Journal of English Language and Education
ISSN 2597- 6850 (Online), 2502-4132 (Print)
Journal Homepage: https://jele.or.id/index.php/jele/index

OPEN ACCESS

Article

Evaluating the Alignment of Language Content and Illustrations in an Indonesian Grade XI English Textbook: A Paradigmatic–Syntagmatic and Multimodal Analysis

https://doi.org/10.31004/jele.v11i6.3330

Risa Anggraini, Anni Holila Pulungan[ab]
[1]Universitas Potensi Utama, Indonesia
[2]Universitas Negeri Medan, Indonesia
Corresponding Author: risa.anggraini005@potensi-utama.ac.id

A B S T R A C T

This study evaluated the alignment between language content and illustrations in an Indonesian Grade XI English textbook, Bahasa Inggris SMA/MA/SMK/MAK Kelas XI by Mahrukh Bashir (revised edition, second printing, 2017), using Cunningsworth's design-and-organization criterion as the evaluative reference. Alignment was examined on two separate dimensions. Verbally, drawing on Halliday's system–structure distinction, each language-focused block or task was coded as paradigmatic-isolated (alternatives without a structural frame), syntagmatic-input-only, or integrated, with integration further classified as contiguous (paradigm and frame in the same table or on the same page), formulaic (each option a complete utterance), or noncontiguous (paradigm and frame separated by intervening pages). Visually, each image–text pairing was coded following Martinec and Salway for status (equal or unequal) and logico-semantic relation (elaboration, extension, or enhancement). Three chapters were selected purposively to represent spoken functional expressions and a written genre under three image conditions: no illustration (Chapter 1), drawn characters (Chapter 2), and an authentic photograph (Chapter 4). Chapter 1 integrates all its paradigms contiguously in tables that map alternatives onto sentence slots. Chapter 2 provides structural frames for only some paradigms, but its characters' gestures, postures and colour-coded speech bubbles elaborate or extend the speech acts being taught. Chapter 4 presents seven lexical paradigms in isolation, realizes four of them noncontiguous, and pairs its reading with a photograph that only enhances the text's factual claim. The textbook therefore applies verbal and visual support unevenly, and revision should target the paradigm–structure and image–task alignment of its written-genre chapter.

Keywords: *Coursebook Evaluation, Multimodal Analysis, Paradigmatic Relations, Systemic Functional Linguistics, Syntagmatic Relations*

Article History:
Received 04th October 2026
Accepted 14th November 2026
Published 16th November2026

Check for updates

INTRODUCTION

English textbooks issued by the Indonesian Ministry of Education and Culture are compulsory student books, so decisions about how they present language reach learners across the country. One such decision concerns how vocabulary and functional expressions are presented. A list of expressions tells learners which items they may choose, but it does not show how each item combines with other words to form an acceptable sentence. The Grade XI textbook examined here, for example, lists Let's … and how about …? as alternative ways of making a suggestion, yet the first is followed by a bare verb (Let's stay at home) and the second by an -Ing form (How about staying at home?). A learner who meets only the list has no basis for knowing this, and is likely to produce forms such as *How about stay at home? Whether a textbook pair each set of choices with the structures that realize it, and whether its illustrations make the intended meaning visible, is therefore a practical question for English language teaching and for the design and revision of textbooks.

This concern is supported by research on vocabulary and formulaic language in EFL materials. In an analysis of three EFL teaching materials used in Swedish primary schools,

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Nordlund & Norberg (2025) found that much of the formulaic language in the texts was neither highlighted in the glossaries nor practiced in the exercises. In Jordanian junior high school textbooks, Eid & Al-Jamal (2023) found that collocations made up only a small proportion of the vocabulary items, that single words were prioritized over word combinations, and that few activities practiced collocations. In Indonesia, Oktavianti & Budiwati (2026) analyzed lexical bundles in three senior high school English textbooks and recommended that such multi-word units be taught explicitly. The first two studies show that EFL materials often present lexis as single items rather than in the combinations in which it is used, and the third shows that this concern applies to Indonesian senior high school textbooks. However, they examine multi-word sequences quantitatively across whole books; they do not examine whether the functional and lexical paradigms that a textbook explicitly teaches are presented together with the sentence structures that realize them.

Systemic functional linguistics (SFL) provides a theoretical basis for this question. Halliday (2013) describes meaning as choice: a speaker selects from a paradigm of alternatives, such as a set of verbs or expressions that can fill the same slot, but the selection becomes meaning only when it is realized in a syntagm, a sequence of elements combined into a structure. The two axes constrain each other. The verbs suggest, recommend and propose belong to the same paradigm precisely because they share a syntagmatic frame (I suggest that he clean his room), while Let's and How about belong to the same functional paradigm but demand different frames. Teaching a paradigm together with its frame therefore shows learners both what they can choose and how the choice becomes a well-formed utterance; teaching the paradigm alone presents choice without meaning. The visual mode can contribute to the same process. Images realize meaning through what they represent and through the stance they construct toward the viewer (Kress & van Leeuwen, 2006), and a verbal text can anchor one reading of an image or let the image relay information the text does not state (Barthes, 1977). Martinec & Salway (2005) make these relations analyzable by describing the relative status of image and text and the logico-semantic relation between them. An image that enacts the communicative act named by a paradigm, such as a speaker gesturing while giving an opinion, supports the paradigm; an image that only documents a topic does not. Developing this kind of multimodal awareness is increasingly regarded as part of English language teaching (Hafner, 2020). In this study, therefore, meaning construction refers to three related dimensions: the selection of an item from a paradigm, the realization of that item in a syntagmatic structure, and the support that an accompanying image gives to the communicative act the item expresses.

