

A CORPUS STYLISTIC AND SEMIOTIC INQUIRY INTO POSTCOLONIAL TEXTS: *SAMPAH* and *TIVI* by SHAHNON AHMAD

Dahlia Janan

Fakulti Bahasa dan Kesusasteraan Melayu Universiti Pendidikan Sultan Idris Malaysia
dahlia@fbk.upsi.edu.my

Received 2025-11-05; Revised 2025-12-12; Accepted 2025-12-12

ABSTRACT

This study investigates the intersection of corpus stylistics and semiotic theory as an integrated methodological framework for interpreting postcolonial Malay literature. Focusing on Shahnnon Ahmad's novels *Sampah* (1974) and *Tivi* (1995), the research combines quantitative corpus technique such as keyword analysis, collocation mapping, and concordance inspection with qualitative semiotic interpretation to reveal how stylistic patterns encode ideological meanings of identity and moral decay in postcolonial Malaysia. Results indicate prominent lexical clusters of contamination and decay, recurrent religious-moral binaries and the media as a sign of cultural colonization. Semiotic analysis shows these patterns function as signs that both diagnose and resist the cultural consequences of modernity. The study argues that corpus stylistics, when coupled with semiotics, offers a robust model for empirically grounded yet culturally sensitive readings of postcolonial texts, enriching theoretical debates in postcolonial stylistics and digital humanities.

Keywords: postcolonial, semiotic, corpus stylistics

ABSTRAK

Kajian ini meneliti persilangan antara stilistik korpus dan teori semiotik sebagai metodologi bersepadu untuk mentafsir kesusasteraan Melayu pascakolonial. Berfokus pada novel Shahnnon Ahmad, *Sampah* (1992) dan *Tivi* (1995), kajian ini menggabungkan teknik kuantitatif korpus seperti analisis kata kunci, pemetaan kolokasi, dan penyemakan konkordans dengan tafsiran semiotik kualitatif untuk menunjukkan corak stilistik menzahirkan makna ideologi tentang identiti dan kemerosotan moral dalam kalangan masyarakat Malaysia pascakolonial. Dapatan menunjukkan gugusan leksikal yang ketara berkaitan pencemaran dan kerosakan, pertentangan agama-moral yang berulang, serta media sebagai tanda penjajahan budaya. Analisis semiotik memperlihatkan corak ini berfungsi sebagai tanda yang mendiagnosis dan penentang budaya dan Kesan modenisasi. Kajian ini juga menggabungkan stilistik korpus dan semiotik yang menawarkan model kukuh untuk data empirikal yang masih sensitif budaya. Gabungan ini memperkayakan perdebatan teori dalam stilistik pascakolonial dan kemanusiaan digital.

Kata kunci: pascakolonial, semiotik, stilistik korpus

INTRODUCTION

Postcolonial studies have persistently interrogated relationships among language, power, and representation in the aftermath of empire. Literary texts from formerly colonised societies are crucial sites where language is reworked to articulate local

identities, contested histories, and cultural imaginaries. Within the Malaysian context, postcolonial discourse is marked by tensions among indigenous linguistic traditions, colonial legacies, and the centrifugal forces of modernization and globalization. Shahnon Ahmad's novels *Sampah* (1992) and *Tivi* (1995) stage such tensions vividly, delivering sustained critique of materialism, corruption, and the erosion of communal and spiritual values.

Traditional criticism on Shahnon has often foregrounded thematic and moral readings; however, there remains a need for empirical, language-based investigations that quantify stylistic tendencies and connect them systematically to meaning. Corpus stylistics, an approach that uses computational methods to detect patterned language use, offers such empirical grounding while preserving interpretive nuance (Baker, 2010; McEnery & Hardie, 2012). When combined with semiotic theory, which situates linguistic signs within systems of cultural signification (Barthes, 1972), a powerful interdisciplinary lens emerges for decoding how stylistic choices operate ideologically in postcolonial literature.

