MALAYSIAN JOURNAL OF APPLIED LINGUISTICS
https://e-journal.uum.edu.my/index.php/mjoal

How to cite this article:
Jati, M. S., Basthomi, Y., Grahmayanuri, N., & Andanty, F. D. (2026). "What do they really need?" The implementation of multimodality in Indonesian higher education of English for specific purposes courses. *Malaysian Journal of Applied Linguistics, 4*, 132-145. https://doi.org/10.32890/mjoal2026.4.9

"WHAT DO THEY REALLY NEED?" THE IMPLEMENTATION OF MULTIMODALITY IN INDONESIAN HIGHER EDUCATION OF ENGLISH FOR SPECIFIC PURPOSES COURSES

[1]Muhammad Suluh Jati, [2]Yazid Basthomi,
[3]Novria Grahmayanuri & [4]Ferra Dian Andanty
[1,2,3,4]Universitas Negeri Malang, Indonesia

[1]Corresponding author: muhammad.suluh.2502219@students.um.ac.id

Received: 16/9/2026 Revised: 20/9/2026 Accepted: 27/9/2026 Published: 31/10/2026

ABSTRACT

This study investigates students' needs in learning English for Specific Purposes (ESP), focusing on the students' needs and supported by visual text theory to enhance students' understanding of the course materials. The problem is that many ESP students struggle not due to limited ability but insufficient support in comprehending learning materials. Employing a descriptive needs analysis, the study involved 92 undergraduate students from a non-English department at an Indonesian university and collected data through a structured questionnaire. The study aims to identify priority language skills, topic preferences, and students' perceptions of visual elements in ESP textbooks. The findings show that students prioritize communicative skills such as speaking, listening, reading, and writing, for academic and professional purposes. However, the analysis also reveals that existing visual resources are largely conceptual and text-dominated. The study concludes that ESP textbook design should be designed with learners' communicative needs through more balanced visual - verbal integration.

Keywords: Visual text, need analysis, ESP, textbook, multimodality.

INTRODUCTION

The integration of multimodal resources in English for Specific Purposes (ESP) instruction has become increasingly significant in higher education, where learners are expected to navigate complex disciplinary discourses. This integration has pushed the boundaries of English language teaching (ELT) become more challenging, and teachers are required to adjust the curricula to meet the needs (Hafner & Ho, 2020). By doing this, teachers and students can collaborate and bring the class. Students can use multiple modes and

resources in learning the materials (Pham & Li, 2023), such as texts, images, emoticons, and others (Kohnke et al., 2021; Pham & Li, 2023).

Multimodal resources are connected with multimodal analysis, which provides a robust analytical framework for examining how meaning is constructed through the interaction of multiple modes. Seminal works by Kress and van Leeuwen (2006) and subsequent developments by Jewitt (2013) conceptualize texts as ensembles of semiotic resources, each contributing distinct but interrelated meanings. In educational contexts, multimodal elements in textbooks such as diagrams, visual representations, layout design, and specialized terminology are not merely illustrative; they are also instructional. Rather, they function pedagogically to scaffold learners' understanding of complex disciplinary concepts and to mediate access to specialized knowledge (Lavrenteva & Orland-Barak, 2023; Lim, 2018; Lim et al., 2012). The use of verbal and visual elements in textbooks also serves as the foundation to enhance students' comprehension of topics (Rafelina & Hermawan, 2022). This aligns with broader shifts in literacy education that emphasize the importance of multimodal literacy in preparing students for academic and professional communication (Hermawan, 2017; Pham & Li, 2022; Jiang & Hafner, 2025).

A growing body of research has examined multimodal texts in English as a Foreign Language (EFL) and ESP settings. Several previous studies have investigated the ideological and representational values embedded in textbooks (as in Xu & Feng, 2022), multimodal practices in high-school (as in Tafrijiyah & Andriani, 2025), discipline-specific ESP courses (as in Arno-Marcia et al., 2020), multimodal pedagogies across educational levels (as in Unsworth et al., 2022), mapped trends in multimodal research through systematic reviews (as in Tandiana et al., 2023), and highlighted diversities by using a bibliometric analysis (as in Liu et al., 2024). Collectively, these studies demonstrate that multimodality plays a significant role in shaping learners' comprehension, engagement, and meaning-making processes.

Despite these advances, the existing research has tended to focus on either linguistic content or visual design, rather than examining the dynamic interplay between verbal and visual modes within instructional materials (Yamada, 2022; Zhang, 2021). Thus, there remain spaces focusing specifically on ESP textbooks designed for undergraduate learners. Moreover, studies about verbal and visual modes were first introduced under the name visual grammar (Kress and Leeuwen, 2006) and have grown into a prominent field of study both in linguistics and English language studies. While the theory of visual grammar has been widely applied (cf. Emiliasari et al., 2025; Fitriana & Wirza, 2020), its application to ESP materials, particularly those used in non-English departments' higher education, e.g., medical, engineering, and business contexts, has been limited. This is supported by the findings from Rahmanu and Molnar (2024), which claim that multimodality has not been fully applied or integrated in higher education. In addition, Sahiruddin et al. (2025) have proven that multimodal literacy is beneficial to lecturers and students in transferring knowledge in the classroom, specifically in conjunction with digital technology. Other studies, such as those by Danielsson and Selander (2021) and Damayanti et al. (2023), emphasize the need to analyze multimodal texts within specific disciplinary and cultural contexts. These previous studies strengthened the need to delve deeper into how multimodality and ESP in this context could work together and bring up new insights between those two.