Research on the visual side of EFL textbooks has grown in recent years. In Indonesia, Elmiana (2019) analyzed 232 pictures in three senior high school EFL textbooks and found that they largely represented Indonesian culture and the realities of 15–17-year-old learners. Purwaningtyas (2020) analyzed 158 images in an Indonesian EFL textbook using Kress and van Leeuwen's framework and recommended that publishers synchronize the written content with the visual representation to avoid misinterpretation. Jamilah et al. (2024) examined the interactive meaning of images in an English textbook for young learners, and Bashir et al. (2025) analyzed visual representation in elementary-level English textbooks from a social semiotic perspective. Other studies have examined teachers' beliefs and perceptions about multimodal literacy (Trisanti et al., 2022). Therefore, these studies show that textbook images carry cultural, social and interactive meanings and that teachers regard multimodality as relevant to language teaching. What they have not yet examined is whether an image supports the specific language point it accompanies, and whether this visual support is related to how the same language point is presented verbally.

To evaluate the textbook rather than merely describe it, this study draws on Cunningsworth (1995) framework for evaluating coursebooks, which covers aims and approach, design and organization, language content, skills, topic, methodology, teachers' books, and practical considerations. Most of these criteria judge what a book teaches or how it teaches it. The design-and-organization criterion is the only one that concerns how the content is arranged and presented on the page, including its layout and visual material, and whether

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

this presentation supports the language content rather than decorating it. Because the present study asks how language items are arranged relative to their structures and to the images beside them, this criterion is the most appropriate evaluative reference. SFL and Martinec & Salway (2005) framework then serve to operationalize it: a chapter meets the criterion when its paradigms are presented with the structures that realize them and its images stand in a supporting relation to those paradigms.

The verbal and the multimodal dimensions are evaluated separately in this study because they describe different characteristics of textbook quality. The verbal dimension concerns the organization of the language content itself: whether a paradigm is presented as a usable system, with the structures that turn a choice into a well-formed utterance. The multimodal dimension concerns the semiotic design of the page: whether an image contributes meaning that bears on the language point, and in what relation. The two dimensions are related, since both fall under Cunningsworth (1995) design-and-organization criterion, but they do not necessarily vary together. A paradigm may be well integrated verbally yet unillustrated, or poorly integrated verbally yet accompanied by a congruent image. Keeping the two dimensions distinct, and then comparing them, makes it possible to identify which kind of support a chapter provides and which kind it lacks.

Previous research has examined lexis in textbooks and images in textbooks largely in separate strands. No study has examined, chapter by chapter within a single secondary-level textbook, whether paradigms are presented with their realizing structures and whether the accompanying images support those same paradigms. The aim of this study was therefore to evaluate how the Grade XI textbook Bahasa Inggris SMA/MA/SMK/MAK Kelas XI (M. Bashir, 2017) aligns paradigmatic choices with syntagmatic structures and with illustrations, and whether this alignment meets Cunningsworth (1995) design-and-organization criterion. Three research questions guided the study: (1) How are paradigmatic and syntagmatic relations presented in the functional-language, grammar, and reading-based tasks of the selected chapters? (2) What do the illustrations and photographs in these chapters represent? (3) How does their representational content relate to the linguistic point being taught?

METHOD

Research design

This study used a qualitative document analysis design, which is suited to examining how meaning is constructed in an authentic, naturally occurring text rather than in data elicited for research purposes (Kyngäs, 2020). Document analysis has been used to investigate how content is organized in school textbooks, treating the textbook's own sections, glossaries and tasks as the units of analysis (Seddighi et al., 2021). The analysis was deductive: coding categories were derived beforehand from Halliday's (2013) system–structure distinction for the verbal strand and from Martinec & Salway (2005) image–text framework for the visual strand, and the two strands were then compared against Cunningsworth (1995) design-and-organization criterion.

Within SFL, the analysis worked with the system–structure distinction at clause rank: the paradigm is the set of alternatives that can fill one slot, and the structure is the sequence of slots in which the selected item is realized. To describe the structural frames precisely, two further SFL categories were used. Mood was used to identify whether a frame realizes a declarative, an interrogative or an imperative clause (e.g., the Finite placed before the Subject in Shall I …? or the imperative Let's …), and projection was used to identify frames in which a mental or verbal process projects a clause (e.g., I suggest that he clean his room; I reckon this could be right). These categories address the first research question because the functional expressions in the sampled chapters differ precisely in their mood and projection structures. Transitivity and theme were not used, because the study asks whether a choice is presented with its realizing structure, not how experiential content or information flow is organized.

