Voices of English Language Education Society
| Volume 10, Number 1| | April 2026 | | E-ISSN 2579-7484 |
DOI: https://doi.org/10.29408/veles.v10i1.31986

How Images and Words Tell Stories Together: A Functional Multimodal Analysis of Let's Read Digital Picturebooks

***[1]Alan Jaelani, [2]Zuriyati, [2]Miftahulkhairah Anwar**

[1]Universitas Ibn Khaldun Bogor, Indonesia
[2]Universitas Negeri Jakarta, Indonesia

Abstract: Although digital picturebooks have attracted increasing scholarly attention, relatively few studies have systematically applied functional image–text relation models to examine how meaning is multimodally organized, particularly in Southeast Asian open-access digital publishing contexts. This study investigates how visual and verbal modes are structurally coordinated in digital children's literature by analyzing image–text relations and logico-semantic patterns in three adventure-themed storybooks published on the Let's Read platform: *A Cry by the Lake*, *Run Away to the Stars*, and *Which Route is Faster?* Drawing on Martinec and Salway's (2005) framework, which extends Halliday's logico-semantic principles to visual–verbal interaction, the study employs a qualitative multimodal discourse analysis. Each page containing a co-present illustration and written text was treated as a multimodal unit of analysis, resulting in a corpus of 50 pages. The findings indicate a recurrent preference for equal–complementary image–text relations, in which visual and verbal modes work interdependently to construct narrative meaning without clear modal dominance. Across the corpus, expansion relations, particularly elaboration and causal enhancement, emerge as prominent strategies for organizing narrative development, while projection–idea relations are concentrated in scenes involving imagination, reflection, and emotional realization. These patterns suggest that the selected digital picturebooks are semiotically designed to support narrative coherence and to represent experiential and interpersonal meanings through patterned visual–verbal coordination. The study highlights the analytical value of functional multimodal frameworks for examining children's digital literature and offers theoretically grounded insights into image–text design practices in Southeast Asian open-access publishing.

Keywords: Digital children's literature, image–text relations, logico-semantic relations, multimodal discourse analysis, visual-verbal interaction.

1. Introduction

In the digital era, children's literature has expanded beyond predominantly print-based storytelling into forms increasingly shaped by screen-based reading practices, digitally mediated circulation, and new modes of reader engagement. As Kucirkova (2019) argues, children's reading now takes place within a broader ecology of digital books and applications, a shift also reflected in studies of digital picturebooks, digital children's literature, and young children's classroom engagement with digital technologies (Serafini et al., 2016; Undheim, 2022; Kachorsky, 2022). Digital picturebooks, in particular, should be understood as multimodal texts in which meaning is produced not through verbal narration alone but through the

***Correspondence:** alan.jaelani@uika-bogor.ac.id

Submitted: August 6, 2025 Revised: March 4, 2026 Accepted, April 2, 2026

coordination of written language, images, design, layout, and other semiotic resources, thereby requiring expanded forms of reading and interpretation (Serafini et al., 2020). As Callow (2020) shows, visual and verbal elements in picture books may intersect in complementary, equivalent, or divergent ways. This suggests that images do far more than decorate the written story; they actively contribute to narrative development, shape interpretive pathways, and position readers' responses. Such multimodal organization is especially important in children's literature because young readers often draw on both visual and verbal cues to infer relationships, follow plot development, and generate meanings that are not explicitly stated in words alone. Similarly, Papen and Peach (2021) demonstrate that children's engagement with picturebooks emerges through interaction with both image and text, while related work on shared reading, multimodal interaction, and young learners' visual meaning-making further supports the view that picturebooks function as complex semiotic artifacts rather than simple illustrated stories (Johnson, 2017; Pantaleo, 2020).

Despite this growing interest in multimodal children's literature, existing scholarship has not always examined image–text relations as a central analytical problem. Many studies on illustrated books, multimodal texts, and children's digital reading have primarily foregrounded pedagogical outcomes, reader engagement, comprehension processes, emotional response, or aesthetic appreciation of picturebook art rather than the formal organization of intersemiotic relations (Egert et al., 2022; Pantaleo, 2018). In many cases, images are discussed in terms of what they depict, how they invite participation and attention, or how they support children's meaning-making and literacy development, while the relationship between image and text is acknowledged but not systematically analyzed (Müller-Brauers et al., 2020; Twiner et al., 2022; Kelly & Kachorsky, 2022). As Serafini and Reid (2024) observe, contemporary picturebooks require multiple analytical lenses; however, comparatively less attention has been directed to the structural organization of image–text relations themselves, especially to the functional and semantic patterning through which visual and verbal modes jointly construct narrative meaning (Kachorsky et al., 2017; Sargeant, 2015). This leaves an important gap in the literature because, without a systematic account of how images and words interact across a narrative, analyses of digital picturebooks may remain largely descriptive, focusing on prominent features, themes, or affordances rather than explaining how multimodal meaning is organized (Serafini et al., 2016; Papen, 2020). Therefore, closer attention to visual–verbal intersections is needed through a more explicit functional approach that examines not only what multimodal elements mean in isolation, but also how they work together to produce coherent and interpretable meaning across an unfolding narrative.

