; peak of 91 in 2023	Expanding field of Malay linguistic studies	Indicates the need to connect Malay linguistics with manuscript-based corpus philology

Taken together, the findings show that the future of Malay philological inquiry depends on the transformation of digitised manuscript images into linguistically searchable, annotated, and contextually interpretable corpora. The DREAMSEA data provide the macro-scale manuscript base, the Royal Asiatic Society data confirm the transnational circulation of Malay textual heritage, the Jawi data demonstrate the methodological complexity of script and orthography, and the bibliometric evidence shows that Malay linguistics is growing but has not yet been fully connected to digitised manuscript corpora.

Discussion

The findings demonstrate that digitisation has significantly expanded access to Malay manuscripts, but access alone does not automatically produce philological knowledge. DREAMSEA's large-scale documentation provides an important empirical foundation for future research, yet manuscript images remain limited as linguistic data unless they are processed through transliteration, orthographic annotation, genre classification, semantic tagging, and discourse interpretation. This supports the central argument of the study: the future of Malay philology should not stop at preservation, but must move toward corpus-based and linguistically accountable knowledge production. In this regard, the study extends preservation-oriented initiatives by asking not only how many manuscripts have been digitised, but also how digitised manuscripts can be made searchable, comparable, replicable, and interpretable.

This interpretation is consistent with recent developments in digital philology, which treat manuscripts not merely as preserved artefacts but as structured data environments for examining textual variation, orthography, and historical meaning (Syahrullah et al., 2024). However, the present study expands that perspective by locating Malay manuscripts within a more specific methodological problem: the need to integrate Jawi-aware transliteration, corpus annotation, historical semantics, and ethical collaboration with manuscript-owning communities. In this framework, digital philology is not only a technical process of converting manuscripts into digital images, but also a critical process of transforming manuscript evidence into linguistic and cultural knowledge.

The RAS data further indicate that Malay textual heritage circulates across transnational archival systems. The addition of eleven Malay manuscripts to the Royal Asiatic Society Digital Library in 2025 shows that Malay philological inquiry cannot be approached solely through national or local frameworks (Royal Asiatic Society, 2025). Malay manuscripts are historically mobile objects shaped by maritime networks, Islamic intellectual circulation, colonial collecting practices, private custodianship, and contemporary digital repositories. Therefore, future Malay philology requires a comparative and spatial approach that can trace how texts, scripts, lexical forms, genres, and knowledge systems move across regions and institutions.

The findings on Jawi reinforce the need to treat script as a central methodological concern. Salehuddin et al. (2025) show that Jawi spelling patterns influence reading accuracy and processing, which has direct implications for philological work. If Jawi forms are mechanically normalised into Rumi without preserving orthographic variation, historically meaningful evidence may be lost. Conversely, if original Jawi variation is preserved without a searchable transliteration layer, the corpus may become difficult to analyse comparatively. For this reason, future Malay philology should adopt a two-layered corpus model: one layer preserves original Jawi forms, while another provides normalised transliteration for search, comparison, and linguistic mapping. This model allows philologists to preserve textual authenticity while also enabling computational and comparative analysis.

The bibliometric evidence also clarifies the gap between Malay linguistic studies and manuscript-based philology. Yaqin et al. (2024) show that Malay linguistics has grown considerably, especially in studies of Malay communities, dialectal variation, identity, structure, and applied linguistics. However, this growth has not yet been strongly integrated with the expanding archive of digitised manuscripts. This study therefore argues that Malay philology and Malay linguistics should not remain separate scholarly trajectories. Manuscripts provide historical linguistic evidence, while

contemporary linguistic methods provide tools for analysing lexical change, semantic fields, discourse patterns, orthographic variation, and sociolinguistic mobility. Integrating these two areas can strengthen both philological interpretation and Malay linguistic scholarship.

