

Developing a corpus-based maritime English dictionary for enhancing engine room crew communication

Dhanan abimanto^{□1}, Hendi Pratama¹, Galuh Kirana Dwi Areni¹

Author Affiliations

¹*Doctoral program of language education, Universitas Negeri Semarang, Semarang, Indonesia*

Abstract This study addresses a critical gap in maritime communication within the Indonesian context, where engine room communication typically involves a hybrid of Bahasa Indonesia and unstandardized English technical terms. While international regulations like the Standard Marine Communication Phrases (SMCP) mandate English, they do not account for this prevalent code-mixing practice, creating significant risks of miscommunication linked to maritime incidents. To mitigate these risks, this research developed a specialized, corpus-based Maritime English dictionary for Indonesian Engine Room Crew (ERC). Adopting a qualitative, corpus-aided lexicographic design, the study compiled and analyzed a corpus of authentic Indonesian engine-room texts including logbooks and technical manuals, to identify high-frequency English technical terms and their contextual usage patterns. The findings confirm the existence of a core lexicon of English terms that are systematically embedded into formulaic Bahasa Indonesia phrases, representing the functional norm of technical discourse. The resulting thematically structured dictionary, validated by subject matter experts, serves as a practical tool to standardize this critical vocabulary. The study's novelty lies in its dedicated focus on this hybrid linguistic domain, aiming to enhance communication clarity, operational safety, and the effectiveness of Maritime English training for Indonesian marine engineering personnel.

Keywords: maritime English, corpus, dictionary, ship engine room

INTRODUCTION

In the globalized maritime industry, English has been firmly established as the essential lingua franca, a non-negotiable tool for ensuring operational safety, efficiency, and international cooperation (Bocanegra-Valle, 2021). This status is codified in international law through the International Maritime Organization's (IMO) International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW), which mandates a working knowledge of English for certified officers (Arbabi et al., 2025). Recognizing that proficiency alone is insufficient without standardization, the IMO developed the Standard Marine Communication Phrases (SMCP) to provide a precise, simple, and unambiguous lexicon for critical communications (Arbabi et al., 2025; Tchkonja, 2019). The SMCP serves as a vital "survival kit" for safety-related verbal exchanges, primarily covering ship-to-shore, ship-to-ship, and on-board navigational communication (Trenkner, 2022).

However, this global standard confronts a unique linguistic reality in specific national contexts, such as Indonesia. As an archipelagic nation, Indonesia's maritime operations are extensive, and on vessels operating within its domestic waters, the de facto language for daily on-board communication is Bahasa Indonesia (Sari & Sari, 2020). This is particularly true in operational spaces like the engine room, where crew members are often entirely Indonesian. Yet, communication is not conducted purely in Bahasa Indonesia. Instead, a hybrid form of communication or "code-mixing" is prevalent, where crew members use Bahasa Indonesia for syntax and general discourse but embed a significant number of English technical terms for machinery, components, and procedures (Sari & Sari, 2020; Simbolon & Yusnita, 2020). This practice arises because many technical manuals, systems, and international standards are in English, making direct translation cumbersome and leading to the adoption of English terminology as the professional norm.

This linguistic situation creates a specific and under-researched communication gap. While the use of Bahasa Indonesia facilitates general understanding among an all-Indonesian crew, the reliance on non-standardized English technical terms introduces significant risks. Inconsistent pronunciation, varied understanding of terms, and the lack of formal, context-specific training on this borrowed lexicon can lead to

critical misunderstandings during high-stakes operations (Rahmawati et al., 2020). This issue is exacerbated by the fact that many Indonesian seafarers report that language barriers are a significant challenge, even if their overall communicative competence is sufficient for social interaction. The problem is not the inability to hold a conversation in English, but the potential for misinterpreting a single, critical technical term within an otherwise fluent Indonesian instruction. A stark example from international contexts illustrates the gravity of this issue: a motorman who understood the order to "evacuate" but not to "close all openings" inadvertently allowed a fire to destroy a vessel (Trenkner, 2022). Such an incident could easily occur in the Indonesian context if an English technical term embedded in a Bahasa Indonesia command is misunderstood.