To address this limitation, a more explicit theoretical framework is needed—one that can explain not only the presence of visual and verbal elements in children's picturebooks, but also the specific ways these modes are structurally related in meaning-making (Unsworth & Macken-Horarik, 2015; Moya-Guijarro & Ventola, 2022). Pagano et al. (2018) provide a useful precedent for this type of functional approach, while Martinec and Salway's (2005) model offers a systematic framework by extending Halliday's logico-semantic principles beyond language to the analysis of image–text relations (Haiyan, 2018; Huang, 2022). Rather than treating words and images as loosely connected or intuitively complementary, this model classifies visual–verbal interaction through two key dimensions: status relations and logico-semantic relations. These dimensions allow equal or subordinate relations between modes, as well as relations of elaboration, extension, enhancement, and projection, to be examined more precisely across picturebook discourse (Moses et al., 2020; Kelly & Kachorsky, 2022). As Gu and Catalano (2022) show, multimodal analysis becomes more explanatory when it focuses on how image and text work together to construct meaning rather than merely cataloguing what each mode represents separately. For this reason, Martinec and Salway's framework is particularly valuable for the present study because it enables the analysis to move beyond surface description toward a more precise explanation of how multimodal meaning is organized across narrative units. In this sense, the study is not primarily concerned with what images or words represent separately, but with how their patterned interaction contributes to the construction of coherent narrative meaning in digital picturebooks (Aguilera et al., 2016).

This functional orientation also helps clarify why Martinec and Salway's framework is more suitable for the present study than earlier approaches to image–text relations. Earlier models have provided important foundations for multimodal analysis. Yet, more recent scholarship suggests that contemporary picture books require finer-grained ways of accounting for how words, images, and design features interact across a narrative sequence (Moses et al., 2020). As McGlashan (2021) notes, simple taxonomies derived from Barthes' concepts of anchorage, illustration, and relay remain influential because they show how verbal language can specify, support, or work alongside images. However, later studies also indicate that these categories are comparatively broad and may leave finer dependencies, contradictions, or degrees of interdependence underdescribed, especially in more complex multimodal texts (Haiyan, 2018; Otto et al., 2020; Grange & Lian, 2022). Likewise, Kress and van Leeuwen's visual grammar has been indispensable for explaining how images realize representational and interactive meanings through resources such as contact, social distance, angle, and power. Nevertheless, recent picturebook research suggests that this model is primarily oriented toward the semiotic organization of visual discourse and is not always sufficient on its own for tracing visual–verbal relations at a more delicate semantic level (Santamaría-García, 2022; Moya-Guijarro & Martínez Mateo, 2022; Huang, 2022). In contrast, Martinec and Salway's framework offers a more explicit functional account of image–text interaction because it combines status relations with logico-semantic relations, allowing analysts to describe equal versus unequal dependencies as well as elaboration, extension, enhancement, and projection more systematically across multimodal units (Hu & Qiu, 2020; Pagano et al., 2018; Haiyan, 2018). For this reason, it provides a suitable analytical basis for the present study, which examines how images and written language are structurally coordinated to produce narrative meaning in digital picturebooks whose design and reading experience are shaped by digital affordances and multimodal narrative features (Sargeant, 2015; Bus & Anstadt, 2021).

Despite the growing scholarly attention to digital picturebooks and multimodal children's literature, existing studies have largely focused on reader engagement, literacy development, pedagogical use, and the aesthetic or thematic roles of illustrations, while giving comparatively limited attention to the functional organization of image–text relations within the narrative structure itself. In particular, few studies have systematically examined how visual and verbal modes are structurally related through status relations and logico-semantic patterns in digital picturebooks, leaving a gap in understanding how multimodal meaning is actually constructed across narrative units. Addressing this gap, the present study aims to analyze the image–text relations and logico-semantic patterns in selected digital children's storybooks published on the Let's Read platform by applying Martinec and Salway's (2005) functional framework, focusing on how visual and verbal modes are coordinated to produce meaning rather than on reader outcomes or pedagogical impact. The novelty of this study lies in its systematic application of a functional multimodal discourse analysis to digital picturebooks, treating each page as a unit of analysis to reveal patterned relations between modes and to provide a more precise explanation of how narrative meaning is organized through visual–verbal interaction. Accordingly, this study addresses the following research questions: (1) What types of image–text relations are present in the selected digital children's storybooks? (2) What logico-semantic relations are used in the selected digital children's storybooks? and (3) How do these patterns contribute to meaning-making in the selected digital children's storybooks?.