To bridge with the previous studies, this study aims to focus on the students' needs in learning English and the effectiveness of using visual texts in their textbook. Following Kress and van Leeuwen's (2006) framework, the research explores how visual elements interact to support specialized language acquisition in shaping student English skills mastery, specifically for non-English department students. In addition, the study incorporates learner perspectives to identify the specific skills required to navigate such texts effectively and to explore how students perceive the contribution of multimodal resources to their English learning process. By foregrounding the multimodal nature of ESP textbooks and situating the analysis within a higher education context, this study aims to contribute to ongoing discussions on multimodal

literacy and ESP material design. The findings are expected to offer theoretically informed and pedagogically relevant insights for instructors, curriculum developers, quality assurance units, and policymakers concerned with improving English language education at the university level. Thus, this study is formulated to answer the following research questions:

RQ1. What are the students' needs in learning English in an ESP course?
RQ2. How do visual texts help students master the material?

Theoretical Background

Multimodality within ESP

The implementation of multimodality in ESP is becoming more and more crucial in higher education, particularly as students must access, comprehend, and create multimodal texts that seamlessly combine linguistic, visual, and layout components. A crucial method for analyzing how meaning is created through the interaction of different modes, including language, images, diagrams, typography, and other spatial aspects within ESP instructional materials, is multimodality, which is based on a social semiotics framework. According to Kress and van Leeuwen (2006) and Jewitt (2013), multimodality is more than just a decorative element; rather, it functions as pedagogical and cognitive scaffolding that successfully supports the acquisition of knowledge in particular fields (disciplinary literacy).

In the context of education, multimodality has transformed and influenced the way knowledge is transmitted and transferred (cf. Beltrán-Palaques, 2024; Lim & Tan-Chia, 2022). Another important research has been conducted by Floyd and Castek (2020). They examined students' learning and drew a conclusion that the integration of multimodal text can enhance students' socioemotional learning and help overcome stress. This means multimodality plays important roles not only as a pedagogy but also helps students understand better as well as cool down their emotional states.

In the context of English for Specific Purposes (ESP), multimodality has increasingly been recognized as a central component in shaping how students construct meaning and engage with disciplinary knowledge. Rather than relying solely on traditional text-based instruction, multimodal approaches integrate visual, auditory, spatial, and digital resources that allow learners to interpret and negotiate meaning through multiple semiotic channels (Kress, 2010). This orientation aligns closely with contemporary ESP pedagogy, which emphasizes not only the development of linguistic competence but also the cultivation of communicative participation within authentic professional and academic practices. In Berezina's (2025) work, she said that ESP instruction should move from merely language assessment tools toward active engagement, enabling students to articulate ideas and respond to specific interdisciplinary scenarios. Therefore, it aligns with the purpose of this research that there's an initial urgency, perhaps, to understand the exact types of teaching materials for teaching ESP.

Students' Needs Analysis on ESP

In the context of mapping out student's need on ESP course, a needs analysis was the best option. Needs analysis has been viewed as an effective medium to understand students' needs. A study by Mao and Zhou (2024) stated that needs analysis is important for covering various aspects and effectively changing students' interests. The need to draw what students really need by using this was also viewed to seek the awareness of need (Hutchinson & Waters, 1991). This means when students know and are aware of their need in learning English, they will understand the materials more easily. To understand their needs fully, the analysis should cover the target needs, including how necessary to learn English, what the current lacks are, and what they really want to learn in an English course.

The Student's Textbook

During the learning process and classroom activities, students need textbooks containing all the information being discussed so they can do it independently at home. Meinawati (2021) added that textbooks are beneficial for students and teachers because they are one of the content sources for teaching and learning processes. Textbooks are designed to help students meet their needs and become teachers' controlling devices to align with the learning program outcomes (Cunningsworth, 1995 in Li et al., 2025). In the context of ESP, specifically in Indonesian universities, textbooks are getting less attention.

However, some studies have analyzed the use of textbooks in ESP contexts. The integration of textbooks and ESP generally showed that multimodal elements, such as images, diagrams, layout, color, and field-specific symbols, play a critical pedagogical function in bridging students' access to new disciplinary knowledge (Hermawan, 2017; Serafini, 2014). However, many previous studies have only focused on linguistic content or visual design separately and have less discussed the interactive nuances between the two modes (Yamada, 2022; Zhang, 2021). By looking at the rising trends in ESP and the growing interest in multimodality, it is necessary to design a specific textbook based on students' needs. The term 'specific' here means that the textbook should be aligned with the disciplinary study students undertake. By doing so, teachers can elevate students' knowledge comprehensively and maximize their skills upon learning English.

METHODOLOGY

To get a vivid and holistic picture of the phenomenon, the present study employed a descriptive approach using a needs analysis to investigate students' skill needs and perceptions regarding multimodal elements in an ESP textbook (Ary et al., 2012). Due to its nature, the data collected for this study are in the form of personal comments and descriptions of respondents' answers qualitatively (Fraenkel & Wallen, 2009). Moreover, the data were gathered through Google Form, which was distributed via Google Classroom. Each participant took approximately 15-30 minutes to complete the form. Considering the time to complete the form and students' daily workload, the form has been extended from three days to seven days prior to those conditions. Therefore, the form is accessible to all participants starting from November 10th to 16th, 2025.