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Data Source

The data source was the official PDF of the student book Bahasa Inggris SMA/MA/SMK/MAK Kelas XI (M. Bashir, 2017). The book was written by Mahrukh Bashir, reviewed by Emi Emilia and Helena I. R. Agustien, and published by the Ministry of Education and Culture (Kementerian Pendidikan dan Kebudayaan, Kemendikbud) in Jakarta under the coordination of Pusat Kurikulum dan Perbukuan, Badan Penelitian dan Pengembangan (Balitbang). The edition analysed is the revised edition (Edisi Revisi), second printing (Cetakan ke-2, 2017) of the book first printed in 2014, with ISBN 978-602-427-106-0 (complete set) and 978-602-427-108-4 (volume 2). It is the compulsory English student book (buku siswa) for Grade XI in all four senior-secondary tracks named in its title: general senior high school (SMA), Islamic senior high school (MA), vocational senior high school (SMK), and Islamic vocational senior high school (MAK). The book was prepared for the 2013 Curriculum (Kurikulum, 2013) and consists of eight chapters: (1) Offers & Suggestions, (2) Opinions & Thoughts, (3) Party Time, (4) National Disaster – An Exposition, (5) Letter Writing, (6) Cause & Effect, (7) Meaning Through Music, and (8) Explain This! followed by an enrichment section.

Three chapters were selected purposively using three criteria. First, each chapter had to contain explicitly presented paradigms, that is, lists or tables of alternative words, expressions or language features, so that the relation between paradigm and structure could be analyzed. Second, the selected chapters had to represent the two broad types of language content in the book: spoken interpersonal functional expressions and the language features of a written genre. Third, the chapters had to show contrasting image conditions, so that the role of images could be compared. Chapter 1 (Offers & Suggestions, PDF pp. 10–14) presents functional expressions entirely through lists and tables without any illustration and therefore serves as a verbal-only baseline. Chapter 2 (Opinions & Thoughts, PDF pp. 25–29) teaches a comparable set of functional expressions but accompanies almost every paradigm with drawn characters. Chapter 4 (National Disaster – An Exposition, PDF pp. 52–57; the chapter heading in the body reads "Natural Disasters – An Exposition") teaches the language features of a written genre and pairs its reading with an authentic photograph.

Chapter 3 (Party Time) was not selected because it teaches accepting and declining invitations, an interpersonal functional paradigm of the same type as those in Chapters 1 and 2; including it would have repeated the functional-language condition already represented rather than adding a new contrast. Chapters 5 to 8 were not selected because the study prioritised depth of analysis, including the examination of every task and image in the sampled chapters, over coverage of the whole book, and the three selected chapters already represent both types of language content and all three image conditions. The findings are therefore interpreted as a documented pattern within these chapters rather than as a claim about the textbook as a whole.

Data Analysis

The units of analysis were defined separately for the two dimensions. Verbal units were every language-focused element in the three chapters: lists of vocabulary or functional expressions, structure tables, model sentences, dialogues and worked exchanges, the "Language Features" box and worked example text in Chapter 4, the reading passages, and the practice and production tasks (e.g., the "Let's Practice" sections). Visual units were every drawn illustration and photograph that accompanies a verbal unit. Within each visual unit, the represented participants and their gesture, posture, facial expression and orientation toward one another were recorded, together with the use of colour (e.g., colour-coded speech bubbles) and the relationship between the image and its adjoining text. The analysis followed five steps.

In the first step, each chapter was read in full and segmented into units, and the PDF page of each unit was recorded. In the second step, every verbal unit was coded into one of four categories derived from Halliday (2013) system–structure distinction: (1) Paradigmatic-isolated (P), a set of alternative words, expressions or forms presented without the syntagmatic frame in which they are combined; (2) Syntagmatic-input-only (S-in), syntagmatic patterning

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

present in the reading input but not targeted by any task; (3) Syntagmatic-task (S-task), a task requiring learners to process or produce a multi-word or clause-level unit without an accompanying paradigm; and (4) Integrated (I), a block or task that pairs a paradigm with the structure that realizes it or requires learners both to select and to combine. Integration was further classified by the location of the structure relative to the paradigm. Integration is contiguous (I-c) when the paradigm and its structural frame appear in the same table or on the same or facing page, with no other content between them, so that both can be read together. It is formulaic (I-f) when each member of the paradigm is already a complete utterance and is therefore its own realization. It is noncontiguous (I-n) when the paradigm and its realization both appear in the chapter but are separated by at least one intervening page or section, and the textbook does not refer from one to the other, so that the learner has to make the connection. At chapter level, verbal integration was described as strong when all of a chapter's paradigms were I-c or I-f, mixed when integrated and isolated paradigms occurred together, and weak when most paradigms were P or only I-n.

In the third step, each visual unit was coded using Martinec & Salway (2005) system of image–text relations, which builds on Barthes (1977) account of anchorage and relay and on Halliday's logico-semantic relations. Status describes whether image and text are of equal status, either independent (neither modifies the other) or complementary (together forming a larger unit of meaning), or of unequal status, where one mode is subordinate to the other. Logico-semantic relation describes how the image relates to the paradigm being taught: elaboration, where the image exemplifies or restates the communicative act the paradigm names; extension, where the image adds new but related meaning, such as an interpersonal stance shown through posture; or enhancement, where the image provides circumstantial or documentary support without depicting the communicative act. For evaluation, these categories were interpreted as follows: an image reinforces the verbal content when it elaborates it, complements it when it is of equal, complementary status, extends it when it adds related meaning, and remains unrelated to the language point when it is independent of the paradigm or only enhances the topic of the text. In the fourth step, the verbal and visual codes of each chapter were entered into a cross-chapter matrix (Table 4) and compared to identify recurring patterns and differences, which were then evaluated against Cunningsworth (1995) criterion.

In the fifth step, the trustworthiness of the coding was supported through the procedures recommended for qualitative content analysis (Kyngäs, 2020; Seddighi et al., 2021). Categories were defined operationally before coding, borderline cases were re-examined against the definitions, and the PDF page, excerpt or image description for each coding decision was retained as an audit trail.