To address this critical safety and training deficit, this study positions itself within the framework of English for Specific Purposes (ESP), tailored to the Indonesian maritime context. The challenge for Indonesian seafarers is not necessarily achieving fluency for international voyages, but mastering the specific English lexicon required for their technical duties at home (Nugroho & Widiastuti, 2025). The most robust methodology for this task is corpus linguistics, which allows for the analysis of authentic language as it is actually used (Stefanowitsch, 2020). This approach directly confronts the lack of authentic teaching materials that reflect the real-world, hybrid communication needs of Indonesian cadets (Fitriati & Rata, 2020). Therefore, this study's novelty lies in its dedicated application of corpus-aided lexicography to the specific linguistic environment of the Indonesian engine room. The aim is to develop the first specialized Maritime English dictionary focused on the technical terms used by Indonesian ERC, a tool designed to standardize this critical vocabulary, enhance communication clarity, mitigate human error, and inform a more effective ESP curriculum for Indonesian marine engineering personnel.

METHODS

This study adopts a qualitative, corpus-aided lexicographic design to develop a specialized dictionary for Engine Room Crew (ERC) operating in the Indonesian maritime context. This approach was selected because it combines the empirical, data-driven analysis of corpus linguistics with the qualitative interpretation and structuring inherent in lexicography. The rationale is to create a descriptive resource that accurately reflects the authentic linguistic practices of marine engineering professionals in Indonesia. That is specifically their use of English technical terms within Bahasa Indonesia discourse, thereby ensuring its relevance and practical utility.

The first phase involved the construction of a specialized, context-specific corpus, the Indonesian Engine Room Corpus (I-ERC). Recognizing that general English corpora are inadequate for this task (Đurović, 2021), the I-ERC was meticulously compiled from a diverse range of authentic texts representative of the Indonesian engine room domain. The sources included: official technical manuals for machinery (often in English, as used on board); engine room logbooks from Indonesian-flagged vessels, which capture the hybrid language of routine operations; official maintenance and incident reports; and, where accessible, anonymized transcripts of on-board communications between Indonesian crew members to include samples of authentic spoken code-mixing. This multi-source approach ensures the corpus is balanced and representative of the varied communicative functions performed in this specific linguistic environment.

The second phase consisted of a systematic analysis of the compiled I-ERC using corpus analysis software. This focused on identifying the key English technical terms and the patterns of their use. First, frequency lists were generated to identify the most common English technical nouns, verbs, and adjectives that form the core borrowed lexicon. Second, collocation analysis was performed to identify recurring partnerships between English terms and the surrounding Bahasa Indonesia words (e.g., "start main engine, *cek* pressure"). Third, as described by Stefanowitsch (2020), lexical bundle analysis was used to extract recurrent multi-word sequences that demonstrate how English terms are embedded in functional Indonesian phrases (e.g., *prosedur untuk* starting, *pastikan* valve). Finally, keyword analysis, comparing the English terms in the I-ERC against a general English corpus, was conducted to isolate terms that are uniquely characteristic of the engine room context.

The final phase involved transforming the analytical data into a user-centered dictionary. The dictionary is structured thematically, organized according to key operational systems (e.g., Propulsion System, Lubricating Oil System) and procedures (e.g., Bunkering, Maintenance). This thematic organization is designed to enhance its utility as a practical, on-the-job reference tool. Each entry for an English technical term includes a clear definition (in both English and Bahasa Indonesia), part of speech, and several authentic example sentences extracted directly from the corpus to provide rich contextual information on how the term is used within Indonesian sentences. To ensure the dictionary's accuracy and relevance, the draft manuscript underwent a validation process with subject matter experts, including experienced Indonesian chief engineers and maritime English instructors. This step ensures the final product is a legitimate and practical tool for its intended users.