Participants in this study were undergraduate students in the Communication Science department at a reputable university in Malang, Indonesia. The initial target population consisted of three classes totaling 102 students. After the research procedure was announced and explained, 10 students chose to withdraw from participation for various reasons, resulting in 92 valid respondents included in the data analysis. Then, the selected participants were given a questionnaire to know their needs in learning English and after taking the English course this semester. The participants were required to answer a questionnaire in the form of a Google Form. The questionnaire uses a Likert scale to draw the students' needs upon learning the ESP course, whether they are strongly agreed, agreed, disagreed, or strongly disagreed (Ary et al., 2012). There are 20 structured questions divided into three sub-sections. Structured questions are useful in their nature of deciding 'frequencies of response' (Cohen et al., 2000). In the first section, participants are required to answer the purpose of the study related questions. The second section is focused on skills needed to learn, and the last section is focused on themes and topics they wish to learn.

All data collected was analyzed in two phases. The first phase is categorizing their needs based on the questionnaires given. The categorization is rooted in the six English skills: speaking, listening, reading, writing, grammar, and vocabulary building. By looking at each skill, the result is designed specifically to map out which skills are the most important based on students' perspectives and needs. Then, in the second

phase, this study tried to analyze and compare the visual text based on the students' needs in their class handout. By incorporating the visual grammar theory, the output is expected to map out the efficiency of the visual mode in the students' book.

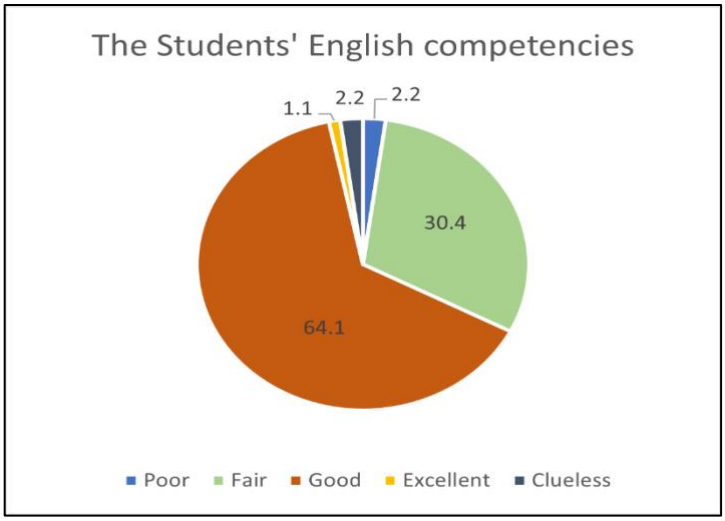
RESULTS

Identifying Students' Needs

The data indicate that, out of the 92 participants, female students outnumbered male students. Specifically, 65 respondents (70.7%) were female, while 27 respondents (29.3%) were male. When asked about prior experience attending English courses or training outside the university, 52.2% of the participants reported that they had never taken any English courses beyond campus. This suggests that the ESP course constituted their first formal exposure to English learning at the university level. To gain a clearer picture of the students' English proficiency, participants were asked to conduct a self-assessment using a self-check form. The results, as presented in Figure 1, show that 64% of the participants (59 students) rated their English proficiency as good, while only 1% reported an excellent level. These findings indicate that most participants possess a basic to moderate foundation in English; however, they still require structured instructional support to enhance their proficiency and build greater confidence in language use.

Figure 1

The Students' English Competencies



In addition, 30% of the participants (28 students) assessed their English ability as fair compared to their peers. Overall, the data suggest that students cluster around two main proficiency levels, namely good and fair. Despite this variation, there was a strong consensus among participants that learning English is both important and relevant for university students. Further questionnaire results reveal that students' primary goals for learning English are closely related to communicative competence and future career development. Most participants emphasized the importance of English as a tool for academic and professional communication, as well as a key asset for their future careers. In contrast, goals related to pursuing higher academic degrees and motivations driven by personal interest elicited more varied responses. This pattern indicates that instrumental motivation is more prominent than affective motivation among the participants.

Alongside these learning goals, the data also highlight clear challenges faced by students in learning English. The most salient difficulties are associated with productive and receptive skills, particularly

speaking, understanding spoken English, and writing sentences accurately. Participants also reported difficulties with pronunciation and reading comprehension. These findings suggest that students' challenges extend beyond basic linguistic knowledge and involve the application of English in more complex communicative contexts. These difficulties are further compounded by limited vocabulary knowledge and weak grammatical competence, which were identified as significant barriers to effective language use. Nevertheless, some participants indicated that their understanding of grammar was relatively stronger than their communicative skills and, in some cases, stronger than other language components. This pattern suggests that many students enrolled in the ESP course possess predominantly passive English competence, characterized by partial grammatical awareness that is not yet fully translated into effective spoken or written communication.

Students' Challenges in English Learning

Regarding reading skills, most participants reported that they did not experience major difficulties. Only a small proportion indicated that they often struggled with reading, particularly when encountering complex academic texts. This suggested that students experienced benefits from increased exposure to structured reading activities aimed at strengthening textual comprehension and familiarity with academic discourse.

In contrast, writing skills exhibited a markedly different pattern. The number of respondents who reported experiencing difficulties increased substantially, with approximately 48 participants indicating challenges in writing. This finding suggested that writing posed greater cognitive and linguistic demands than reading, as students were required to control grammatical structures, select appropriate vocabulary, organize ideas coherently, and transform those ideas into written form. The complexity of these simultaneous demands has made writing skills one of the most challenging skills for ESP learners.