This paper presents a balanced integration of corpus stylistics and semiotics applied to *Sampah* and *Tivi*. The research aims to demonstrate that Shahnon's linguistic choices of lexical repetition, collocational patterns, syntactic rhythm, and register mixing able to act as semiotic resources that encode concerns about cultural authenticity, moral order, and the colonising effects of media and consumerism. The study addresses three principal research questions: (1) What stylistic patterns do corpus stylistic analyses reveal in *Sampah* and *Tivi* that pertain to postcolonial themes of power and identity? (2) How do these stylistic features operate as moral commentary, and cultural critique? (3) In what ways does integrating corpus stylistics and semiotic theory contribute to methodological and theoretical understandings of postcolonial meaning-making in Malaysian literature?

Methodologically, the project compiles digital corpora of the two novel, employing AntConc for keyword, collocate, and concordance analyses. Semiotic

interpretation is then used to contextualise corpus patterns within Malaysia's sociocultural setting and postcolonial theoretical frameworks (Bhabha, 1994; Ngūgĩ, 1986). The significance of this research lies in its demonstration that empirical textual analysis and cultural semiotics are mutually reinforcing: corpus tools surface patterned linguistic evidence, while semiotics situates these patterns within ideological and symbolic fields. In doing so, the study situates Shahnnon Ahmad's works within global and local debates on the politics of language, modernization, and moral imagination.

THEORETICAL FRAMEWORK

This study is anchored in an interdisciplinary theoretical synthesis that brings together corpus stylistics, semiotic theory, and postcolonial literary studies. Each strand contributes critical conceptual resources: corpus stylistics offers tools for detecting systematic linguistic patterning (Semino & Short, 2004; McEnery & Hardie, 2012, McIntyre, D., & Walker, R., 2019), semiotics furnishes interpretive mechanisms for how language functions as signs within cultural systems (Barthes, 1972), and postcolonial theory frames the sociohistorical dynamics that make such signs politically charged (Ashcroft, Griffiths, & Tiffin, 2002; Bhabha, 1994).

Corpus Stylistics: Language as Patterned Data

Corpus stylistics regards style as patterned and amenable to empirical detection. Keywords, collocations, semantic prosodies, and concordance patterns reveal regularities that contribute to aesthetic effect and ideological orientation (Mahlberg, 2013; Semino & Short, 2004). In postcolonial contexts, such empirical methods can uncover lexical traces of colonial discourse, mixture of registers (Baker, 2010). The approach privileges neither mechanical quantification nor purely impressionistic reading; instead, it provides reproducible data to support interpretive claims.

Semiotics: Signs, Myths, and Cultural Meaning

Semiotics examines how signs (words, metaphors, motifs) generate meaning through their relations and contexts Barthes's (1972). Eco (1976) emphasises textual polysemy and the openness of interpretation. In postcolonial literature, semiotics helps

decode how linguistic choices transform everyday objects (e.g., “*tivi*”, “*sampah*”) into loaded cultural symbols that embody collective anxieties, desires, and power relations.

Three-tier Conceptual Model

The study operates with a three-tier model: (1) Corpus Level is the identification of statistical stylistic markers; (2) Semiotic Level is the interpretation of these markers as signs within cultural codes; (3) Postcolonial Level is the situating signs within Malaysia’s historical, moral, and ideological terrain. This model ensures data-driven claims remain culturally contextualized.

METHODOLOGY

A mixed-method design integrates corpus stylistic techniques with semiotic analysis to examine Shahnnon Ahmad’s *Sampah* and *Tivi*. The procedural steps encompass corpus construction, analytical routines, and interpretive strategies.

Corpus Construction

The primary corpus comprises digitised full texts of *Sampah* and *Tivi*, each treated as a separate sub-corpus and together forming a combined corpus for cross-textual analysis. Texts were cleaned of paratextual material (pagination, editorial notes) and normalised for orthographic consistency. Basic statistics (word count, type-token ratio, lexical density) were computed to profile stylistic variation.