The novelty of this study lies in the proposed model of critical digital Malay philology. Unlike previous approaches that focus primarily on manuscript preservation, catalogue development, access, textual editing, or literary-historical interpretation, this model integrates classical philology, corpus linguistics, Jawi-sensitive annotation, historical semantics, sociolinguistics, discourse analysis, semiotics, digital humanities, and ethical community collaboration. The contribution of this model is not the claim that digitisation itself is new, but that digitised Malay manuscripts require a more rigorous interpretive framework before they can function as reliable evidence for language, law, religion, diplomacy, knowledge transmission, and civilisation. In this sense, the model reframes Malay manuscripts as living archives of linguistic and cultural action rather than static textual objects.

The theoretical implication of this study is that Malay philology needs to be repositioned as an interdisciplinary field that connects textual scholarship with linguistic theory and digital humanities. Classical philology remains essential because it provides careful attention to textual transmission, variant readings, scribal practice, and historical context. However, classical philology must now be expanded through corpus-based methods that allow researchers to compare larger sets of manuscripts across script, genre, region, and semantic category. This expansion does not replace traditional philological expertise; rather, it strengthens it by providing more systematic and replicable ways to interpret manuscript evidence.

The practical implication is that manuscript digitisation projects should be designed beyond image preservation. Future projects should include metadata enrichment, diplomatic transcription, normalised transliteration, Jawi orthographic tagging, genre classification, semantic annotation, and documentation of community ownership. Such procedures would allow manuscripts to be searched, compared, and reanalysed by different researchers. This approach is also important for education and public access because it can help non-specialist users engage with Malay manuscript heritage without erasing the complexity of Jawi and historical language variation.

The study also has ethical implications. Since many manuscripts originate from private owners, local communities, pesantren, royal collections, and family archives, philological research should not treat digitised manuscripts merely as institutional data. Ethical collaboration with custodians is necessary to ensure that digitisation, interpretation, and publication respect community ownership, cultural sensitivity, and local knowledge. Critical digital Malay philology therefore requires not only technical competence and linguistic expertise, but also social accountability.

This study has several limitations. First, it develops a conceptual and methodological model rather than producing a full diplomatic edition or complete annotated corpus of a specific manuscript collection. Second, the analysis relies on selected institutional data, previous Jawi literacy findings, and bibliometric evidence; therefore, the proposed model still needs to be tested on a defined manuscript corpus. Third, the study does not yet provide a detailed computational workflow for corpus construction, such as software selection, annotation schema, or database architecture. Future studies should apply the proposed model to a specific set of Malay manuscripts, test a two-layered Jawi-Rumi corpus system, compare orthographic variation across regions and genres, and develop a

replicable digital workflow for Malay manuscript annotation. Through such work, critical digital Malay philology can move from a conceptual proposal toward an empirically tested research practice.

Conclusion

This study concludes that the future of Malay philological inquiry depends on its transition from manuscript preservation toward linguistically grounded, digitally mediated, and ethically accountable knowledge production. The findings show that large-scale digitisation initiatives, such as DREAMSEA's manuscript documentation and the Royal Asiatic Society's digital collection, have created a substantial empirical foundation for corpus-oriented Malay philology. However, digitised images and metadata can only become meaningful philological evidence when processed through transliteration, orthographic annotation, genre classification, semantic tagging, and historical interpretation. The study also highlights Jawi as a crucial methodological variable because spelling variation affects readability, searchability, and interpretation. The proposed model of critical digital Malay philology contributes by integrating classical philology, corpus linguistics, historical semantics, sociolinguistics, digital humanities, and ethical community collaboration. This study is limited by its conceptual and documentary scope, as it does not yet test the model on a fully annotated manuscript corpus. Future research should empirically apply and validate this model across broader manuscript collections, linguistic datasets, and digital philology platforms to strengthen its methodological reliability and scholarly contribution..

