

ENGLISH FOR ISLAMIC STUDIES IN INDONESIAN ESP CONTEXT: EXPLORATION OF CORPORA DIVERSITY IN TEXTBOOKS

Wanda Rohmatun Hamidah¹, Siti Asmiyah², Arba'iyah Yusuf³

Universitas Islam Negeri Sunan Ampel Surabaya, Jl. A. Yani 117 Surabaya, East Java, Indonesia
¹wandahamidah07@gmail.com, ²asmi.iain@gmail.com, ³arba.gusti@gmail.com

Abstract

Focusing on disciplinary English, students' learning ESP needs to be equipped with sufficient knowledge of disciplinary vocabulary. In this context, the use of corpora in ESP context is deemed to be important. This quantitative content analysis study explores the diversity of corpora in English for Islamic studies, focusing on three books from different Islamic universities: the State Islamic University of Sunan Ampel Surabaya (UINSA), University of Muhammadiyah Malang (UMM), and IAIN Walisongo. The books published by these top-ranked Islamic-based universities may be exemplary in the context of English for Islamic studies. Using AntConc, the study analyzes the language and terminology used in these books. The findings reveal some diversities, including unique topics, different terms with similar meanings, and variations in word frequencies from the same word. The study has implications for ESP in Islamic Studies. It provides reference words for ESP practitioners teaching Islamic knowledge. The corpus can be published for learners interested in English words related to Islamic studies. It offers an overview of how these words are used.

Keywords: Corpora Diversity, English for Islamic Studies, ESP, Textbooks Content

INTRODUCTION

The use of corpora in English for Specific Purposes (ESP) is particularly crucial because every discipline may have its discipline-specific vocabulary items. ESP is one of the branches of English Language teaching that concentrates on communicative English needs of particular jobs, disciplines or purposes such as law, science, business, and Islamic studies. ESP helps people to know the language for their job or purposes and teaches them that language (Hyland, 2007). ESP is vital in understanding the people about how language is used in different jobs or fields because every field has its own language or words. ESP also can help people master language directly to their needs. Better learning of ESP can be facilitated through the use of corpora.

Corpora come from the word corpus. The term "*corpus*" has its roots in Latin, where it means "body" or "collection." Corpus linguistics is a language study that focuses on the extensive collection of texts called corpora using software from computers to analyze and find the patterns in how words are used in texts (McEnery & Wilson, 2011). Several software tools for analyzing the corpora include TagAnt, AntConc, and Wordsmith.

This text is taken from reading or spoken text such as books, television conversations, social media, articles, and many others. Corpora can identify the most frequent and relevant words used in a specific field of study or profession. This situation hence may result in corpora diversity. For students and professionals, corpora helps in focusing their learning on the suitable vocabulary to their needs. Moreover, by knowing

specific vocabulary in their field study, students can effectively communicate with their colleagues and understand texts in their area of expertise.

Corpora diversity refers to the range of sources, genres, registers, dialects, and language varieties used in a corpus. A corpus is a collection of texts or spoken language compiled and analyzed for linguistic research, natural language processing, and machine learning. Corpora diversity is essential because it provides a more representative sample of language use in different contexts, cultures, and communities. Corpora must be comprehensive or heterogeneous to avoid biased or inaccurate results. Diverse corpora can seize the richness and complexity of language variation and change. Corpora diversity can also facilitate cross-linguistic and cross-cultural comparison, language teaching and learning, and corpus-based applications in various fields (Biber et al., 1998; Crystal, 1995; Dash & Arulmozi, 2018; Hunston, 2002; McCarthy, 2004; McEnery & Wilson, 1996; Sinclair, 1991). In summary, diverse collections of language samples (corpora) are essential in linguistic research. It helps us understand how language is used, prevents biases, and allows us to study how language can be different and change over time.

Several studies have highlighted corpora in various aspects (Cao, 2021; Hoang, 2023; Weisser, 2021), such as multiword units across textbook and learner corpora, lexical bundles across paradigms and disciplines, and profiling learners through pragmatically. Also, there is a study about ESP textbook evaluation and vocabulary comparison (Istiqomah & Maunah, 2020; Ningsih, 2023). However, there have yet to be any attempts to research comparing the differences and similarities in textbooks of English for Islamic studies from many Islamic institutes or universities.