2. Method

2.1 Research Design

This study employed a qualitative multimodal discourse analysis to examine how image–text relations and logico-semantic patterns are organized in selected digital picturebooks. This design was appropriate because the study focused on how narrative meaning is constructed through the interaction of visual and verbal modes, rather than through reader responses, classroom practices, or learning outcomes. Multimodal discourse analysis enables researchers to investigate how different semiotic resources, including written language, images, layout, visual composition, and design, work together to produce meaning in a text (Jones,

2012; O'Halloran, 2008). In digital picture books, such an approach is particularly relevant because meaning is not carried by written language alone but is distributed across images, verbal narration, and page-based visual design.

The study was grounded in Systemic Functional Linguistics and functional multimodal theory, particularly the view that meaning emerges through patterned semiotic choices across modes (O'Halloran, 2008). To analyze visual–verbal interaction, the study applied Martinec and Salway's (2005) image–text relation framework, which classifies image–text relations into two major dimensions: status relations and logico-semantic relations. Status relations explain whether image and text contribute equally or unequally to meaning, while logico-semantic relations explain how the two modes expand or project meaning. This framework has been widely used in multimodal research because it provides a systematic way to examine how visual and verbal modes are functionally coordinated rather than treated as separate semiotic elements (Martinec & Salway, 2005; Haiyan, 2018; Huang, 2022; Pagano et al., 2018). The original manuscript also identifies the study as a qualitative multimodal discourse analysis using Martinec and Salway's framework to analyze three Let's Read digital picturebooks.

2.2 Data Source and Corpus

The data were taken from three adventure-themed digital storybooks published on the Let's Read platform: A Cry by the Lake, Run Away to the Stars, and Which Route is Faster? These storybooks were selected because they contain page-based digital narratives in which written text and illustrations appear together and jointly contribute to story development. Digital picturebooks are well-suited to multimodal analysis because they function as integrated semiotic artifacts in which meaning emerges through the orchestration of verbal, visual, and design resources (Serafini et al., 2016; Serafini et al., 2020; Kachorsky, 2022). The selected texts were also considered appropriate because they present clear narrative sequences, character actions, spatial movement, emotional moments, and problem-solving events, allowing for the systematic examination of image–text relations and logico-semantic patterns.

The English-language versions of the storybooks were used to maintain consistency across the corpus. The corpus consisted of 50 narrative pages, comprising 21 pages from A Cry by the Lake, 14 pages from Run Away to the Stars, and 15 pages from Which Route is Faster? Cover pages and publication credit pages were excluded because they did not form part of the main narrative sequence. However, narrative closure pages containing factual or contextual information were included when they contributed to the story's overall meaning-making. Platform navigation features were not analyzed because the study focused specifically on image–text relations within the narrative content. This delimitation is consistent with multimodal discourse analysis, which requires the researcher to define a bounded corpus and determine which semiotic resources are relevant to the research focus (Jones, 2012; O'Halloran, 2008). The original corpus and inclusion criteria are consistent with the manuscript's description of 50 analyzed pages across the three storybooks.

2.3 Unit of Analysis

The unit of analysis was the individual narrative page containing both illustration and written text. Each page was treated as a single multimodal unit because the image and written text appeared together, forming a single meaning-making ensemble. This page-based unit was appropriate because picturebook meaning is often constructed through the simultaneous interaction of visual and verbal elements within a shared page or screen space (Serafini et al., 2018; Kachorsky et al., 2017; Moses et al., 2020). In digital picturebooks, readers typically encounter illustration and verbal narration as co-present semiotic resources; therefore, analyzing them together allows the researcher to examine how the two modes jointly construct narrative meaning. Each multimodal unit was examined through two analytical dimensions. First, the analysis identified the status relation between image and text, including whether the relation was equal or unequal. Equal relations were further interpreted as equal–complementary when image and text depended on each other to construct meaning, or equal–independent when both modes contributed meaning without strong dependency. Unequal relations were identified when one mode was dominant, and the other was

subordinate, such as an image subordinate to text or text subordinate to an image. Second, the analysis identified the logico-semantic relation between image and text, including expansion relations—elaboration, extension, and enhancement—and projection relations—locution and idea. These categories follow Martinec and Salway's functional model of image–text relations (Martinec & Salway, 2005).