FINDINGS AND DISCUSSION

The findings are presented according to the three research questions and then compared across chapters. In brief, the three chapters present paradigmatic and syntagmatic relations in different ways (RQ1): following the definitions given in the Method, verbal integration is strong in Chapter 1, mixed in Chapter 2 and weak in Chapter 4. Their illustrations also differ (RQ2): Chapter 1 has none, Chapter 2 uses drawn characters in interaction, and Chapter 4 uses an authentic documentary photograph. These images relate to the language point in different ways (RQ3): Chapter 2's images elaborate or extend the speech acts being taught, while Chapter 4's photograph only enhances the reading's factual claim. Table 1 summarizes the verbal coding, Tables 2 and 3 the multimodal coding, and Table 4 the cross-chapter comparison.

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Table 1. Coding Summary of Language-Focused Tasks in Chapters 1, 2, and 4

Code	Illustrative source in the textbook	Approx. instances
Integrated, contiguous (I-c)	Ch.1 suggestion paradigm (PDF p. 10) + Table 1.1 structure (PDF p. 11); offer paradigm (PDF pp. 12–13) + Table 1.4 structure (PDF p. 13); Let's Practice cloze task (PDF p. 14); Ch.2 opinion-verb paradigm + Table 2.1 structure (PDF p. 27)	4
Integrated, formulaic (I-f)	Ch.2 ready-to-use complete expressions where each paradigm member is already its own syntagmatic realization: personal/general-viewpoint stems modeled in use through a worked boy–girl exchange (PDF p. 26); agreeing/disagreeing expression sets (PDF p. 29)	3 sets
Integrated, noncontiguous (I-n)	Ch.4 Table 4.1 annotated worked exposition example (PDF pp. 56–57) colour-coding mental verbs, conjunctive relations, causal conjunctions, generic reference, expert opinion, and present tense, separated from the paradigm lists on PDF p. 55 by an intervening page	1 (covering most, not all, of p. 55's paradigms)
Paradigmatic-isolated (P)	Ch.2 personal-viewpoint expression stems, marked with trailing "…" rather than worked into full sentences (PDF p. 27); Ch.2 general-viewpoint expression stems with no accompanying structure table at all (PDF p. 28); Ch.4 "Language Features" descriptive-synonym and attitude-marker lists, not revisited anywhere in Table 4.1 (PDF p. 55)	4 lists
Syntagmatic-input-only (S-in)	Ch.4 pre-reading passage "Global Warming" (PDF p. 52), rich in evidential and causal syntagmatic patterning (e.g., "there is irrefutable evidence that…") never revisited as a language-focused target	2+ instances
Syntagmatic-task (S-task)	Post-reading discussion questions for "Global Warming" (PDF p. 52) address comprehension and text-type identification only	0 (identified as language-focused)

Paradigmatic and syntagmatic relations in the verbal tasks (RQ1)

Chapter 1 (Offers & Suggestions) presents every paradigm together with its structure. The suggestion paradigm on PDF page 10 lists expressions such as Let's …, Why don't we …? How about …? and I suggest that …. On the next page, Table 1.1 (Sentence structure to suggest something) arranges the reporting verbs suggest, recommend, propose, put forward, advise and advocate in one column and places them in the slots Subject–Verb–That–Object, with complete examples such as I suggest that he clean his room and We recommend that we work together on this. Table 1.2 (PDF p. 12) then sets out complete exchanges in three columns, for example Let's go to movies. – Yes, let's go. – No, thank you. I do not feel like going. The offer paradigm follows the same pattern: the list on PDF page 12 (May I …? Can I …? Shall I …? Would you …?) is followed on PDF page 13 by Table 1.3 (Responding to offers), with exchanges such as Shall I bring you some tea? – Yes, please. That would be lovely., and Table 1.4 (Sentence structure to offer something), which places the modal verbs in a Modal Verb–Subject–Object frame (Shall we take you there? Could I offer you something?). In SFL terms, the tables supply the structure that each paradigm requires: the reporting verbs project a clause (I suggest + that he clean his room), Let's realizes an imperative, and Shall I …? and Can I …? realize interrogatives with the Finite before the Subject. Because each paradigm appears with its realizing structure in the same table or on the facing page, all of these blocks were coded as Integrated, contiguous (I-c).

The tables are sufficient to show integration in this chapter because of how they are organized, not merely because they are tables. Each table lays the paradigmatic axis out vertically, as a column of alternatives that can fill the same slot, and the syntagmatic axis horizontally, as the sequence of slots that make up the sentence, so that reading down shows what can be chosen and reading across shows how the choice is combined. The response tables add a further syntagmatic layer by placing each expression in an exchange with an accepting and a declining reply. The integration is also carried into practice. In the "Let's Practice" task (PDF p. 14), learners complete the reply I can't, I am low on cash right now. ____ stay at home and watch TV instead by choosing among How about, Let's, what about and I think. All four

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

options belong to the suggestion paradigm, but only Let's fits the bare verb stay, so the task requires both paradigmatic selection and syntagmatic fit. At the same time, the tables show integration at the level of presentation; they say little about when each alternative is appropriate, and the present analysis cannot show whether learners actually use the frames.