English for Islamic Studies (EfIS) is a part of English for Specific Purposes (ESP) that uses literature about the Islamic context as the medium of learning English language. EfIS gives insight into Islamic materials for students who are studying Islamic education by enriching vocabulary (Istiqomah & Maunah, 2020).

Therefore, this current study explores the diversity of corpora between three books of English for Islamic studies from different publishers, such as the State Islamic University of Sunan Ampel Surabaya (UINSA), University of Muhammadiyah Malang (UMM), and IAIN Walisongo.

The reason for choosing these books is that three are from Islamic publishers, such as UINSA and UMM. At the same time, English for Islamic Studies is from Djamaluddin Darwis, published by cooperating with IAIN Walisongo. Despite three being from Islamic institute publishers and having the same title as each other, it would be interesting to explore diverse corpora and content if they have any differences and similarities.

METHOD

This study used quantitative content analysis Coe & Scacco (2017) and selected three English books for Islamic studies from different publishers: the State Islamic University of Sunan Ampel Surabaya (UINSA), the University of Muhammadiyah Malang (UMM), and Djamaluddin Darwis published by cooperating with IAIN Walisongo. The books were converted into *.txt format and analyzed using Antconc to generate a list of lexical items used in each book. The researcher carefully examined the list to identify the variations in lexical items among the books, focusing on removing

similarities and highlighting diversity. The diversity shows the unique characteristics of each book of English for Islamic studies. The diversity informs the way in which the discourse of Islamic studies varies within the context of Indonesian ESP.

FINDINGS AND DISCUSSION

The data from three textbooks in English for Islamic Studies were extracted to obtain the words and their numbers by utilizing AntConc to study the variety of Islamic terms and compare them to identify their differences. The top 100 most frequent words are found and selected by the criteria that the words must be about or refers to Islamic studies topics or context. The top ten Islamic words are found and shown in Table 1.

Table 1 The Top Ten Vocabulary Featuring The Discourse of Islamic Studies

No.	Islamic word	Frequency	No.	Islamic word	Frequency
1.	Islam	294	6.	Muhammad	130
2.	Muslim	180	7.	Islamic	127
3.	Prophet	175	8.	Al-Qur'an	96
4.	Allah	169	9.	Religion	64
5.	God	164	10.	Fasting	41

The findings showed ten common Islamic words in the textbook. Islamic words refer to vocabulary or terms specifically associated with Islam; the religion founded by Prophet Muhammad in the 7th century. These words may include terms related to Islamic theology (such as *Tawhid*, *al-Qadar*, *afterlife*, *revelation*), rituals (such as *Shahada*, *Salah*, *Sawm*, *Fasting*, *Zakah*, *Hajj*), practices, concepts, holy books (*Tawrah*, *Zabur*, *Injil*, *Al-Qur'an*), and Islamic culture.

In addition to the above words clearly showing words related to Islam, some other function words, such as pronouns, nouns, determiner, and auxiliary verbs seem unrelated to Islamic studies. However, close observation of the context of the words shows that the words actually refer to topics and contexts related to Islamic studies. The top ten function words are taken and shown in Table 2.

Table 2 The Top Ten Function Words Related to Islamic Studies

No.	Function word	Frequency	No.	Function word	Frequency
1.	his	227	6.	them	53
2.	he	192	7.	all	52
3.	they	103	8.	this	46
4.	their	79	9.	be	45
5.	him	76	10.	people	42