Analytical Tools and Procedures

Analyses were conducted in AntConc (version 3.6.0). The principal techniques were:

- Keyword analysis to identify lexical items significantly over-used in the target novels.
- Collocation analysis to discover habitual co-occurrences and semantic associations.
- Concordance inspection to contextualise keyword appearances and interpret semantic prosody and syntactic patterns.
- Semantic field mapping to group keywords into conceptual clusters (e.g., contamination, religion, media).

Quantitative findings guided the subsequent semiotic interpretation.

Semiotic Interpretation Strategy

Semiotic reading followed three steps: (1) classify recurrent lexical clusters as semiotic markers (e.g., *sampah* cluster, *tivi* cluster); (2) contextual re-signification such as interpret markers within Malaysia's sociohistorical conditions (modernisation, Islamisation, neoliberal consumerism); (3) map inter-semiotic relations such as binary oppositions and symbolic networks (purity/corruption, sacred/profane, rural/urban).

Limitations and Ethics

Limitations include the small number of primary texts (two novels), which restricts broad generalisation. Interpretation involves researcher reflexivity; semiotic readings are context-sensitive and open to alternative interpretations. Texts were used under fair academic use; all procedures ensure methodological transparency to facilitate replicability.

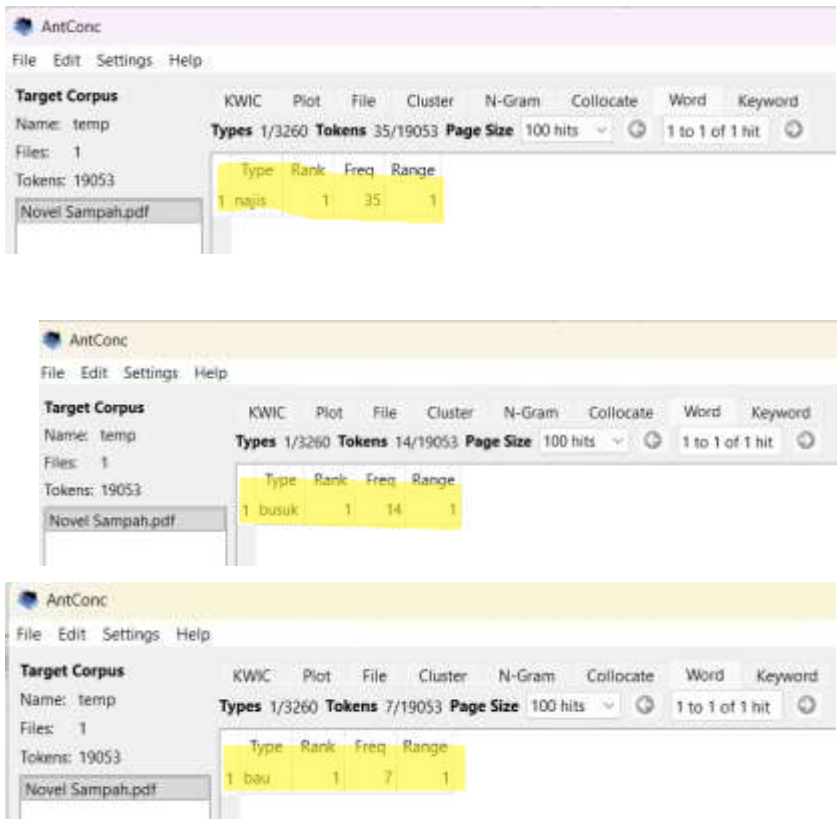
FINDINGS AND DISCUSSION

This section integrates corpus evidence and semiotic interpretation. Analysis identified three interrelated thematic domains: (1) lexical patterns of *contamination and decay* (metaphors to describe the degradation of society or an institution morals); (2) religious and moral discourse markers; (3) media, materialism, and cultural alienation.