Chapter 2 (Opinions & Thoughts) integrates paradigm and structure contiguously for only one of its paradigms. Table 2.1 (Sentence structure to express opinions, PDF p. 27) pairs the opinion verbs agree, believe, reckon, doubt, assume and think with a Subject–Verb–Object frame and a complete example for each (e.g., I reckon this could be right considering the reasons you have provided), which was coded as I-c. The worked opinion exchange between a boy and a girl on PDF page 26 (e.g., I like Harry Potter movies because the magic seems so real. – I don't agree with you. Harry Potter movies are just overrated.) and the sets of expressions for agreeing and disagreeing on PDF page 29 (e.g., I couldn't agree more; I am not convinced that ….) consist of complete utterances, so the paradigm and its realization are the same object; these three sets were coded as Integrated, formulaic (I-f). In contrast, the personal-viewpoint stems on PDF page 27 (In my opinion …, To my mind …, As I see it …) and the general-viewpoint stems on PDF page 28 (Most people do not agree …, Some people say that …) end in ellipses and are never completed anywhere in the chapter; both lists were coded as Paradigmatic-isolated (P).

Chapter 4 (National Disaster – An Exposition) shows the weakest verbal integration. Its "Language Features of an Exposition Text" box (PDF p. 55) lists seven paradigms in succession: descriptive synonyms (e.g., instead of bad, use appalling, ghastly, terrible), present-tense usage, mental verbs (I believe, I doubt), saying verbs (people say, research indicates), connecting words (additionally, moreover, firstly), causal conjunctions (because, consequently, despite) and attitude markers (will, usually, in all probability). None of them is accompanied by a structural frame or example sentence at the point of presentation, so all seven were coded as P. Two pages later, Table 4.1 (PDF pp. 56–57) presents a worked exposition arguing that motorbikes should be banned from housing areas, with colour-highlighted words matched to labels such as mental verbs, causal conjunctions and present tense. Four of the six labelled categories correspond to paradigms on page 55, giving those four a delayed, noncontiguous realizations (I-n); the remaining paradigms receive no label, and even ghastly, which reappears in the worked text, is left unmarked. The pre-reading passage "Global Warming: Is It an End to Our World?" (PDF p. 52) is written in the same persuasive register (e.g., there is irrefutable evidence that human activities have changed the atmosphere of our earth), but its three post-reading questions address only comprehension and text type, so its patterning was coded as Syntagmatic-input-only (S-in).

Chapter 1 contains no illustration accompanying its paradigms. Chapter 2 is illustrated throughout with drawn characters engaged in interaction. On PDF page 26, two characters, a boy and a girl, face each other with speech bubbles; the boy's turns are consistently boxed in blue and the girls in orange across the three worked topics (Figure 1). On PDF page 27, a single boy with one arm raised in an open-palm presenting gesture is surrounded by the personal-viewpoint stems (Figure 2).

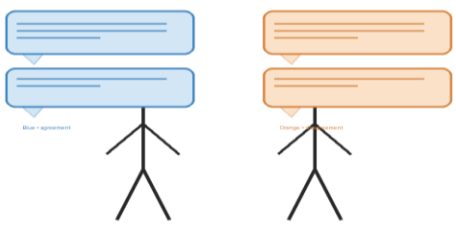

Figure 1. Schematic Reconstruction of The Colour-Coded Opinion Exchange (Boy's Turns in Blue, Girl's Turns in Orange, Across Three Worked Topics).

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

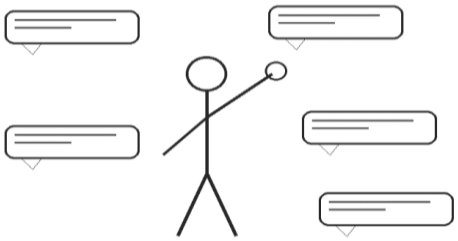

Figure 2. Schematic Reconstruction of The Composition on A Single Figure with an Open-Palm Gesture Surrounded by Personal-Opinion Expression Stems.

The dialogue between Siti and Jane on PDF page 25 represents the two speakers through contrasting postures: Siti with arms crossed and Jane with a hand on her hip and mouth open, again with each speaker's turns boxed in a consistent colour (Figure 3). The general-viewpoint stems on PDF page 28 are accompanied by three characters in varied postures, and the agreeing and disagreeing expressions on PDF page 29 are accompanied by a single figure with arms crossed and by two figures at facing lecterns, one of them pointing. Across the chapter, the images represent people performing interpersonal acts, and colour is used systematically to track speaker identity.

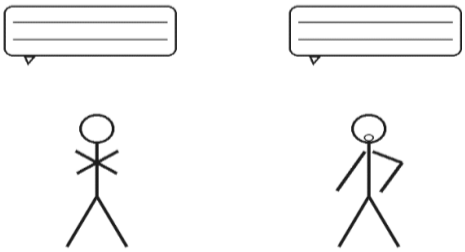

Figure 3. Schematic Reconstruction of The Contrasting Postures (Siti, Arms Crossed, Vs. Jane, Hand on Hip, Mouth Open); in the Original, Siti's Bubbles Are Boxed in Blue and Jane's in Orange Throughout.

Chapter 4 uses a different kind of image. Its pre-reading passage is accompanied by an authentic photographic pair (Picture 4.1, source: nationalgeographic.com) showing the same glacier, "Glacier AX010, Nepal," in 1978 and 2004, visibly reduced in the later photograph (Figure 4). The image represents physical evidence of environmental change rather than people communicating.