The pronouns found in the three books are (e.g., "his," "he," "they," "their," "him," and "them"). The pronoun "his," "he," and "him" in the book refers to Muslim expertise (e.g., Al-Khawarizimi, Abu Kamil, al-Hasan, Muhammad ibn Musa, etc.), the Prophet's

name (e.g., Muhammad, Lut, Ismail, Adam), Muhammad's companion (e.g., Abu Bakr, Zaid ibn Harithah), Muhammad's closest uncle (Abu Talib), the recipient of Zakah (e.g., the wayfarer, the insolvent, the needy), Islamic ritual (e.g., the pilgrim), creature mentioned from the holy book (the Devil) and god (Allah), Hadith Narrator (e.g., Muslim), non-believer (*kaafir*). Moreover, the pronoun "they", "their," and "them" in the book refers to Muslim women's roles in Islamic civilization, the believers (members of *the ummah*), the five pillars of Islam, the Zalimun (e.g., polytheists, hypocrites, and wrongdoers), creatures mentioned from the holy book (the angels). The prophets and messengers of God (Adam, Noah, Abraham, Ishmael, Isaac, Jacob, Moses, Jesus), The four books of Allah (Tawrah, Zabur (Psalms), Injil (Gospel), and the Qur'an).

The determiners found in the three books are "all" and "this." The determiner "all" in the book refers to various aspects related to Islam and Muslims. It encompasses the entirety or collective group of Muslims, the entirety of Islamic culture or practices, and the collective factors contributing to technological achievements among Muslims. "All" also encompasses Muslims and their obligations, such as being educated and performing the Hajj pilgrimage. Furthermore, it refers to references to the Qur'an, which contains factual statements considered absolute truths by believers and verses demonstrating the prophet's humbleness and nobility in his mission to spread the Message of Islam. Additionally, "all" extends to encompass women's participation in the social and cultural activities of early Islam, Muslim states (Indonesia), forbidden and detestable acts (e.g., inciting terror, destroying properties, and harming innocent people.)

The determiner "this" in the book refers to various aspects related to Islamic studies. It includes the acceptance of E-Commerce as a new way or technology to facilitate economic transactions based on the Syariah (Islamic Law), the involvement of Muslim women in military affairs, the understanding of "Trading Data Management" through "Islamic Business Transactions" (*Fiqh al-Mu'amalat*), the belief in Divine Predestination (*al-Qadar*), and the Islamic idea consisting of the creed or doctrine (*aqeedah*) and a system of rules and regulations based on this creed (*shari'a*). Moreover, it encompasses obligations such as *zakah* (support of the needy) and the testimony of faith. Ultimately, "this" refers to Islam, the Islamic finance concept (*Mushaarakah, Murabahah*), natural inborn belief in Allah (*Fitrah*), and many others.

The auxiliary verb "be" refers to respectful salutation to the prophet (Peace be upon him/them), Arabic greetings in Islam (peace be upon you), expressing praise to Allah (Glory be to Him. may He be glorified).

The noun "people" in the book refers to collective people (*ummah*) such as, people of Luth, the people of Mecca, Arabic people, and Quraysh people.

After finding the vocabulary and function words related to Islamic studies, comparing these three books can be made to highlight their difference. Three varieties of differences are found in these Islamic Studies books: First, unique topics. There is one topic in one book that the other does not have. The topics are shown in Table 3.

Table 3 The Special Topics in Three Textbooks in English for Islamic Studies

UINSA Book		
Word	Classification	Example
They	Referring to "Muslim expertise"	al-Hasan, Abu Sa'id al-Darir al-Jurfani, Abul Abbas Ahmad Ibn Muhammad, Abu Kamil, Muhammad Ibn Musa, Al-Khawarizmi
UMM Book		
Word	Classification	Example
They	Referring to Ashara Mubashara, "al-Mubasharabil Jannah" (the ten who were promised Paradise)	Abu Bakr, Ali ibn Abi Talib, Talha ibn Ubaydullah, Abdur Rahman ibn Awf, Sa'id ibn Zayd, Abu Ubaydah ibn al-Jarrah, Umar ibn al-Khattab, Sa'd ibn Abi Waqqas, Uthman ibn Affan, Zubayr ibn al-Awwam,
IAIN Walisongo Book		
Word	Classification	Example
They	Referring to the four books of Allah in the Quran,	Tawrah (Torah), Zabur (Psalms), Injil (Gospel), and the Quran itself

From this data, it can be concluded that the word "they" has different contexts in each of the three books related to Islamic studies. "They" in each book shows different aspects of Islamic knowledge, including expertise, righteous individuals, and sacred texts. The UINSA book refers to Muslim expertise. The UMM book refers to the ten individuals promised Paradise; in other words, they can be called Muhammad's companions. The IAIN Walisongo book refers to the four holy books of Allah mentioned in the Quran.