For speaking and listening skills, responses were relatively distributed between the categories rarely and often. In speaking, around 42 participants experienced difficulties, while in listening, approximately 40 participants indicated challenges in understanding spoken English. This pattern highlighted that oral skills presented considerable challenges, as effective communication required learners to produce language and process incoming information in real time. These difficulties were commonly associated with limited fluency, slow processing and response speed, and problems understanding different accents or rapid speech.

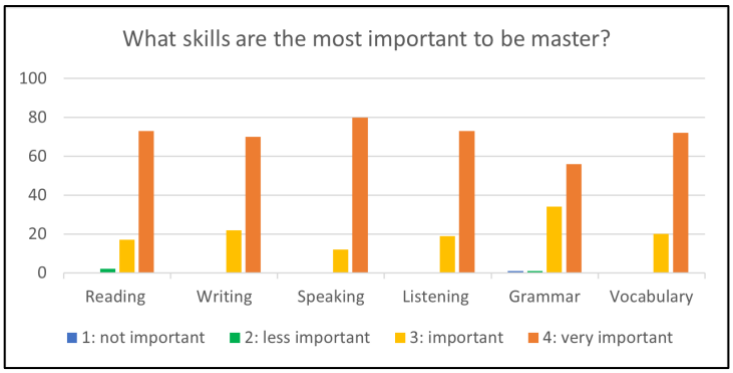
Shockingly, the most significant obstacle identified among students concerned with grammar skills. Many participants reported experiencing difficulties *often* (approximately 56 participants) or *very often* (around 17 participants). Notably, none of the participants selected *never*, underscoring that grammar constitutes a persistent and widespread source of difficulty. This challenge was further intensified by institutional requirements that mandate proof of English proficiency, often in the form of standardized certification like TOEFL, as a prerequisite for the graduation process. For students from non-language departments, this requirement tends to amplify anxiety and reinforce the perception of grammar as particularly difficult.

Skills Needed to Improve

In this section, participants were asked to indicate which language skills they wished to develop through the ESP course. Six skills were included, such as reading, writing, speaking, listening, grammar mastery, and vocabulary development, rated on a four-point Likert scale ranging from not important to very important. The results show that all six skills were perceived as important to very important by most respondents, indicating that students regard each language component as relevant in the context of ESP learning (see Figure 2).

Figure 2

Students' Desired Skills Mastery



More specifically, speaking, listening, and reading emerged as the highest priorities, as reflected in the large proportion of very important responses. Approximately 80 and 73 participants identified these skills as central and interrelated in learning English. For speaking, activities such as public speaking were highlighted as particularly desirable, as they were perceived to enhance confidence. In listening and reading, participants expressed strong interest in understanding spoken language in media contexts (e.g., social media, YouTube, films) and developing comprehensive and fast reading skills. These preferences reflect students' orientation toward communicative needs in both academic and professional settings, such as presentations, classroom discussions, and institutional interactions. The prominence of these skills suggests that students view English primarily as a functional tool for communication, rather than merely an object of study, which aligns with the core principles of ESP.

At the same time, writing and vocabulary development were also considered important. A total of 67 participants emphasized the need to improve these skills, particularly for academic writing tasks such as essays, practicum reports, journal articles, and theses. Students perceived writing and vocabulary as closely connected, noting that expanded lexical knowledge supports idea development, word choice, and sentence construction. Interestingly, some participants also described writing as a form of personal reflection or self-healing, suggesting that English learning is valued not only for academic purposes but also as a means of personal expression. This finding reinforces the continued importance of writing and vocabulary in ESP, especially in supporting academic success and professional growth.

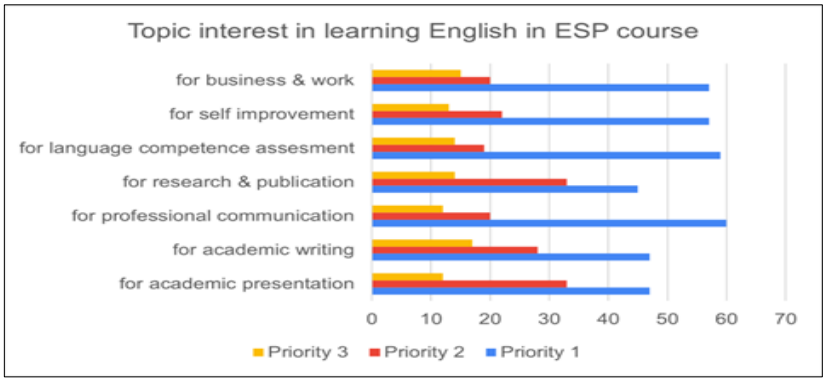
In contrast, grammar tended to be positioned as a supporting skill. Responses varied, indicating that students view grammar primarily to achieve accuracy and clarity in communication rather than as a goal. While many participants expressed a desire to write grammatically correct texts and use grammar appropriately in context, others reported anxiety about making grammatical errors. Overall, grammar was perceived as instrumental in supporting effective communication. These findings indicate that ESP students prioritize skills that directly support academic and professional communication, particularly speaking and listening in formal and institutional contexts. This pattern underscores a strong instrumental orientation toward language learning, where the emphasis lies on using English effectively in discipline-specific and real-world situations rather than on mastering linguistic systems in isolation.