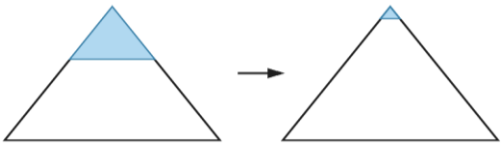

Figure 4. Schematic Reconstruction of The Before/After Comparison

Relation between the Images and the Linguistic Point (RQ3)

In Chapter 2, most images stand in a supporting relation to the paradigm they accompany. The boy–girl exchange on page 26 is of equal, complementary status with the text and elaborates it: the two characters facing each other visually enact the act of giving and countering an opinion that the exchange models. The open-palm figure on page 27 also elaborates its paradigm, since the gesture enacts the act of offering a viewpoint (Kress & van Leeuwen, 2006). This pairing is notable because the verbal paradigm here is isolated (P): the image supports a speech act that the text leaves structurally incomplete, showing that verbal integration and image–text congruence vary independently. The postures of Siti and Jane on page 25 extend the dialogue by adding an interpersonal stance that the words alone do not state, and the agreeing and disagreeing figures on page 29 elaborate the two paradigms they accompany. The only exception is the group of three figures on page 28, which is of equal but independent status: the varied postures may echo the plurality of most people and some

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

people, but the textbook does not establish this link, so no logico-semantic relation is realized and the general-viewpoint paradigm remains unsupported both verbally and visually.

Table 2. Status and Logico-Semantic Coding of Every Illustrated Unit Across the Three Chapters

Chapter/image	Status; logico-semantic relation	Relation to the verbal paradigm-structure point
Ch.1 (Building Blocks, offers/suggestions)	No image	Not applicable; integration is achieved entirely through tabular text.
Ch.2, p. 25 (Siti and Jane dialogue)	Equal (complementary); extension	Crossed arms (Siti) vs. hand-on-hip/open-mouth (Jane) add interpersonal stance not recoverable from the dialogue alone; Siti's turns are consistently boxed in blue, Jane's in orange, an unstated visual device tracking who is speaking.
Ch.2, p. 26 (boy/girl opinion exchange)	Equal (complementary); elaboration	Two characters visually enact giving and countering an opinion across three worked topics; the boy's turns are boxed in blue, the girls in orange, consistently distinguishing the two speaker roles the paradigm teaches.
Ch.2, p. 27 (single gesturing figure)	Equal (complementary); elaboration	Open-palm gesture visually enacts the communicative act (giving a personal opinion) that the adjoining paradigm names.
Ch.2, p. 28 (three figures, varied posture)	Equal (independent); no realized relation (elaboration only plausible)	Multiple postures plausibly echo the semantic plurality of "most people / some people / the majority", though the textbook does not state this correspondence.
Ch.2, p. 29 (agreeing/disagreeing figures)	Equal (complementary); elaboration	A single arms-crossed figure anchors the agreement paradigm; two figures at facing lecterns, one pointing, anchor the disagreement paradigm, again split by consistent colour-coding.
Ch.4, p. 52 (glacier before/after, Nepal, 1978 vs. 2004)	Unequal (image subordinate to text); enhancement	Warrants the reading passage's factual claim; not linked to any causal-conjunction or descriptive-synonym paradigm the chapter teaches.

Table 3 sets out the specific visual features of Chapter 2 and how each relates to the linguistic meaning of the activity it accompanies. Three kinds of feature carry this support. Gesture enacts the speech act itself: the open palm on page 27 shows a speaker offering a view, which is the function of in my opinion … and As I see it …. Posture and orientation show the interpersonal relation between speakers: the facing characters on pages 26 and 29 enact an exchange in which one speaker responds to another, and the contrasting postures of Siti and Jane on page 25 add a stance that the words do not state. Color supports the structure of the exchange: because each speaker's turns are always boxed in the same color, learners can see that an opinion is followed by an agreeing or disagreeing response, which is the sequence of moves the chapter teaches. Color thus makes visible a syntagmatic relation at the level of the exchange, even where the verbal paradigms themselves are not given a sentence frame.

Table 3. Visual Features in Chapter 2 and Their Relation to the Linguistic Point

Page and image	Visual feature	Language item and function	How the feature relates to the linguistic meaning
p. 25, Siti and Jane dialogue	Contrasting postures (arms crossed vs. hand on hip, mouth open); each speaker's bubbles in a consistent colour	Dialogue in which the two speakers exchange views	Posture adds an interpersonal stance not stated in the words (extension); colour tracks which speaker holds which view across turns
p. 26, boy and girl exchange	Characters facing each other; boy's turns in blue, girls in orange	Giving an opinion and countering it (I like … because …; I don't agree with you …)	Facing orientation enacts the interaction; colour separates the opinion move from the response move (elaboration)
p. 27, single boy	Raised arm with an open-palm presenting gesture	Personal-viewpoint stems (In my opinion …; To my mind …; As I see it …)	The gesture enacts offering a view (elaboration), although the stems are not completed verbally
p. 28, three figures	Three characters in varied postures; no link stated	General-viewpoint stems (Most people …; Some people say that …)	The plurality of figures may echo the plural subjects, but no relation is established (independent)

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Page and image	Visual feature	Language item and function	How the feature relates to the linguistic meaning
p. 29, agreement	A single figure in a consistent outline colour	Agreeing (I couldn't agree more)	Anchors the agreement expressions to one speaker (elaboration)
p. 29, disagreement	Two figures at facing lecterns, one pointing; two bubble colours	Disagreeing (I am not convinced that …)	Facing positions and pointing enact opposition; colours keep the opposing speakers distinct (elaboration)