References

- Amanda, A. P., Siregar, I., & Rachmawati, K. (2025). Perubahan makna kata pada berita olahraga di laman Sindonews.com. *Hortatori: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 9(1), 112–120. <https://doi.org/10.30998/jh.v9i1.3683>
- Anjani, S. D., & Siregar, I. (2023). Health vitality of the Betawi language in the future in Jakarta: A sociolinguistic study. *Formosa Journal of Sustainable Research*, 2(3), 623–640. <https://doi.org/10.55927/fjsr.v2i3.3521>
- Anjani, S. T., & Siregar, I. (2023). The existence of Palang Pintu culture in the opening procession of Betawi traditional weddings: Case study: George Herbert Mead's symbolic interactionism. *Formosa Journal of Sustainable Research*, 2(3), 641–666. <https://doi.org/10.55927/fjsr.v2i3.3517>
- Bania, A. S., & Akob, B. (2025). Preserving the Jawi script in Aceh: Assessing literacy, cultural heritage, and modern paradigm challenges. *Studies in English Language and Education*, 12(1), 457–470. <https://doi.org/10.24815/siele.v12i1.36629>
- Burhanudin, J. (2022). Two Islamic writing traditions in Southeast Asia: Kitab Jawi and Kitab Kuning with reference to the works of Da'ud al-Fatani and Nawawi al-Bantani. *Al-Jami'ah: Journal of Islamic Studies*, 60(1), 1–28. <https://doi.org/10.14421/ajis.2022.601.1-28>
- Daneshgar, M. (2022). A very old Malay Islamic manuscript. *Indonesia and the Malay World*, 50(147), 161–172. <https://doi.org/10.1080/13639811.2021.1975968>
- Daneshgar, M. (2023). Persianate Fiqh in Indonesia: Majmu'ih-yi Khani as a rare legal manuscript in a cosmopolitan context. *International Journal of Islam in Asia*, 2(2), 144–169. <https://doi.org/10.1163/25899996-20223006>
- Dewi, E., Munthe, B., Yasin, A., & Afriza, A. (2023). The contestation of the Malay Arabic and Roman literature: Shifts in culture, literature, and religion. *Arabiyatuna: Jurnal Bahasa Arab*, 7(1), 1. <https://doi.org/10.29240/jba.v7i1.5654>
- DREAMSEA. (2024a, October 1). *Empowering scholarship: Insights on Malay manuscript preservation from DREAMSEA*. <https://dreamsea.co/empowering-scholarship-insights-on-malay-manuscript-preservation-from-dreamsea/>