2.4 Analytical Framework

The analysis adopted Martinec and Salway's (2005) framework of image–text relations as the main analytical framework. This framework was selected because it provides a functional classification of how visual and verbal modes interact. In the first dimension, status relations refer to the relative contribution of image and text. An equal relation occurs when image and text contribute relatively balanced meanings, while an unequal relation occurs when one mode is more dominant than the other in conveying meaning. This distinction is useful in picturebook analysis because images and words may sometimes work interdependently, while at other times one mode may carry stronger narrative, emotional, spatial, or explanatory meaning (Callow, 2020; Kelly & Kachorsky, 2022; Moya-Guijarro & Martínez Mateo, 2022).

In the second dimension, logico-semantic relations explain the semantic connection between image and text. Expansion occurs when one mode develops the meaning of the other. It includes elaboration, where one mode restates, specifies, or clarifies the meaning of the other; extension, where one mode adds new but related information; and enhancement, where one mode adds circumstantial information such as time, place, cause, condition, or manner. Projection occurs when one mode represents speech, thought, perception, imagination, or inner experience. It includes locution, which relates to represented speech or verbal expression, and idea, which relates to thought, imagination, reflection, or mental representation (Martinec & Salway, 2005). This framework is particularly relevant to picture books because visual and verbal modes often work together to represent not only events and actions but also emotions, spatial relations, causal links, and characters' internal experiences (Papen & Peach, 2021; Serafini & Reid, 2024).

The framework was used not merely to count relation types but to interpret their narrative functions. Equal–complementary relations were examined in terms of how image and text jointly constructed plot continuity, character action, and setting. Enhancement relations were examined for their support of temporal, spatial, and causal narrative development. Projection–idea relations were examined in terms of their representation of imagination, reflection, emotional realization, and problem-solving. This functional orientation allowed the analysis to move beyond descriptive labeling toward an explanation of how multimodal coordination contributes to meaning-making in digital children's literature.

2.5 Data Analysis Procedures

Data analysis was conducted through several stages. First, the three selected storybooks were read repeatedly to gain an overall understanding of their narrative structure, character development, and visual–verbal organization. Repeated close reading is important in qualitative multimodal analysis because it allows the researcher to examine how meaning emerges across semiotic resources and narrative sequences rather than from isolated elements (O'Halloran, 2008; Jones, 2012). Second, each page containing both images and written text was coded as a single multimodal unit. The researcher examined the illustration, written text, spatial arrangement, character action, setting, emotional expression, and narrative sequence on each page. The purpose of this stage was to determine how the visual and verbal modes contributed to the meaning of the page. Third, the status relation of each page was coded. The researcher identified whether image and text had an equal or unequal relationship. When both modes were interdependent and contributed complementary meanings, the relation was coded as equal–complementary. When both modes contributed meaning but remained relatively independent, the relation was coded as equal–independent. When one mode carried more central meaning than the other, the relation was coded as unequal, either image-dominant/text-subordinate or text-dominant/image-subordinate.

Fourth, the logico-semantic relation of each page was coded. The researcher identified whether the image–text relation involved expansion or projection. Expansion was coded as elaboration, extension, or enhancement, while enhancement was further interpreted as causal, spatial, or temporal when relevant. Projection was coded as locution or idea, depending on whether the multimodal relation represented speech, thought, imagination, reflection, or emotional experience. This coding procedure followed Martinec and Salway's (2005) categories and was informed by subsequent functional multimodal studies that emphasize systematic classification and interpretation of image–text relations (Haiyan, 2018; Hu & Qiu, 2020; Huang, 2022; Pagano et al., 2018). Fifth, the coded data were compared across the three storybooks to identify recurring patterns and story-specific variations. Frequencies of relation types were noted to show dominant tendencies, but they were interpreted descriptively rather than statistically. This decision was appropriate because the study aimed to explain functional meaning-making within a bounded corpus, not to generalize statistically across all digital picturebooks. Finally, the patterns were interpreted in relation to narrative meaning-making, including how image–text relations supported coherence, emotional depth, spatial orientation, causal logic, imagination, reflection, and problem-solving. The original manuscript confirms that the corpus analysis involved 50 pages and examined each page for both status relations and logico-semantic relations.

2.6 Trustworthiness of Analysis

Several strategies were used to enhance the trustworthiness of the analysis. First, the study used a clearly defined analytical framework, namely Martinec and Salway's image–text relation model, so that coding decisions were theoretically guided and systematically applied. Second, repeated close reading and visual inspection were conducted to reduce the risk of superficial or single-pass interpretation. Third, analytic memos were used during coding to document the reasons for classifying particular pages as equal, unequal, elaborative, enhancing, or projective. Memoing is useful in qualitative analysis because it helps make analytic reasoning more explicit and traceable (Birks et al., 2008). Fourth, coding patterns were compared across the three storybooks to examine whether similar image–text relations served similar narrative functions across the texts. This cross-text comparison supported interpretive consistency. Fifth, frequency counts were used only as descriptive support, while the main interpretation focused on functional meaning-making. This prevented the analysis from reducing multimodal interpretation to numerical dominance alone. Sixth, representative examples from each storybook were presented to illustrate key patterns, including equal–complementary relations, image-dominant relations, causal enhancement, elaboration, and projection–idea. In qualitative multimodal research, the use of representative examples helps readers evaluate the plausibility and transparency of the interpretation (O'Halloran, 2008; Serafini & Reid, 2024).