Students' Proposed Topic Interest

Based on the data presented in Figure 3, students demonstrate a strong needs orientation toward the use of English for academic and professional purposes. To identify topic preferences, participants were asked to rank learning topics into three priority levels. Priority 1 represents the topics students strongly wish to learn, Priority 2 indicates secondary preferences, and Priority 3 refers to topics students would like to study after their first and second priorities have been addressed.

Figure 3

Topic Interest in ESP Course



Topics ranked as Priority 1 were dominated by English for professional communication, English for language competence assessment, and English for self-improvement. The strong interest in professional communication indicates that students require practical skills such as CV writing, formal email correspondence, and job interview preparation. Meanwhile, the high priority placed on language competence assessment reflects an instrumental need to prepare for standardized tests such as TOEFL and IELTS, which are commonly required for academic progression and professional opportunities. Interest in self-improvement, particularly in developing speaking fluency, suggests that students also view English as a means of continuous personal development. At the Priority 2 level, English for research and publication and English for academic presentation showed a noticeable increase in interest. This indicates that students seek to develop skills related to conducting surveys and field observations, understanding academic ethics, using anti-plagiarism tools, and participating in academic seminars and conferences. These topics appear to be perceived as advanced competencies that become relevant once foundational communicative skills have been strengthened. Finally, English for business and work was most frequently placed under Priority 3, suggesting that students view business-related English as a contextual need that becomes more relevant after entering the workforce. Overall, this prioritization pattern confirms that students approach ESP learning pragmatically, focusing first on competencies that directly support their immediate academic demands and near-future professional goals.

Visual Text in Textbook

Students' Perspectives of Visual Text in Textbooks

Overall, the findings indicate that students show a strong preference for visual representations that are informative, engaging, and supportive of content comprehension, with relatively balanced preferences across several major categories. The most preferred type of visual is illustrations, cartoons, and comics, selected by 34.8% of respondents. The dominance of this category suggests that students perceive narrative-based visual illustrations as effective tools for simplifying concepts, enhancing the attractiveness of learning modules, and reducing cognitive load when engaging with English learning materials. This preference highlights students' need for materials that are not only informative but also engaging and communicative, particularly within the ESP learning context.

Realistic photographs ranked second, preferred by 22.8% of respondents, followed closely by diagrams, tables, and infographics at 21.7%. The strong interest in these visual types indicates that students value visuals that are representational and explanatory in nature. Realistic photographs are perceived as providing authentic contextualization, while diagrams and infographics assist in presenting complex

information in a concise, systematic, and accessible manner, especially for academic and professional content. Meanwhile, colorful design elements were selected by 19.6% of respondents, suggesting that visual aesthetics remain an important consideration in textbook design. Visually appealing layouts are believed to enhance learning motivation and reading comfort. In contrast, screenshots received minimal preference, and maps were rarely selected. This pattern suggests that highly technical or narrowly contextualized visuals are considered less relevant to ESP learning needs within textbook-based instruction.

Visual Text in ESP Textbook

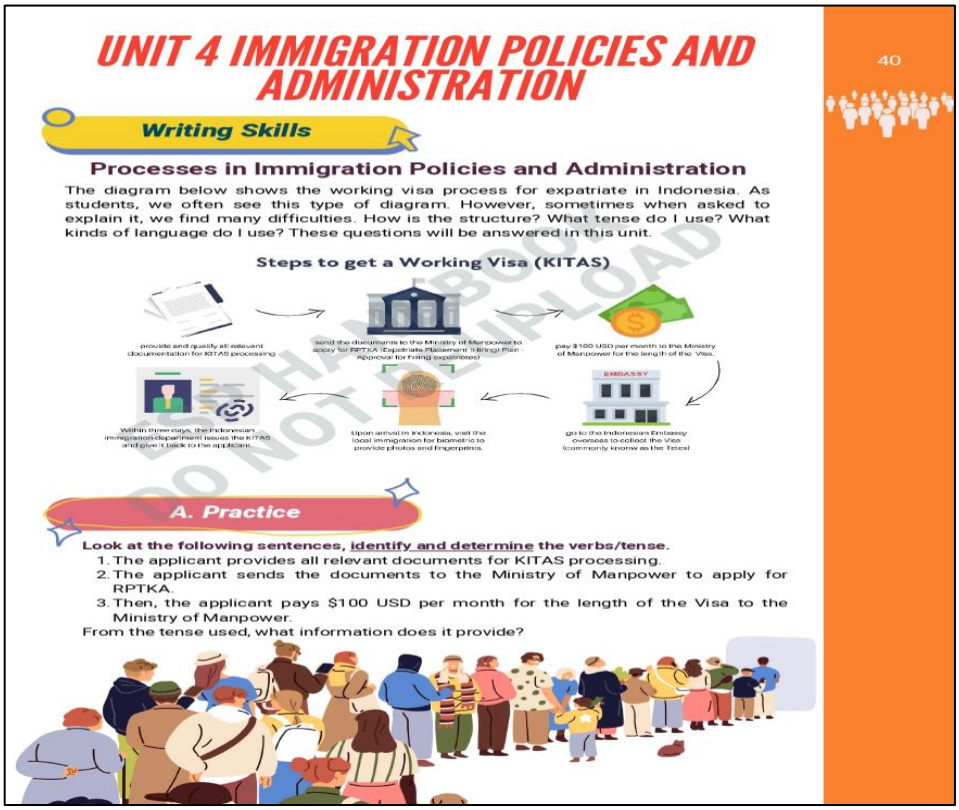
Kress and van Leeuwen (2006) introduced the theory of visual grammar, which focuses on three metafunctions: representational meaning, interactive meaning, and compositional meaning. Drawing on this framework, the present analysis examines recurrent visual patterns across the students' textbook (e.g., diagrams, photographs, charts, and layouts) to explain how visual resources contribute to—or constrain—meaning-making within the textbook.