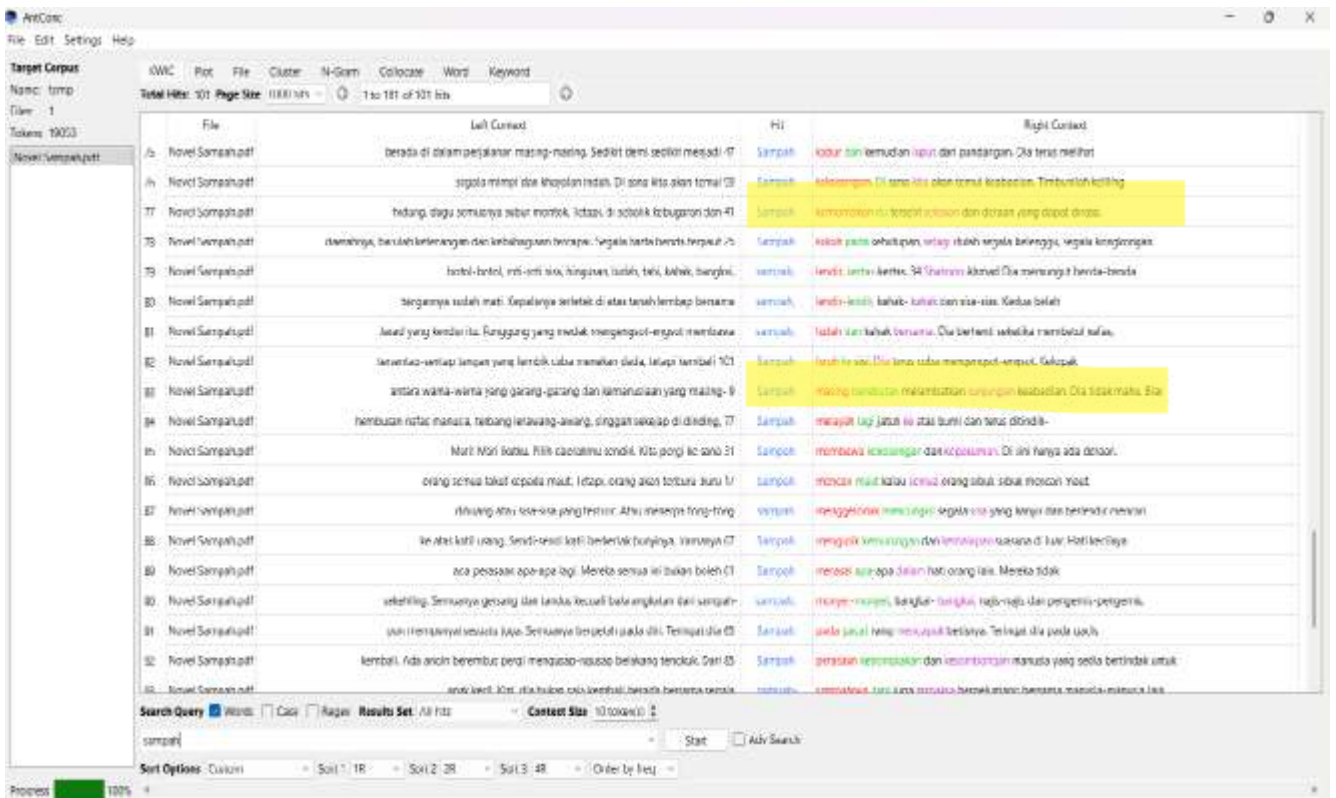
Lexical Patterns of Contamination and Decay

Keyword analysis positions *sampah* and related vocabulary (*busuk, najis, bau*) as markedly over-represented in *Sampah*. Collocation patterns (e.g., *sampah membawa kekosongan dan kepakuman* (85) and *sampah kukuh pada kehidupan* (78)) demonstrate metaphorical extension: physical filth becomes symbolic of moral degradation. In semiotic terms (Barthes, 1972), these material signs are made symbolic into larger ideological claims about national decay.