FINDINGS AND DISCUSSION

The analysis of the Indonesian Engine Room Corpus (I-ERC) reveals a highly specialized and hybrid linguistic landscape. Communication is predominantly conducted in Bahasa Indonesia but is systematically interspersed with a core lexicon of English technical terms. The findings are presented in two parts: the identification of this core English lexicon and the discovery of prevalent hybrid phrasal patterns. These findings are then discussed in relation to their implications for maritime safety and professional training within the Indonesian context.

The Core English Lexicon of the Indonesian Engine Room

The frequency analysis of the I-ERC confirmed the existence of a dense, domain-specific lexicon of borrowed English terms. This core lexicon represents the fundamental conceptual world of the Indonesian marine engineer, where English terms are used to name critical components, systems, and processes. Table 1 displays the top 20 most frequent and salient English technical terms identified in the corpus. The prevalence of terms such as “engine”, “pump”, “valve”, “pressure”, and “temperature” underscores the constant focus on the state of machinery, while verbs like check, monitor, start, and stop reflect the action-oriented nature of the work. These are not just occasional borrowings but are the standard, high-frequency terms used for technical communication.

Table 1. The top 20 most frequent and salient technical terms

Rank	Term	Frequency (per million words)	Part of Speech
1	engine	3,450	Noun
2	pump	2,890	Noun
3	valve	2,560	Noun
4	oil	2,310	Noun
5	pressure	2,180	Noun
6	system	2,050	Noun
7	water	1,970	Noun
8	tank	1,840	Noun
9	fuel	1,760	Noun
10	check	1,650	Verb
11	temperature	1,590	Noun
12	line	1,420	Noun
13	filter	1,330	Noun
14	control	1,210	Noun/Verb
15	start	1,150	Verb
16	stop	1,120	Verb
17	main	1,080	Adjective
18	monitor	990	Verb

19	level	950	Noun
20	alarm	910	Noun

Prevalent Hybrid Collocational and Phrasal Patterns

Moving beyond individual words, the analysis revealed that engine room communication relies on highly conventionalized patterns of code-mixing. Indonesian marine engineers embed these English terms into predictable Bahasa Indonesia grammatical structures to perform specific communicative functions. These hybrid phrases form the backbone of procedural and operational discourse. The analysis identified numerous examples of these patterns, where an Indonesian verb or function word is paired with an English technical noun phrase. Table 2 provides examples of these significant hybrid units, categorized by their primary communicative function. This evidence suggests that communicative competence in the Indonesian engine room depends on mastering this specific repertoire of hybrid, functional phrases.

Table 2. Examples of these significant multi-word units

Communicative Function	Example Hybrid Phrases (Bahasa Indonesia + English Term)
Monitoring & Reporting	<i>tolong cek pressure gauge</i> ; monitor temperature mesin; oil level sudah turun; <i>lapor ke anjungan</i>
Maintenance & Repair	<i>laksanakan</i> maintenance; <i>kita</i> overhaul main engine; <i>ganti</i> filter element; <i>bersihkan</i> sea strainer
Operational Procedures	<i>prosedur untuk</i> starting the generator; start up boiler; <i>segera</i> shut down generator; <i>sesuai dengan</i> manual book
Safety & Emergency	<i>dalam keadaan</i> emergency; <i>aktifkan</i> fire detection system; <i>bunyikan</i> alarm; <i>pakai</i> emergency bilge suction

Discussion

The findings provide robust, empirical validation for the study's central premise: a specialized, hybrid linguistic system exists within the engine rooms of Indonesian vessels, and the English component of this system is critical but unstandardized. The prevalence of English technical terms (Table 1) and their systematic integration into Indonesian phrasal patterns (Table 2) demonstrates a unique communication reality not addressed by global standards like the SMCP or traditional Maritime English curricula. While the SMCP is designed for full English communication between multinational crews (Tchkonina, 2019), it is largely irrelevant for an all-Indonesian crew communicating primarily in Bahasa Indonesia. The real risk in this context is not a failure of general English conversation, but a failure to correctly understand the specific, high-stakes English terms embedded within that conversation.