Representational Meaning

From a representational perspective, the textbook predominantly employs conceptual–analytical representations rather than narrative ones. Visuals such as process diagrams (see Figure 4) present disciplinary knowledge as systems of ordered stages and relations, emphasizing classification, sequencing, and part–whole structures. In these visuals, participants are not depicted as human agents performing actions but are instead abstracted into icons, labels, or generic symbols. The absence of vectors indicating action and the lack of identifiable actors suggest a preference for non-transactional processes, in which meaning is constructed through logical relations rather than social interaction.

Figure 4

Example of Diagram in Student Textbook



This design choice positions knowledge as institutional and decontextualized, aligning with the goals of ESP to convey procedural clarity and conceptual accuracy. However, it simultaneously reduces opportunities for learners to engage with the experiential dimension of communicative practices. As noted in recent multimodal research, conceptual representations support cognitive mapping but often limit learners' understanding of how knowledge is enacted in real-world contexts (Jewitt, Bezemer, & O'Halloran, 2025). In ESP materials, such abstraction may be effective for introducing disciplinary frameworks, yet it may be insufficient for fostering communicative competence, which requires sensitivity to roles, agency, and situational dynamics (Li, Gibbons, & Pham, 2023). Moreover, the scarcity of narrative visuals where human participants interact within recognizable contexts suggests that the textbook frames communication primarily as a structured system rather than a social practice. This framing has implications for students' ability to transfer knowledge from the classroom to professional or academic settings, where meaning is negotiated through interaction. Thus, while representationally coherent, the visual design privileges systemic knowledge over situated practice.

Interactive Meaning

Interactive meaning concerns how visuals establish relationships between depicted participants and viewers. Across the textbook, photographs and illustrations generally adopt an impersonal interactive stance. Human figures, when present (e.g., images of volunteering activities), rarely establish direct eye contact with the viewer. This absence of gaze constitutes an offer rather than a demand, positioning images as sources of information rather than invitations to engage. As a result, the viewer is constructed as an observer, not a participant.

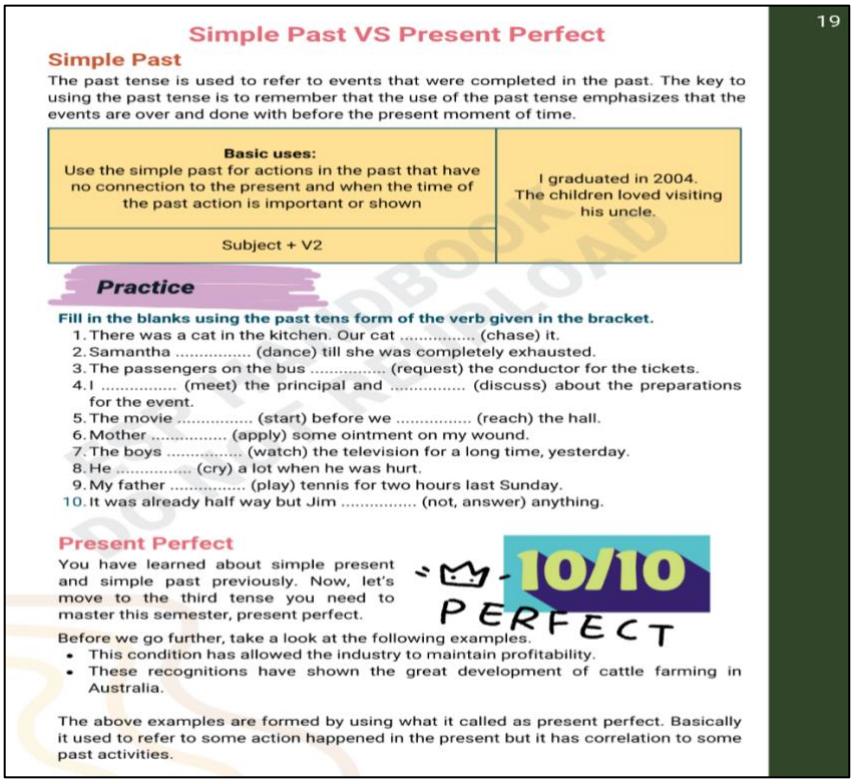
In terms of social distance, most images are framed in medium to long shots, which create a formal and detached relationship between viewers and represented participants. Combined with neutral, eye-level perspectives, this distance reinforces objectivity and aligns with academic conventions. The modality of the image, characterized by realistic color palettes and naturalistic settings, supports credibility but does not compensate for the lack of interpersonal engagement. As Djonov and Zhao (2014) argue, educational visuals that maintain strict neutrality often prioritize informational clarity at the expense of dialogic interaction. In the present textbook, this results in a visual environment that informs without inviting. While appropriate for conveying factual content, such an approach may not sufficiently motivate learners or support the development of interpersonal communication skills that are central to ESP.

Compositional Meaning

The textbook exhibits a highly structured and hierarchical layout, with verbal text occupying a dominant position. Visuals are typically placed at the top or center of the page as contextual anchors *(Given information)*, while explanatory text and exercises appear below *(New information)* (see Figure 5). This layout reflects a linear instructional logic, guiding learners from overview to elaboration.