Sampah: Keywords analysis



Sampah: Collocation patterns



In *Tivi*, contamination becomes psychological: collocates around *Tivi* example *jantan*, *dajal*, *ganas* that indicate mind overload with negativeness. The television is a vector for cultural colonisation, an internalised force that mediates desires and erodes agency. Corpus evidence thus reveals a consistent lexical ideology: filth and pollution function as semiotic sites for critiquing modernity’s corrosive effects.

Religious and Moral Discourse Markers

Both novels exhibit dense clusters of religious lexis (*Allah*, *sembahyang*, *masjid*, *telekung*, *jemaah*). Concordance lines reveal frequent opposing structures such as *telekung itu pun sudah hapak* and *telekung dilonggokan*, which linguistically stage moral confrontations through the neglect of proper prayer. The use of intensifiers (*amat*, *terlalu*, *benar-benar*) amplifies negative evaluations, producing a rhetorical rhythm of moral urgency. Semiotic analysis interprets these oppositions as signs of a society caught between spiritual norms and secular temptations which a recurring postcolonial narrative of ethical loss.

TIVI: Concordance lines

| File | Left Context | Hit | Right Context |
|-------------------|---|----------|--|
| 1 Novel Tivi.pdf | telekung mak-pun ada itu juga. Dan, telekung serta anak | telekung | itu pun sudah hapak. Mak-jarang-jarang susah beribad |
| 2 Novel Tivi.pdf | di-paku di-biang-pempuru-itu. Telekung mak-dan anak | telekung | mak-pun ada itu juga. Dan, telekung serta anak |
| 3 Novel Tivi.pdf | macam-kain-burak-saja. Kertapas, jubah-ayah, telekung-emak-dan | telekung | Kain-kain-pun dilonggokkan-bersama-sejatan-burak. Hampas-bertimbun |
| 4 Novel Tivi.pdf | mak-dan anak-telekung mak-pun ada itu juga. Dan, | telekung | serta anak-telekung itu-pun sudah hapak. Mak-jarang- |
| 5 Novel Tivi.pdf | sejatan-sejatan-itu di-masa-masa-pun. Longgok-dulu-bersama | telekung | dan-ku-ku-ku-ku-ku-ku-ku. Tengok-Ti-itu-dulu |
| 6 Novel Tivi.pdf | aduk-aduk-dimantar-lakap-di-perjuru-bersama-sama-dengan-longgokkan-sejatan, | telekung | dan-jubah-ayah. Zil-700-Perjuru-di-sebetah-matahari-jatuh |
| 7 Novel Tivi.pdf | dan-mak-pembahyang. Senanglah-kalau-Ti-ada-di-situ. Sejadah, | telekung | dan-kertapas-ayah-burak-diletakkan-di-masa-masa-saja |
| 8 Novel Tivi.pdf | ke-terpang-lain, termasuk-tali-sembang, sejatan, al-Quran, kain | telekung | latah, kelabu-dan-berbulu-nya. Batu-itu-ditukap-tempat-keluarga |
| 9 Novel Tivi.pdf | saja. Labakkan-saja-terlalu-itu-di-perjuru-tersebut. Ruang-sejadah, | telekung | jubah, burak, dan-al-Quran-itu-se-tempat-lain. |
| 10 Novel Tivi.pdf | juhan-urang-ayah-bertangung-di-paku. di-Fang-perjuru-itu | Telekung | mak-dan anak-telekung mak-pun ada itu juga. |
| 11 Novel Tivi.pdf | sejadah-itu-bertonggok-macam-kain-burak-saja. Kertapas, jubah-ayah, | telekung | emak-dan-telekung-ku-itu-pun dilonggokkan-bersama-sejatan. |
| 12 Novel Tivi.pdf | paku-itu-emak-batang-dekat-dinding-itu-sesuai-sejatan-dan | telekung | -betah-ke-kita-di-situ. Sudah-ayah-pun-betah-diangkut- |

Media, Materialism, and Cultural Alienation

In *Tivi*, the television is central: *tv* appears with high frequency and collocates with *di rumah, dibeli, tersergam*. The novels' paratactic, repetitive syntax like short clauses and repeated verbs will mimics the mechanical cadence of media consumption (e.g., *tengok tv, tengok gambar, tengok, tengok sampai habis, tengok lama-lama*). This stylistic mimicry performs a mimesis of media trance, while semiotic analysis positions the *tivi* as a symbol of an ideological invasion where global consumer images colonise local modes of valuation (Baudrillard, 1983).