In Chapter 4, the glacier photograph is of unequal status, subordinate to the reading passage, and relates to it by enhancement. Its dated, source-attributed realism functions like a citation that supports the passage's claim about global warming, but it does not elaborate or extend any paradigm the chapter teaches. The photograph is not captioned with any of the causal conjunctions or descriptive synonyms listed in the Language Features box, although such a caption (e.g., The glacier has shrunk dramatically as a consequence of rising temperatures) would have linked the image directly to the language point. Chapter 2's images are simpler in composition but congruent with the language being taught, whereas this photograph is representationally richer but disconnected from it. Table 2 summarizes the status and logico-semantic coding of every illustrated unit.

Table 4 brings the two dimensions together. Across the three chapters, one pattern recurs: the strength of support depends on the type of paradigm being taught. Chapter 1's offers and suggestions form small, closed sets of fixed expressions, each fitting a predictable sentence pattern, so a table can display the whole system and an image would add little. Chapter 2's opinions are interpersonal acts expressed through tone, posture and reaction, which is what drawn characters can show (Kress & van Leeuwen, 2006); here the images compensate for paradigms whose structural realization is weak. Chapter 4's language features are large, open sets used across extended written text, which no single table or image could capture completely. This explains why the chapter relies on lists and a documentary photograph, but it does not explain the absence of the simple support that the book provides elsewhere, such as an example sentence beside each list or a caption that uses the chapter's own causal conjunctions.

Table 4. Cross-Chapter Comparison of Verbal and Visual Support

Dimension	Chapter 1 (Offers & Suggestions)	Chapter 2 (Opinions & Thoughts)	Chapter 4 (National Disaster – An Exposition)
Type of paradigm	Small, closed sets of functional expressions	Functional expressions for giving and responding to opinions	Large, open lexical sets for written exposition
Verbal integration (RQ1)	Strong: integrated, contiguous (I-c) for all paradigms; reinforced by a practice task	Mixed: one I-c, three I-f, two P	Weak: mostly P at presentation; four of seven realized noncontiguously (I-n); reading input S-in
What the images represent (RQ2)	No illustration	Drawn characters interacting through gesture, posture and colour	Authentic before-and-after photograph of a retreating glacier
Image–text relation (RQ3)	Not applicable	Mostly equal (complementary) and elaborating or extending; one equal (independent) with no realized relation	Unequal (image subordinate) and enhancing only
Fit with the design-and-organization criterion	Met through verbal presentation	Met mainly through image–text congruence	Not met: weak verbal and visual support

The coexistence of mixed verbal integration and consistent visual congruence in Chapter 2 suggests that the two kinds of support are organized on different principles. The illustrations follow the communicative function of each section: every speech act the chapter introduces (giving an opinion, agreeing, disagreeing) receives a character who performs it. The structural

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

frames, in contrast, follow the grammar of the paradigm: a frame is given where the paradigm consists of verbs that project a clause (Table 2.1), but not where it consists of sentence stems or complete formulaic expressions. The images therefore show learners what each expression does, while the verbal design shows how to build a sentence with only some of them. This pattern indicates that the textbook's visual design and its language-content design were aligned at the level of function rather than at the level of structure.

Compared with Chapters 1 and 2, the weaker integration in Chapter 4 involves three distinct differences. The first is distance: in Chapters 1 and 2 the integrated paradigms and their frames appear together, whereas in Chapter 4 four paradigms are realized only two pages later in Table 4.1, and three are never realized. The second is the absence of direct visual support: no image accompanies the Language Features box, while in Chapter 2 an image accompanies almost every paradigm. The third is the limited pedagogical relevance of the only image: the glacier photograph supports the topic of the reading but is not connected to any language item, whereas Chapter 2's images enact the very speech acts being taught. All three differences point in the same direction, which is why Chapter 4 is the only chapter that does not meet the design-and-organization criterion on either dimension.

The comparison therefore specifies the unevenness of the textbook. The book applies verbal integration consistently to closed functional paradigms, applies visual support consistently to interpersonal paradigms, but applies neither to the open lexical paradigms of the written genre, which are arguably the hardest for learners to use without support. Following Halliday (2013) view of meaning as choice realized in structure, a learner who meets consequently only in a list has been shown the choice without being shown how it becomes meaning. Studies of multimodal literacy in Indonesian EFL contexts likewise emphasize that learners need guidance in reading how images and words work together (Damayanti & Febrianti, 2020; Jayanti & Damayanti, 2023). Because the textbook already demonstrates both kinds of support in its earlier chapters, closing the gap in Chapter 4 would require extending an existing model rather than introducing a new design approach.

CONCLUSIONS

This study found that the alignment between paradigmatic choices, syntagmatic structures, and illustrations varied across the three chapters. Chapter 1 consistently integrated paradigms with sentence structures, while Chapter 2 showed partial integration and Chapter 4 presented several language features without sufficient sentence-level realization. Illustrations in Chapter 2 generally supported interpersonal language functions, whereas the photograph in Chapter 4 reinforced the reading content rather than the language focus. These findings indicate specific weaknesses rather than general textbook deficiencies, suggesting targeted revisions, including adding example sentences, completing unfinished lists, and explicitly connecting illustrations to relevant language features. Theoretically, the study demonstrates how paradigmatic–syntagmatic analysis and image–text relations can operationalize Cunningsworth (1995) textbook evaluation criteria through observable categories. However, the study was limited to three purposively selected chapters, document analysis, and a single analyst. Therefore, its findings cannot be generalized to other chapters, textbooks, or classroom effectiveness. Future research should examine additional textbooks and investigate how teachers and learners use linguistic and visual resources in classroom practice.