Figure 5

Visual Order (given and new information)



Salience within the visual design is relatively low. The layout primarily uses a white background with black font; contrast is minimal, and visual elements rarely exceed textual components in size or prominence. Consequently, visuals function mainly as illustrative support rather than as independent meaning-making resources. Framing is strong, with clear borders, tables, and segmented sections that tightly control the flow of information. Such framing enhances clarity and reduces ambiguity, which is advantageous for instructional purposes. However, the dominance of text over visuals indicates that the textbook has not fully embraced multimodal pedagogy, in which images, diagrams, and layout actively shape interpretation. In this textbook, conservative compositional choices ensure readability but limit the pedagogical potential of visual resources to support deeper engagement and multimodal meaning-making.

CONCLUSION

This study demonstrates that ESP students in higher education exhibit a strong needs orientation toward the use of English for academic and professional purposes, and it could be boosted if their textbook is designed by employing visual text. Although most students fall within a mid-level proficiency range, they continue to face substantial challenges, particularly in productive and interactive skills such as speaking and writing, as well as in grammar and vocabulary use. These findings suggest that students' English competence remains largely passive, with linguistic knowledge not yet fully realized in authentic academic and professional communication. The needs analysis further reveals that students prioritize skills that support functional and contextual communication, especially speaking and listening in formal and institutional settings, followed by reading and writing for academic purposes. Grammar is generally positioned as a supporting skill that serves to ensure accuracy and clarity rather than as a primary learning objective. This pattern highlights a predominantly instrumental orientation to ESP learning, consistent with the discipline-specific and use-oriented nature of ESP pedagogy.

From a multimodal perspective, students express a clear preference for visual texts that are informative, engaging, and communicative. However, the visual analysis of the ESP textbook indicates a partial misalignment between students' needs and existing visual design practices. Drawing on Kress and van Leeuwen's visual grammar framework, the textbook visuals are largely characterized by conceptual–analytical representations, impersonal interactive relations, and compositional arrangements that privilege verbal text as the primary source of meaning. As a result, the potential of visual elements to enhance engagement, contextualization, and multimodal literacy has not been fully realized.

Last but not least, this study underscores the value of integrating needs analysis by means of evaluating and developing ESP learning materials. Pedagogically, the findings point to the importance of designing ESP textbooks that are more closely aligned with learners' communicative needs, through greater use of narrative visuals, representations of authentic social practices, and more balanced visual–verbal integration. Extending this line of inquiry to a wider range of disciplinary contexts and institutional settings, and adopting complementary methodological approaches, may further deepen understanding of how visual texts can support effective and sustainable ESP learning.

DECLARATION OF COMPETING INTEREST

There are no competing interests among authors.

DECLARATION OF GENERATIVE AI USE

This article uses generative AI, Grammarly, to improve the clarity of sentences and check for grammatical errors.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Arnó-Macià, E., Aguilar-Pérez, M., & Tatzl, D. (2020). Engineering students' perceptions of the role of ESP courses in internationalized universities. *English for Specific Purposes, 58*, 58-74. https://doi.org/10.1016/j.esp.2019.12.001

Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. A. (2012). *Introduction to research in education* (9th ed.). Cengage Learning.

Beltrán-Palanques, V. (2024). Assessing video game narratives: Implications for the assessment of multimodal literacy in ESP. *Assessing Writing, 60*, Article 100809. https://doi.org/10.1016/j.asw.2024.100809

Berezina, Z. I. (2025). Content-based instruction for psychology majors: Issues and solutions for ESP curriculum development. Training, *Language and Culture, 9*(3), 99–112. https://doi.org/10.22363/2521-442X-2025-9-3-99-112

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). Routledge.

Damayanti, I., Febrianti, Y., Suryatama, K., Dewi, F., & Lubis, A. H. (2023). "Is Kancil kind?": Exploring the interplay of verbal and visual texts in a picture book for teaching English to young learners. *TEFLIN Journal*, *36*(1), 21–43. https://doi.org/10.15639/teflinjournal.v36i1/21-43

Danielsson, K., & Selander, S. (2021). *Multimodal texts in disciplinary education*. Springer.

Djonov, E., & Zhao, S. (2014). *Critical multimodal studies of popular discourse*. Routledge.

Emiliasari, R. N., Emilia, E., & Hermawan, B. (2025). Integrating SFL and visual grammar in genre-based pedagogy to foster EFL students' critical reading of multimodal texts. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya, 11*(2), 626–639. https://doi.org/10.22219/kembara.v11i2.41286

Fitriana, W., & Wirza, Y. (2020). An analysis of multimodal text in EFL textbooks of secondary school in Indonesia in assisting students' text understanding. *Advances in Social Science, Education and Humanities Research*, *546*. https://doi.org/10.2991/assehr.k.210427.095

Floyd, R., & Castek, J. (2020). Academic, emotional, and social growth in the second language classroom: A study of multimodality. In C. McCray (Ed.), *Multifaceted strategies for social-emotional learning and whole learner education* (pp. 163–188). IGI Global.

Fraenkel, J. R., & Wallen, N. E. (2009). *How to design and evaluate research in education* (7th ed.). McGraw-Hill.