This directly connects the study's findings to maritime safety in Indonesia. The lack of standardization for these borrowed English terms means that pronunciation, meaning, and usage can vary from person to person and ship to ship, creating a latent risk of miscommunication (Rahmawati et al., 2020). By codifying the most frequent and critical English terms and providing clear definitions and authentic usage examples, the proposed dictionary serves as a direct safety intervention. It aims to create a common, standardized technical vocabulary for all Indonesian marine engineers, reducing the ambiguity that contributes to human error (Fan et al., 2020; Luo & Shin, 2019). The dictionary becomes a tool for professional standardization, ensuring that when an engineer says to "check pressure," every member of the crew understands the command in exactly the same way.

Furthermore, these findings have profound implications for Maritime English training in Indonesia. The persistent challenges in Indonesian maritime ESP, which is often stemming from a disconnect between classroom materials and real-world needs, can be addressed by focusing on this hybrid linguistic reality (Fitriati & Rata, 2020; Wahyuni et al., 2021). Instead of teaching comprehensive English grammar and conversation that may rarely be used on domestic voyages, the curriculum can be reformed to prioritize what is actually needed: mastery of the core English technical lexicon. Instruction should focus on the correct pronunciation, definition, and contextual use of the terms identified in the corpus. This targeted, task-based approach is not only more efficient but also more motivating for cadets, as it equips them with the tangible, job-ready language skills they will use daily (Nugroho & Widiastuti, 2025).

This study contributes a valuable case study to the field of ESP by examining a professional community where code-mixing is the norm. It demonstrates how corpus methodologies can be adapted to analyze hybrid language use and develop specialized lexicographical tools for such contexts (Đurović, 2021). The process of building a specialized corpus that reflects this reality and involving local subject matter experts in validation ensures the final product is a practical tool co-created with and for the Indonesian maritime community it is intended to serve.

CONCLUSION

This study was initiated to address a specific communication challenge within the Indonesian maritime industry: the non-standardized use of English technical terms within Bahasa Indonesia-dominant communication in the ship's engine room. While global standards focus on English as a complete lingua franca, the reality on many Indonesian vessels is a hybrid language. This practice, while functional, carries inherent risks of miscommunication due to a lack of standardized vocabulary. By employing a corpus-aided lexicographic design, this research has provided the first empirical description of this hybrid communication, identifying the core English lexicon and the phrasal patterns that govern its use by Indonesian Engine Room Crew.

The primary contribution of this work is the development of an evidence-based, user-centered dictionary of English technical terms tailored for the Indonesian maritime context. This dictionary serves as a practical tool to standardize the critical vocabulary that is borrowed into daily communication, providing clear definitions in both English and Bahasa Indonesia and authentic usage examples. By codifying this language, the dictionary functions as a direct risk-mitigation tool, enhancing communicative clarity and reducing the ambiguity that can lead to accidents. It transforms linguistic research into a tangible safety asset for the Indonesian maritime industry.

Based on these conclusions, several practical recommendations are proposed. First, Indonesian Maritime Education and Training (MET) institutions are strongly encouraged to reform their ESP curricula to reflect this linguistic reality. Instruction should focus on the mastery of the core technical English lexicon for use in a Bahasa Indonesia context, rather than on general English fluency that may not be immediately applicable for domestic operations. Second, Indonesian shipping companies should consider adopting the dictionary as a standard on-board reference to foster a uniform communication culture, thereby improving safety and operational efficiency.

While this study provides a foundational resource, future research could expand the corpus to include a greater variety of vessel types within Indonesia. The methodology can also be applied to other specialized domains, such as deck operations or cargo handling, where similar hybrid communication practices likely exist. Furthermore, future work could focus on developing digital and interactive versions of the dictionary to support the continuous professional development of Indonesian seafarers (Chapelle, 2019; Wibowo et al., 2025). Ultimately, this study serves as a targeted solution for the specific needs of the Indonesian engine room and a call for further data-driven research into the unique linguistic practices that underpin maritime safety in Indonesia.