Second, several different terms but similar meanings are found across different books. For example, "*the hajj*" in the UMM book is equivalent to "the pilgrim" in the UINSA book. Likewise, "prayer" in the UINSA book corresponds to "*salah*" in the UMM book. "*sawm*" in the UMM book is the same as "fasting" in the IAIN Walisongo book. Similarly, "the pillars of Islamic faith" in the IAIN Walisongo book aligns with "six basic beliefs in Islam" in the UINSA book.

Third, there are the same words but have different frequencies in three books, especially word or vocabulary related to Islamic studies. These vocabularies are shown in Table 4.

Table 4 The Same Words With Diverse Frequencies in Three Textbooks in English for Islamic Studies

No.	Islamic word	Frequency		
		UINSA	UMM	IAIN Walisongo
1.	Islam	100	110	84
2.	Muslim	89	61	30
3.	Prophet	33	47	95
4.	Allah	32	54	83
5.	God	53	32	79
6.	Muhammad	40	32	58
7.	Islamic	35	83	9
8.	Al-Qur'an	3	47	48
9.	Religion	17	17	30
10.	Fasting	8	18	15

To sum up the results, there are diversities found in these three books, such as; first, unique topics from each book; second, different terms with similar meanings; last, same words with different frequencies.

The study aims to reveal diversity between English for Islamic studies books from three universities. In line with Nagano's (2015) statement, different fields of study have their own ways of writing and communicating ideas.

This current results study agrees with several previous studies, first from Cao, (2021). Previous results show how the terms in articles are impacted based on the method scholars conduct and the subject they study. Similarly, one of the current findings found there are different terms but similar meanings in these three textbooks, and it showed that the terms used in Islamic Studies textbooks are impacted based on their publisher.

This current result study also showed that several different terms each other, such as "the hajj" with "the pilgrims", "prayer" and "salah", "sawm" and "fasting", "the pillars of Islamic faith" and "six basic beliefs in Islam" has a similar meaning. When students use textbooks from many publishers, there will be two possibilities: 1) the students can enrich their vocabulary, and 2) the students get confused in understanding some terminology because one meaning can be conveyed by many terms. This situation is also found in the Hoang's (2023) study. In that finding study, students encountered difficulties with article and word group combinations.

Each book that is the source of this data has a unique topic and a wealth of terminology. These books can be used as a reference for ESP, particularly for learning disciplinary vocabulary. Therefore, when interacting, either verbally or in writing, the students will use words relevant to the topic of discussion in their discipline. Thus, words that do not contribute to conversational interactions no longer appear. As Weisser (2021) discovered, people learning a language commonly use many words to help them plan what to say. However, these words only sometimes contributed much to the conversation regarding structure or interaction. Many of the words used do not contribute to the

conveyance of the message. This may be because of their unfamiliarity with the common words used in the discipline. With the use of corpora in the three books, students learning English for Islamic studies may avoid the use of words that do not contribute to their written or verbal communication in the discipline.

A study by Istiqomah & Maunah (2020) evaluated textbooks focusing more on Islamic education vocabulary. While their focus may to some extent within the scope of Islamic studies, their findings on the textbook is further expanded with the findings of corpora in the current study. While Istiqomah & Maunah (2020) study is limited to the corpora in Islamic education, their finding in combination with the corpora in this current study will further enrich reference learning for Islamic studies. Given the varieties of topics and unique terminology used in each of the book, students have more options on varieties of more vocabulary items used in the discipline.

The findings of this current research also further strengthen the critical value of using corpora based on particular disciplines, especially in English for Islamic Studies. Ningsih's finding (2023) shows that that 45% of the most common words used by students matched the most common words in the textbooks. This means that despite the abundance of resources available, the use of textbook in English learning could not be ignored. Hence, the corpora of the three books serve as good learning resources for students to produce written and verbal information more accurately in their discipline.