- DREAMSEA. (2024b, August 9). *Reviving history: DREAMSEA's digital manuscript grant to Indonesia's National Library*. <https://dreamsea.co/reviving-history-dreamseas-digital-manuscript-grant-to-indonesias-national-library/>
- Fajriah, F. N., Siregar, I., & Susanto, A. (2025). Implikatur variasi bahasa dan pengaruh gender pada komentar akun media sosial X @Collegemfs: Kajian sosiopragmatik. *Pena Literasi: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 8(2), 164–175.
- Gallop, A. T. (2021, November 28). *Who reads Malay manuscripts? The impact of the Bollinger digitisation project at the British Library*. IFLA. <https://www.ifla.org/news/who-reads-malay-manuscripts-the-impact-of-the-bollinger-digitisation-project-at-the-british-library/>
- Guenduez, A. A., Fuchs, S., Mergel, I., Mettler, T., & Frowein, S. (2025). From patterns to meaning: A mixed-methods framework that integrates computational and qualitative text analysis. *Quality & Quantity*, 60(1), 2713–2738. <https://doi.org/10.1007/s11135-025-02357-7>
- Hizbullah, N., Suryaningsih, I., & Mardiah, Z. (2019). Manuskrip Arab di Nusantara dalam tinjauan linguistik korpus. *Arabi: Journal of Arabic Studies*, 4(1), 65. <https://doi.org/10.24865/ajas.v4i1.145>
- Kalsum, N. U., & Rachmadini, N. (2023). The threat of discontinuity ancient scripts in the Islamic-Malay manuscript. *Arabiyatuna: Jurnal Bahasa Arab*, 7(1), 363. <https://doi.org/10.29240/jba.v7i1.7327>
- Korompot, C. A., Siregar, I., Khursanov, N. I., Abdullaev, D., & Mohamed, K. (2024). Investigating gender DIF in the reading comprehension section of the B2 First Exam. *International Journal of Language Testing*, 14(2), 57–66. <https://doi.org/10.22034/ijlt.2023.421011.1301>
- Maeir, N. (2025). Material philology and Syriac excerpting practices: A computational-quantitative study of the digitized catalog of the Syriac manuscripts in the British Library. *PLOS ONE*, 20(3), Article e0320265. <https://doi.org/10.1371/journal.pone.0320265>
- Pangaribuan, R. E., Siregar, I., & Susanto, A. (2025). Tindak tutur ekspresif dalam serial film *Nanti Kita Cerita tentang Hari Ini* pada YouTube kanal Toyota Indonesia: Kajian pragmatik. *J-Symbol: Jurnal Magister Pendidikan Bahasa dan Sastra Indonesia*, 13(1), 208–225. <https://doi.org/10.23960/symbol.v13i1.460>
- Rachmawati, C. R., Sosrohadi, S., & Siregar, I. (2025). Analisis semiotika dan retorik dalam sampul majalah *Tempo* edisi “Tentakel Judi Kamboja.” *Triangulasi: Jurnal Pendidikan Kebahasaan, Kesastraan, dan Pembelajaran*, 5(2), 93–99. <https://doi.org/10.55215/triangulasi.v5i2.14>
- Royal Asiatic Society. (2025). *New additions to the Digital Library*. <https://royalasiaticsociety.org/new-additions-to-the-digital-library/>
- Salehuddin, K., Jaafar, N. M., Jaafar, S. R. S., & Shamsudin, N. A.-D. (2025). Investigating the spelling pattern that can potentially propagate Jawi among speakers of Malay. *Reading and Writing*. <https://doi.org/10.1007/s11145-025-10742-9>
- Shofiyanti, L., & Rahman, H. (2026). Hibah naskah DREAMSEA dan strategi pelestarian ingatan kolektif Nusantara. *Manuskripta*, 15(2), 269–293. <https://doi.org/10.33656/manuskripta.v15i2.173>
- Siregar, I., Anjani, S. D., & Yahaya, S. R. (2023). Projection of the vitality of the Betawi language in the future time in Jakarta. *International Journal of Linguistics, Literature and Translation*, 6(3), 39–46. <https://doi.org/10.32996/ijllt.2023.6.3.4>
- Siregar, I., & Yahaya, S. R. (2023). The Betawi lexicon of kinship and greeting. *Formosa Journal of Sustainable Research*, 2(3), 667–680. <https://doi.org/10.55927/fjsr.v2i3.3522>
- Syahrullah, S., Muhammad, W. I., Nugraha, E., & Jannah, A. R. (2024). Corpus Coranicum and digital philology: A methodological model for advancing Qur'anic manuscript studies in Indonesia. *Khazanah Theologia*, 6(3), 201–222. <https://doi.org/10.15575/kt.v6i3.45610>
- Tang, X., & Li, J. (2024). Toward integrative triangulation in discourse-historical approach. *Discourse & Society*. <https://doi.org/10.1177/09579265231215473>
- Wahyuni, S. S., Siregar, I., Rachmawati, K., & Hamzah, N. H. (2024). Written language errors of

- dysgraphic children aged 9–12 years in inclusive classrooms. *Jurnal Onoma: Pendidikan, Bahasa dan Sastra*, 10(3), 3420–3430. <https://doi.org/10.30605/onoma.v10i3.4126>
- Wahyuni, S. S., Siregar, I., Rachmawati, K., Hamzah, N. H., & Yahaya, S. R. (2025). Kesalahan bahasa tulis anak disgrafia umur 9–12 tahun siswa di kelas inklusi SDN Parakan 01. *Stilistika: Jurnal Pendidikan Bahasa dan Sastra*, 18(1), 191–202. <https://doi.org/10.30651/st.v18i1.23343>
- Wardani, E. T., Siregar, I., Susanto, A., Hamzah, N. H., & Yahaya, S. R. (2025). Strategi dan peranan orang tua terhadap pemerolehan bahasa ABK dalam komunitas masyarakat Betawi. *Stilistika: Jurnal Pendidikan Bahasa dan Sastra*, 18(1), 131–146.
- Yaqin, L. N., Yusof, B., & Damit, A. R. (2024). “Status quo” of Malay linguistics studies: A bibliometric analysis. *Journal of Language and Literature Studies*, 4(2). <https://doi.org/10.36312/jolls.v4i2.1854>