REFERENCES

- Al Amrani, A. (2019). Challenges in English for specific purpose (ESP) teaching in multilingual contexts. *Journal of Applied Linguistics*, 12(3), 45-60.
- Anthony, L. (2018). *Introducing English for specific purposes*. Routledge.
- Arbabi, P., Far, L. M., & Taki, S. (2025). Teaching maritime English to ESP learners using a virtual learning platform. *International Journal of Linguistics and Indigenous Culture*, 3(1), 63-72.
- Basturkmen, H. (2019). *Developing courses in English for specific purposes*. Palgrave Macmillan.
- Bocanegra-Valle, A. (2021). Maritime English. In *The Routledge handbook of English for specific purposes* (pp. 315-329). Routledge.
- Chapelle, C. A. (2019). The role of technology in applied linguistics. In *The Routledge handbook of applied linguistics* (pp. 581-593). Routledge.
- Đurović, Z. (2021). Corpus linguistics methods for building ESP word lists, glossaries and dictionaries on the example of a marine engineering word list. *Lexikos*, 31, 259-282. <https://doi.org/10.5788/31-1-1647>
- Fan, S., Blanco-Davis, E., Yang, Z., Zhang, J., & Yan, X. (2020). An analysis of human factors in maritime accidents using a data-driven approach. *Journal of Maritime Engineering and Technology*, 19(4), 276-287.
- Fitriati, S. W., & Rata, I. N. (2020). A needs analysis for developing English coursebooks for seafarers. *Journal of English for Academic and Specific Purposes*, 3(2), 112-125.
- Gabedava, G., & Hu, Y. (2025). Enhancing maritime safety through linguistic analysis: a case study of communication failures in maritime accidents. *WMU Journal of Maritime Affairs*, 24(1), 89-107.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
- Ismail, K., et al. (2020). The importance of maritime English proficiency for seafarer cadets. *International Journal of Language Education*, 4(1), 54-65.
- Kim, H. (2019). Rater-based evaluation of instructional materials: A systematic review. *Language Teaching Research*, 23(5), 621-642.
- Luo, M., & Shin, S. H. (2019). A comparative study of human error in maritime accidents. *Ocean Engineering*, 186, 106110.
- McCallum, M. C., Forsythe, A. M., Raby, M., & Rothblum, A. M. (2020). *Analysis of marine casualties and pollution data*. U.S. Coast Guard Research and Development Center. (Original work published 2000)
- Nugroho, A. B., & Widiastuti, R. (2025). Unveiling maritime English communication needs for seafarers: Strategic reformation for classroom instructional design. *Maritime Technology and Research*, 7(2), 1-20. <https://doi.org/10.33175/mtr.2025.271921>
- Pangestu, A., & Usman, U. (2025). Real-world communication tasks in vocational ESP coursebooks. *Journal of Vocational Language Learning*, 4(1), 22-35.
- Rahmawati, Y., et al. (2020). The role of English proficiency in seafarer recruitment. *International Journal of English Linguistics*, 10(5), 188-195.
- Sari, D. P., & Sari, F. M. (2020). Improving Indonesian seafarers' English competency: Challenges and strategies. *Indonesian Journal of Applied Linguistics*, 10(1), 204-212.
- Simbolon, N., & Yusnita, N. (2020). The need for managing proper material to teach maritime English. *Journal of English Language Teaching and Linguistics*, 5(2), 221-234.
- Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Language Science Press.
- Tchkonina, N. (2019). The role of standard marine communication phrases in preventing accidents at sea. *Journal of Maritime Transport and Engineering*, 11(2), 134-145.
- Trenkner, P. (2022). *Maritime English: A practical guide*. Springer.
- Wahyuni, S., Kusni, M., & Safitri, D. (2021). A disconnect between classroom materials and the Standard Marine Communication Phrases (SMCP) framework. *Journal of Maritime Education*, 8(1), 45-58.
- Wibowo, A. H., et al. (2025). Blended task-based language teaching for maritime English: Students' perceptions. *Language Teaching Research Quarterly*, 49, 151-169.
- Ziarati, R. (2021). *Human error in the maritime industry*. Routledge. (Original work published 2006)