2.7 Scope and Delimitation

This study was delimited to the visual and verbal relations found in three selected English-language digital picture books from the Let's Read platform. The analysis focused only on co-present written text and illustration within narrative pages. It did not examine reader responses, children's comprehension, classroom use, animation, audio narration, interactive features, or platform navigation. Therefore, the findings should be understood as an account of semiotic organization and multimodal meaning-making within the selected storybooks rather than as evidence of their pedagogical effectiveness or impact on children's literacy development. This delimitation is important because digital picturebooks can be studied from multiple perspectives, including reader engagement, literacy learning, multimodal design, and platform affordances (Kucirkova, 2019; Serafini et al., 2016; Bus & Anstadt, 2021), whereas the present study specifically focuses on functional image–text relations.

3. Findings

3.1 Image–Text Relations in the Selected Digital Picturebooks

The analysis shows that the three digital picturebooks were predominantly organized through *equal–complementary image–text relations.* In this pattern, images and written text worked interdependently to construct narrative meaning. The written text generally provided explicit narrative direction, while the illustrations supplied visual information about setting, character movement, emotional expression, spatial orientation, and contextual detail. This means that neither mode functioned merely as an accessory to the other; rather, meaning was produced through their coordinated interaction.

Table 1. Summary of Image–Text Relations across the Three Storybooks

Storybook	Corpus size	Dominant relation	Notable unequal relation	Functional tendency
A Cry by the Lake	21 pages	Equal–complementary	Text-dominant pages for factual/descriptive information; image-dominant pages for emotional scenes	Image and text jointly construct exploration, friendship, and emotional reflection
Run Away to the Stars	14 pages	Equal–complementary	Image-dominant pages in imaginative scenes	Visual mode foregrounds fantasy, dreams, and Juno's inner world
Which Route is Faster?	15 pages	Equal–complementary	Image-dominant page for route comparison/map interpretation	Image and text jointly support action sequence, spatial reasoning, and problem-solving

In *A Cry by the Lake*, most pages displayed equal–complementary relations. For example, in the picnic scene, the text presents the family's activity, while the image shows Mike's facial expression and body posture, allowing readers to infer his mood. This relation shows how visual and verbal modes together construct both the event and the scene's emotional tone. Unequal relations appeared only in specific narrative functions. Text-dominant relations occurred when the story provided explanatory or factual information, such as the tea-making process or context about *Patenggang Lake* and *the Javan gibbon*. Image-dominant relations appeared in emotionally significant moments, particularly when Mike first encountered Spike and when Spike walked away sadly. In these moments, the illustration carried a stronger affective meaning than the verbal text.

Figure 1. Examples of image–text status relations in A Cry by the Lake

In *Run Away to the Stars*, equal–complementary relations also formed the dominant pattern, especially in scenes where Juno builds the rocket, experiments with materials, and moves through the imagined journey. These relations helped maintain narrative continuity because the text described Juno's actions while the images visualized her emotional state, imaginative world, and problem-solving process. However, image-dominant relations appeared in moments of intensified imagination, such as Juno visualizing her rocket and

launching into space. In these scenes, the visual mode became more prominent because fantasy and emotional intensity were more effectively represented through illustration than through verbal explanation alone. One equal–independent relation was also identified when Juno's sock collection was visually represented as part of her personality rather than as a direct plot-driving event.