CONCLUSION

This study explored a variety of corpora between three books of English for Islamic studies from different universities, such as the State Islamic University of Sunan Ampel Surabaya (UINSA), University of Muhammadiyah Malang (UMM), and IAIN Walisongo. The research focuses on three main aspects. First, it examines the unique topics covered in each book. Second, it explores using different terms with similar meanings in the books. Lastly, it compares the frequency of occurrence for the exact words in the books, noting any variations.

The study has significant implications for the field of English for Specific Purposes (ESP), especially in Islamic Studies. By examining the language and terminology used in these books, ESP practitioners have relevant reference words specific to Islamic knowledge for their students. Moreover, the corpus could be published for learners interested in learning English words about Islamic studies, providing an overview of how words related to Islamic studies are used.

However, the limited number of books may only partially represent part of the range of Islamic literature or other specialized domains. This study only took the top ten frequencies that may only represent some ranges of Islamic studies and focus on one discipline. For the suggestion, Future research could compare English terminology in different ESP contexts, such as legal, medical, or business domains. Also, further research can extend the number of texts to enlarge the corpora.

REFERENCES

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

- Cao, F. (2021). A comparative study of lexical bundles across paradigms and disciplines. *Corpora*, 16(1), 97–128. <https://doi.org/10.3366/cor.2021.0210>
- Coe, K., & Scacco, J. M. (2017). Content Analysis, Quantitative. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods* (1st ed., pp. 1–11). Wiley. <https://doi.org/10.1002/9781118901731.iecrm0045>
- Crystal, D. (1995). *The Cambridge Encyclopedia of the English Language*. Cambridge University Press.
- Dash, N. S., & Arulmozi, S. (2018). Definition of ‘Corpus.’ In N. S. Dash & S. Arulmozi, *History, Features, and Typology of Language Corpora* (pp. 1–15). Springer Singapore. https://doi.org/10.1007/978-981-10-7458-5_1
- Hoang, H. (2023). *A corpus-based investigation of multiword units across textbook and learner corpora* [PhD Thesis, The University of Queensland]. <https://doi.org/10.14264/3363a38>
- Hunston, S. (2002). *Corpora in Applied Linguistics* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524773>
- Hyland, K. (2007). English for Specific Purposes: Some Influences and Impacts. In J. Cummins & C. Davison (Eds.), *International Handbook of English Language Teaching* (pp. 391–402). Springer US. https://doi.org/10.1007/978-0-387-46301-8_28
- Istiqomah, F., & Maunah, B. (2020). An ESP textbook evaluation: The case of English for Islamic Insight. *JETLe (Journal of English Language Teaching and Learning)*, 1(2), 11. <https://doi.org/10.18860/jetle.v1i2.8262>
- McCarthy, M. (2004). *Touchstone: From corpus to course book* (1a ed). Cambridge University Press.
- McEnery, T., & Wilson, A. (1996). *Corpus linguistics*. Edinburgh University Press.
- McEnery, T., & Wilson, A. (2011). *Corpus linguistics: An introduction* (2. ed., repr). Edinburgh Univ. Press.
- Ningsih, F. (2023). Mapping Vocabulary: A corpus-based comparison of Indonesian fourth-grade students’ writing and English language textbooks. *Journal of English Language Teaching and Learning (JETLE)*, 4(2), 136–150. <https://doi.org/10.18860/jetle.v4i2.20683>
- Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford university press.
- Nagano, R. L. (2015). Research Article Titles and Disciplinary Conventions: A Corpus Study of Eight Disciplines. *Journal of Academic Writing*, 133–144. <https://doi.org/10.18552/joaw.v5i1.168>
- Weisser, M. (2021). Chapter 6. Profiling learners through pragmatically and error annotated corpora. In P. Pérez-Paredes & G. Mark (Eds.), *Studies in Corpus Linguistics* (Vol. 102, pp. 121–148). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.102.06wei>