Hafner, C. A., & Ho, W. Y. J. (2020). Assessing digital multimodal composing in second language writing: Towards a process-based model. *Journal of Second Language Writing, 47*, Article 100710. https://doi.org/10.1016/j.jslw.2020.100710

Hermawan, B. (2017). Multimodality: Menafsir verbal, membaca gambar, dan memahami teks. *Jurnal Pendidikan Bahasa & Sastra*, *13*(1). https://doi.org/10.17509/bs_jpbsp.v13i1

Hutchinson, T., & Waters, A. (1991). *English for specific purposes* (6th ed.). Cambridge University Press.

Jewitt, C. (2013). Multimodal methods for researching digital technologies. In S. Price, C. Jewitt, & B. Brown (Eds.), *The SAGE handbook of digital technology research* (pp. 250–265). SAGE Publications.

Jewitt, C., Bezemer, J., & O'Halloran, K. (2025). *Introducing multimodality* (2nd ed.). Routledge.

Jiang, L., & Hafner, C. (2025). Digital multimodal composing in L2 classrooms: A research agenda. *Language Teaching*, *58*, 528–546. https://doi.org/10.1017/S0261444824000107

Kohnke, L., Jarvis, A., & Ting, A. (2021). Digital multimodal composing as authentic assessment in discipline-specific English courses: Insights from ESP learners. *TESOL Journal.* https://doi.org/10.1002/tesj.600

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Routledge.

Lavrenteva, E., & Orland-Barak, L. (2023). Conceptual-analytical framework for exploring culture in EFL coursebooks: Analysis of teaching materials from a multimodal perspective. *Social Sciences & Humanities Open*, *7*(1), Article 100441. https://doi.org/10.1016/j.ssaho.2023.100441

Li, M., Gibbons, J., & Pham, Q. N. (2023). Re-exploring writer-reader interaction: Analyzing metadiscourse in EAP students' infographics. *Journal of English for Academic Purposes*, *66*, Article 101303. https://doi.org/10.1016/j.jeap.2023.101303

Li, Y., Darmi, R. H., & Thai, Y. N. (2025). Systematic literature review on evaluation of English language textbooks: A decade of research. *World Journal of English Language*, *15*(3), 65–78. https://doi.org/10.5430/wjel.v15n3p65

Lim, F. V. (2018). Developing a systemic functional approach to teach multimodal literacy. *Functional Linguistics*, *5*, Article 13. https://doi.org/10.1186/s40554-018-0066-8

Lim, F. V., O'Halloran, K. L., & Podlasov, A. (2012). Spatial pedagogy: Mapping meanings in the use of classroom space. *Cambridge Journal of Education, 42*(2), 235–251. https://doi.org/10.1080/0305764X.2012.676629

Lim, F. V., & Tan-Chia, L. (2022). *Designing learning for multimodal literacy: Teaching viewing and representing*. Routledge.

Liu, H., Liu, L., & Li, H. (2024). Multimodal discourse studies in the international academic community (1997–2023): A bibliometric analysis. *SAGE Open, 14*(4). https://doi.org/10.1177/21582440241305454

Mao, F., & Zhou, J. (2024). A needs analysis of ESP courses in colleges of art and design: Consensus and divergence. *PLOS ONE*, *19*(6), Article e0305210. https://doi.org/10.1371/journal.pone.0305210

Meinawati, S. (2021). ESP textbook: What the users' need? *Jurnal Pendidikan Humaniora*, *9*(2), 77–84.

Rafelina, D. P., & Hermawan, B. (2022). The ideational meanings of multimodal texts presented in an EFL textbook. *Teaching & Learning English in Multicultural Contexts*, *6*(2).

Rahmanu, I. W. E. D., & Molnar, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon, 10*(19), Article e38357. https://doi.org/10.1016/j.heliyon.2024.e38357

Sahiruddin, S., Sudarwati, E., Lestari, I. P., & Daniswara, R. A. (2025). Digital multimodal literacy model of English language education in Indonesian higher education. *JEELS, 12*(2), 615–640. https://doi.org/10.30762/jeels.v12i2.2612

Serafini, F. (2014). *Reading the visual: An introduction to teaching multimodal literacy*. Teachers College Press.

Tafrijiyah, K., & Andriani, M. (2025). Multimodal discourse analysis of an English language learning textbook in Indonesia. *Lire Journal (Journal of Linguistics and Literature), 9*(1), 61–79. https://doi.org/10.33019/lire.v9i1.391

Tandiana, S. T., Abdullah, F., Andriani, A., Hidayati, A. N., & Rosmala, D. (2023). Mapping multimodal teaching materials for Indonesian EFL students: A need analysis. *Vision - Journal for Language and Foreign Language Learning*, *12*(2), 147–168. https://doi.org/10.21580/vjv12i219990

Unsworth, L., Tytler, R., Fenwick, L., Humphrey, S., Chandler, P., Herrington, M., & Pham, L. (2022). *Multimodal literacy in school science*. Routledge.

Xu, Y., & Feng, D. (2022). Values in business English textbooks: A multimodal analysis approach. In T. Xiong, D. Feng, & G. Hu (Eds.), *Cultural knowledge and values in English language teaching materials*. Springer. https://doi.org/10.1007/978-981-19-1935-0_2

Yamada, M. (2022). The role of multimodality in ESP textbook design. *Journal of English for Academic Purposes*, *57*, Article 101072. https://doi.org/10.1016/j.jeap.2022.101072

Zhang, J. (2021). Multimodal pedagogy in ESP: A review of recent trends. *Journal of Language Teaching and Research*, *12*(4), 678–686. https://doi.org/10.17507/jltr.1204.XX