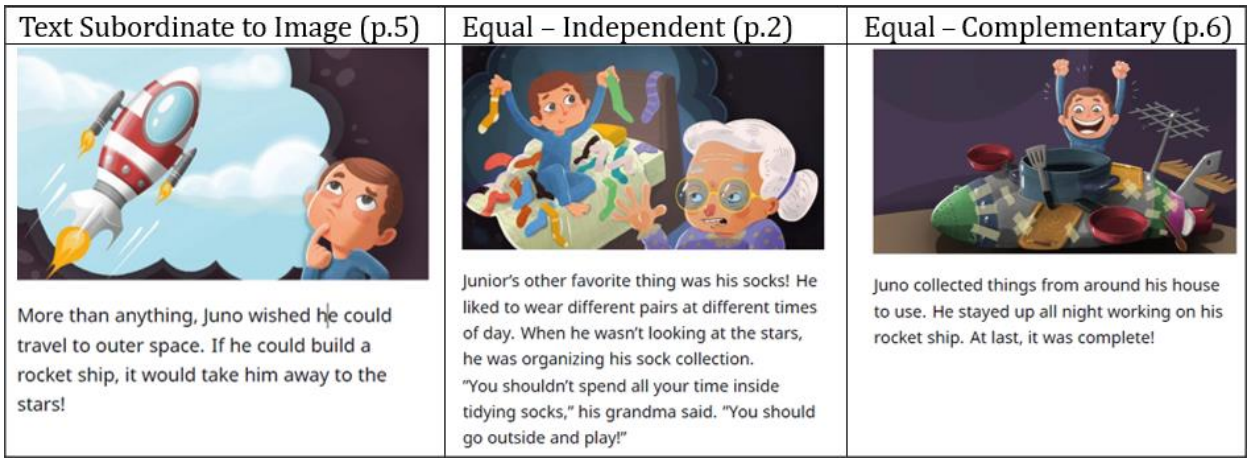

Figure 2. Examples of image–text status relations in Run Away to the Stars

In *Which Route is Faster?,* equal–complementary relations were even more dominant, appearing in nearly all narrative pages. The written text explained the children's actions, while the images clarified spatial movement, route choices, obstacles, and consequences. Only one strongly image-dominant relation appeared when the two alternative routes were visually mapped. In this case, the illustration conveyed the central meaning because comparing routes, distances, directions, and potential obstacles was easier to understand visually than in text alone. This indicates that image dominance occurred not because the verbal mode was weak, but because spatial reasoning was more effectively communicated visually. Across the three storybooks, unequal relations were not random deviations from the dominant pattern. Instead, they appeared at narratively strategic points. Text became dominant when factual explanation or contextual information was required. Image became dominant when the story needed to represent emotion, imagination, fantasy, or spatial comparison. This pattern suggests that the selected digital picturebooks are designed around modal specialization: text tends to anchor sequence and explanation, while images intensify affect, imagination, spatial orientation, and character experience.

3.2 Logico-Semantic Relations across the Storybooks

The second research question examined the logico-semantic relations used in the selected digital picture books. The findings show that *expansion relations* were the dominant logico-semantic pattern across the corpus. Expansion appeared through elaboration, extension, and enhancement. Projection relations were less frequent, but they played important narrative roles in representing imagination, reflection, speech, and emotional realization.

Table 2. Summary of Logico-Semantic Relations across the Three Storybooks

Logico-semantic relation	Function in the corpus	Storybook tendency	Narrative contribution
Expansion–elaboration	Clarifies or specifies characters, settings, events, and factual information	Prominent in all three storybooks	Supports narrative clarity and contextual understanding
Expansion–extension	Adds related information or shifts narrative focus	Appears in character-building and movement sequences	Expands plot, setting, and character details

Logico-semantic relation	Function in the corpus	Storybook tendency	Narrative contribution
Enhancement–causal	Shows cause-and-effect relations among actions, obstacles, and outcomes	Strongly visible in Which Route is Faster? and Run Away to the Stars	Builds narrative logic and consequence
Enhancement–spatial	Clarifies location, route, direction, or movement	Visible in A Cry by the Lake and Which Route is Faster?	Supports spatial orientation
Enhancement–temporal	Mark's sequence and progression of events	Appears in action sequences	Supports plot development
Projection–idea	Represents thought, imagination, reflection, emotional realization, or problem-solving	Prominent in Run Away to the Stars and selected scenes in A Cry by the Lake	Gives access to characters' inner worlds
Projection–locution	Represents speech or vocal expression	Limited occurrence	Supports reinterpretation of sound or verbal meaning

In *A Cry by the Lake*, expansion relations were used to support setting, action, and narrative development. Elaboration appeared when images and text jointly specified scenes such as the family trip, tea plantation, and picnic, as well as factual information about the Javan gibbon. Enhancement relations were also important, particularly causal and spatial enhancement. For instance, Mike's hearing a mysterious sound functions as a causal trigger for the subsequent narrative movement, while crossing the lake and entering the forest are supported through spatial enhancement. Projection relations appeared in emotionally significant scenes involving Mike's encounter with Spike, his desire for friendship, Spike's departure, and Mike's later reflection. These projection–idea relations shifted the narrative from external action to emotional and psychological meaning.