DECLARATION

Author Contributions: Risa Anggraini conducted the textbook examination, applied the analytical categories, and prepared the manuscript. Anni Holila Pulungan provided scholarly direction, assessed the analytical framework, and offered critical revisions to the final discussion.

Conflict of Interest: The authors have no interests or affiliations that could compromise the objectivity of this publication.

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Funding: The research was completed using the authors' own resources, without external sponsorship.

Ethical Approval: No ethical review was required because the research material consisted of an existing English textbook.

Informed Consent: Not applicable, since the study did not recruit participants or collect personal information.

Use of Generative AI: The authors retained full responsibility for the textbook evaluation and multimodal interpretation. AI assistance, if any, was used only for editorial support.

REFERENCES

Barthes, R. (1977). *Image, music, text*. Fontana Press.

Bashir, A., Aziz, A., Imran, M., & Almusharraf, N. (2025). A multimodal discourse analysis of visual representation in English language textbooks at the elementary level: A social semiotic perspective. *Khazar Journal of Humanities and Social Sciences*, *28*(3), 1. https://kjhss.khazar.org/journal/vol28/iss3/1

Bashir, M. (2017). *Bahasa Inggris SMA/MA/SMK/MAK kelas X*I (Edisi revisi). Kementerian Pendidikan dan Kebudayaan. https://static.buku.kemdikbud.go.id/content/pdf/bukuteks/k13/bukusiswa/B%20Inggris%20Kelas%20XI%20BS%20press.pdf

Cunningsworth, A. (1995). *Choosing your coursebook*. Macmillan Heinemann.

Damayanti, I. L., & Febrianti, Y. (2020). Multimodal literacy: Unfolding reading path in children's picture book and its potential for EFL classrooms. *Indonesian Journal of Applied Linguistics*, *9*(3), 616–627. https://doi.org/10.17509/ijal.v9i3.23212

Eid, I. A., & Al-Jamal, D. A. (2023). Exploring collocations in EFL textbooks and providing teaching guidelines. *Theory and Practice in Language Studies*, *13*(11), 2941–2951. https://doi.org/10.17507/tpls.1311.25

Elmiana, D. S. (2019). Pedagogical representation of visual images in EFL textbooks: A multimodal perspective. *Pedagogy, Culture & Society*, *27*(4), 613–628. https://doi.org/10.1080/14681366.2019.1569550

Hafner, C. A. (2020). Digital multimodal composing: How to address multimodal communication forms in ELT. *English Teaching*, *75*(3), 133–146. https://doi.org/10.15858/engtea.75.3.202009.133

Halliday, M. A. K. (2013). Meaning as choice. In L. Fontaine, T. Bartlett, & G. O'Grady (Eds.), Systemic functional ling*uistics: Exploring choice (pp. 15–36). Cambridge University* Press https://doi.org/10.1017/cbo9781139583077.003

Jamilah, S., Ismail, N. M., & Faizah, C. (2024). Multimodal analysis of an English textbook used for EFL young learners. *New Language Dimensions*, *5*(1), 50–63. https://doi.org/10.26740/nld.v5n1.p50-63

Jayanti, E. D., & Damayanti, I. L. (2023). Exploring teachers' perceptions of integrating multimodal literacy into English classrooms in Indonesian primary education. *Child Education Journal*, *5*(2), 98–109. https://doi.org/10.33086/cej.v5i2.5240

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Kyngäs, H. (2020). Qualitative research and content analysis. In H. Kyngäs, K. Mikkonen, & M. Kääriäinen (Eds.), *The application of content analysis in nursing science research* (pp. 3–11). Springer. https://doi.org/10.1007/978-3-030-30199-6_1

Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. Visual Communication, 4(3), 337–*371 https://doi.org/10.1177/1470357205055928*

Nordlund, M., & Norberg, C. (2025). Formulaic language in EFL teaching materials. *Education Inquiry*. https://doi.org/10.1080/20004508.2025.2593701

Oktavianti, I. N., & Budiwati, T. R. (2026). Investigating lexical bundles in Indonesian EFL textbooks for senior high schools: A corpus-based study. *XLinguae*, *19*(1), 128–147. https://doi.org/10.18355/XL.2026.19.01.09

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Anggraini & Pulungan.

Purwaningtyas, T. (2020). Didactic symbol of visual images in EFL textbook: Multimodal critical discourse analysis. *Pedagogy: Journal of English Language Teaching*, *8*(1), 51–63. https://doi.org/10.32332/pedagogy.v8i1.1959

Seddighi, H., Yousefzadeh, S., & López López, M. (2021). Qualitative content analysis as a research method to investigate hazards information in school textbooks. *MethodsX*, *8*, 101559. https://doi.org/10.1016/j.mex.2021.101559

Trisanti, N., Suherdi, D., & Sukyadi, D. (2022). Multimodality reflected in EFL teaching materials: Indonesian EFL in-service teacher's multimodality literacy perception. *Language Circle: Journal of Language and Literature*, *17*(1), 129–139.

OPEN ACCESS

© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.