TIVI: Repetitive syntax

| File | Left Context | Hit | Right Context |
|----------------|--|--------|---|
| Novel Tivi.pdf | rumah mereka. Sebaliknya, banyak orang akan datang ke rumah untuk | tengok | TV yang dibelikan oleh Chah itu. Sekarang Jeha lega |
| Novel Tivi.pdf | depannya dan memontonnya. Ataupun baring-baring - slang dan malams - | tengok | apa yang sedang dimainkannya. Hasanah tentu suka, dan Hassan |
| Novel Tivi.pdf | Bukan 33 Tivi saja Mat Isa berdiri lama-lama di situ | tengok | apa yang sedang berlaku dalam TV, tetapi dia sendiri |
| Novel Tivi.pdf | TV. Bukan saja 55 Tivi nak tengok TV itu, tetapi nak | tengok | apa yang ty sanah tengok. Waktu itu matahari sudah |
| Novel Tivi.pdf | Chahiah yang akan tengok sendiri. Dan, apabila mak Chah sedang | tengok | babak itu tentu mak Chah naik getahah. Waktu itu |
| Novel Tivi.pdf | Tetapi, Hassan berjanji yang dia akan centa kelak, apabila dia | tengok | centa itu sampai habis. Dan kalau Kak Chah berdegit |
| Novel Tivi.pdf | Semua tengok babak itu, tetapi semuanya melongong saja. Mak Chah | tengok | knah itu macam tengok nani sejuk dalam pinggan saja. |
| Novel Tivi.pdf | tak boleh masuk. Mana ada yalan masuk. Ini rumah mahu | tengok | tibi pun tak ada jalan ka?" "itu jalan ada!" |
| Novel Tivi.pdf | yang agak kuat. Tak siapa yang berkata apa-apa. Hendak | tengok | terus pun sudah terasa malu, tetapi hendak meninggalkan babak |
| Novel Tivi.pdf | tengok gambar jenis itu. Mat Isa membiar terus. Meminta nak | tengok | lagi pun tidak. Tak minta pun tidak juga. Dan |
| Novel Tivi.pdf | macam tak ada apa-apalah. Beberapa kali sewaktu semua sedang | tengok | TV, Chah dan jetri duduk rapat-rapat di sebelah |
| Novel Tivi.pdf | besar. Ada skru panjang, ada skru pendek. Jadi, kami jarang | tengok | Kak Chah balik lagi kecuali sebulan dua bulan sekali. |
| Novel Tivi.pdf | kawan-kawan. Jetri ini kawannyaalah. Taikan tak boleh berkawan. Cuba | tengok | dalam TV. Banyak jaritan dan betina berkawar, malah berpegang |
| Novel Tivi.pdf | saja babak cium dan dakap malah macam-macam lagi boleh | tengok | dalam TV. Ada yang terkangtang terkoyak sampai nampak segala |
| Novel Tivi.pdf | tu. Itulah apabila Chah kata nak beli TV untuk manya | tengok | dan ayahnya tengok sama, ayahnya setuju sangat-sangat. Kita |
| Novel Tivi.pdf | Chah kata nak beli TV untuk manya tengok dan ayahnya | tengok | sama, ayahnya setuju sangat-sangat. Kita mesti ikut apa |
| Novel Tivi.pdf | dan tumbuh gambar lain lagi. Mak Chah bersingut. Dia hendak | tengok | juga gambar orang basuh kain pakai sabun proton itu. |
| Novel Tivi.pdf | gambar-gambar itu keluar. Mat Isa tak tahan lagi nak | tengok | gambar-gambar itu. Hasanah kemudian bangun dan berlari-lari |
| Novel Tivi.pdf | saja. Nampak macam fak ada ana-ana tingkasan mat. Semua | tengok | babak itu, tetapi semuanya itu loooooo saia. Mak Chah tengok |