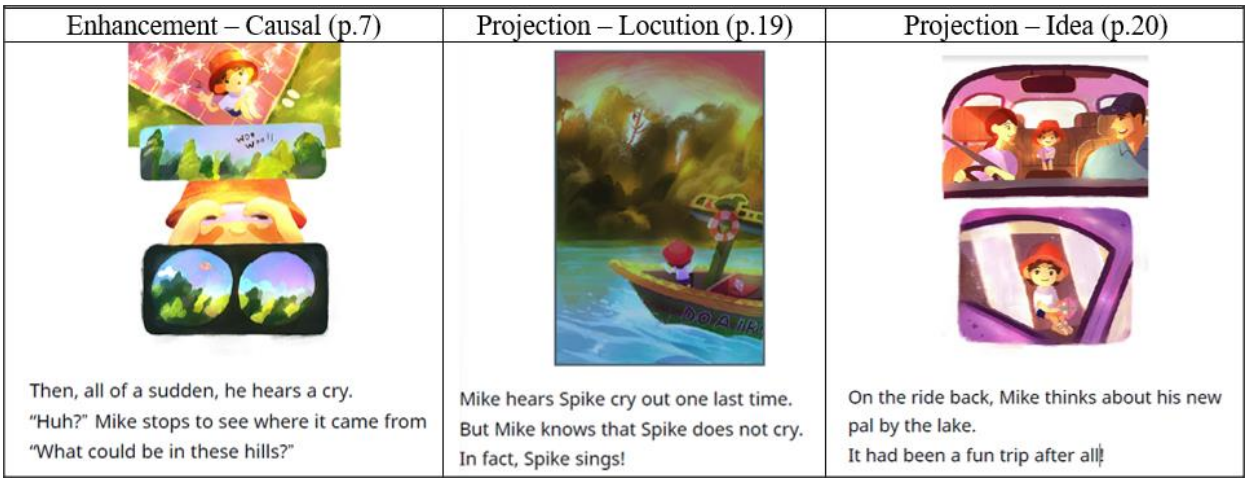

Figure 3. Examples of logico-semantic relations in A Cry by the Lake

In *Run Away to the Stars*, the projection–idea was particularly salient because the story centers on Juno's dreams, desires, and imaginative escape to outer space. Projection–an idea appeared in scenes where Juno looked at the stars, imagined building a rocket, launched into space, and later realized that the journey was a dream. These relations positioned Juno's inner world as the story's thematic center. At the same time, enhancement–causal relations structured the plot by linking Juno's obsession with stars, failed attempts to power the rocket, problem-solving actions, and eventual return to reality. Expansion relations supported character development and exploratory movement, including Juno's sock collection, rocket-building process, and movement through imagined galaxies.

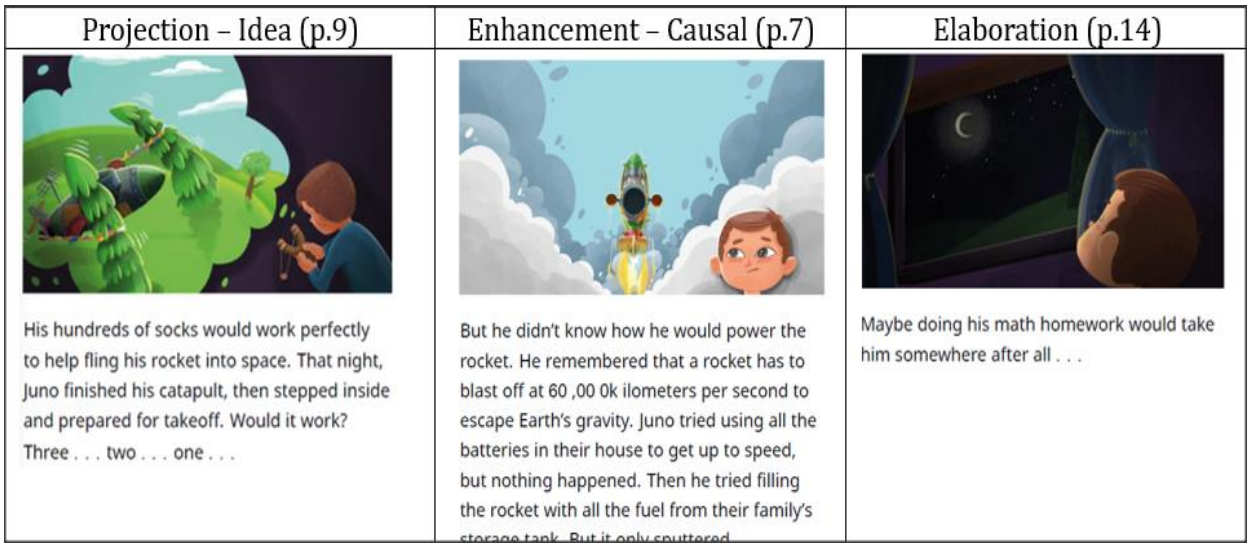

Figure 4. Examples of logico-semantic relations in Run Away to the Stars

In *Which Route is Faster?*, enhancement–causal relations were especially prominent. The story is built around choices, obstacles, delays, and outcomes; therefore, causal relations are central to its narrative structure. Pages involving ants, puddles, friends, a torn map, and the final arrival at the aunt's house all demonstrate how image and text work together to show consequence. Expansion–exemplification also appeared in the route comparison and environmental cues, such as the guava tree and the gate. These visual-verbal relations helped young readers understand that the fastest route is not determined only by distance but also by obstacles and decisions along the way. Projection–an idea appeared once, when a character generated a new idea after losing the map. Although limited, this relation marked an important moment of problem-solving and internal cognitive engagement.