Synthesis: Corpus Evidence as Semiotic Signage

Collectively, the data show that Shanon's stylistic practice acts both diagnostically and prescriptively: he diagnoses moral malaise and prescribes moral vigilance. Corpus

patterns (keyword prominence, collocational networks, repetitive syntax) supply empirical backbone; semiotic reading embeds these patterns in cultural meaning-making. The novels thus enact **corpus semiotic**: linguistic patterns are data, and those patterns are also signs that challenge cultural colonisation and ethical erosion.

CONCLUSION AND IMPLICATIONS

By integrating corpus stylistics with semiotic analysis, this study traced how Shahnon Ahmad's *Sampah* and *Tivi* use patterned language to generate postcolonial critique. Empirical analyses revealed recurrent contamination metaphors, religious binaries and media-centred networks of meaning. All these serving as semiotic devices that narrate and resist the moral consequences of modernity and commodification.

Theoretical and Methodological Contributions

The paper demonstrates that corpus stylistics can substantively enrich postcolonial literary interpretation without reducing it to mere counts: quantitative patterns direct attention to meaningful stylistic features, and semiotics supplies interpretive depth. This balanced approach helps develop the study of postcolonial stylistics, and shows how digital tools can be useful for studying literatures written in languages other than English.

Cultural and Literary Significance

Shahnon's novels exemplify how language functions as a site of decolonization. The use of local rhythm, religious words, and satire helps reaffirm local moral values in the face of global influences. The *sampah* and *tivi* motifs remain potent metaphors for contemporary debates about development, identity, and moral authority in Malaysia.

Future Research Directions

Future studies could expand the corpus to include more Malaysian authors for comparative stylistic mapping, incorporate multilingual corpora to examine code choice across Malay and English literatures, or combine corpus results with reader-response methods to assess how stylistic patterns affect reception. Integrating critical discourse analysis would also deepen the socio-political reading of lexical networks.

REFERENCES

- Ashcroft, B., Griffiths, G., & Tiffin, H. (2002). *The Empire Writes Back: Theory and Practice in Post-Colonial Literatures* (2nd ed.). Routledge.
- Baker, P. (2010). Corpus Methods in Critical Discourse Analysis. In Wodak, R., & Meyer, M. (Eds.), *Methods for Critical Discourse Analysis* (2nd ed.). Sage. (Chapter reference for corpus approaches to ideology.)
- Baudrillard, J. (1983). *Simulations* (P. Foss, P. Patton, & P. Beitchman, Trans.). Semiotext(e). (Original work published 1981)
- Barthes, R. (1972). *Mythologies* (A. Lavers, Trans.). Hill and Wang.
- Bhabha, H. K. (1994). *The Location of Culture*. Routledge.
- Eco, U. (1976). *A Theory of Semiotics*. Indiana University Press.
- Mahlberg, M. (2013). *Corpus Stylistics and Dickens: Another Approach to Victorian Fiction*. Palgrave Macmillan.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.
- McIntyre, D., & Walker, R. (2019). Corpus stylistics and the integration of quantitative and qualitative methods. In *The Routledge Handbook of Stylistics* (pp. 1–20). Routledge.
- Ngũgĩ wa Thiong’o. (1986). *Decolonising the Mind: The Politics of Language in African Literature*. James Currey.
- Semino, E., & Short, M. (Eds.). (2004). *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. Routledge.
- Shahnon, A. (1995). *Sampah*. Dewan Bahasa dan Pustaka. Shahnon, A. (2001). *Tivi*. Dewan Bahasa dan Pustaka.

Acknowledgments:

This research was funded by MegaConsult Sdn Bhd (MCSB) under the External Research Grant [2024-0159-107-29].