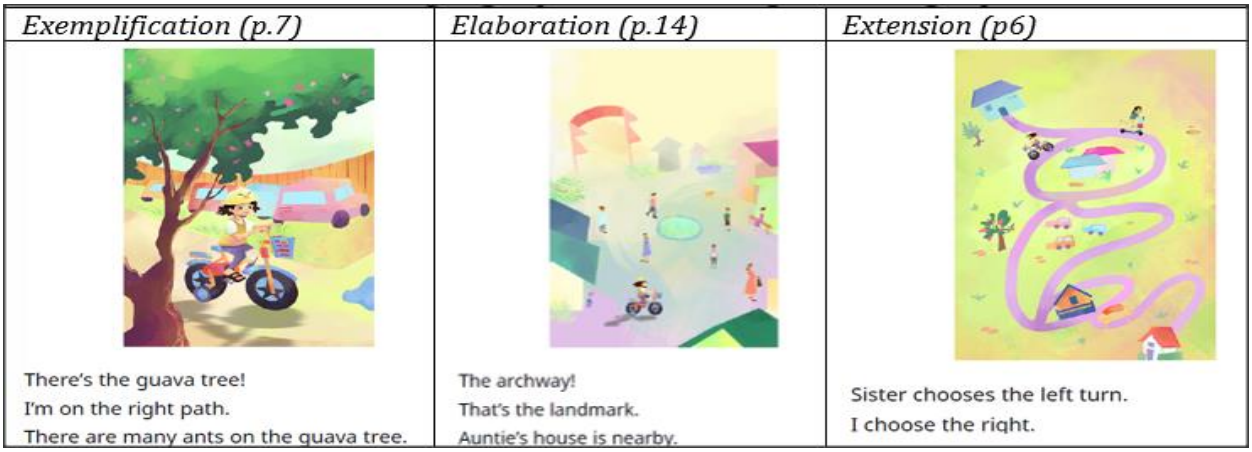

Figure 5. Examples of logico-semantic relations in Which Route is Faster?

The logico-semantic analysis shows that each storybook employs expansion and projection in line with its narrative focus. *A Cry by the Lake* combines elaboration, spatial enhancement, causal enhancement, and projection to develop exploration and emotional reflection. *Run Away to the Stars* relies strongly on projection–idea to represent dreams and imagination, while enhancement–causal structures Juno's attempts and failures. *Which Route is Faster?* depends heavily on causal enhancement and exemplification because its narrative meaning is built around route selection, obstacles, and problem-solving.

3.3 Contribution of Image–Text and Logico-Semantic Patterns to Meaning-Making

The third research question asked how image–text and logico-semantic patterns contribute to meaning-making in the selected digital picture books. The findings indicate that visual–verbal relations contribute to meaning-making in three major ways: supporting narrative coherence, representing experiential and interpersonal meaning, and guiding young readers' interpretation of action, emotion, and thought. First, equal–complementary relations supported narrative coherence. In most pages, text and image shared the work of storytelling. The text typically sequenced events, named actions, or provided verbal explanation, while the image supplied contextual details, facial expressions, body movement, spatial layout, and visual cues. This relation helped readers follow the narrative without relying exclusively on verbal information. For young readers, this is particularly important because visual details can support inference, comprehension, and continuity across pages. The dominance of equal–complementary relations across the three storybooks indicates that the stories were designed as integrated multimodal narratives rather than as written stories with decorative illustrations.

Second, expansion relations organized experiential meaning by clarifying events, extending information, and specifying circumstances. Elaboration helped explain who was involved, what was happening, and how a scene should be understood. An extension added related information that expanded characterization, setting, or narrative direction. Enhancement was particularly important because it constructed time, place, cause, and consequence. In *Which Route is Faster?,* causal enhancement was central to the story's meaning because the narrative teaches that route efficiency depends on obstacles, decisions, and consequences. In *A Cry by the Lake*, spatial enhancement helped readers understand Mike's movement from family space to the lake and forest. In *Run Away to the Stars*, causal enhancement helped structure Juno's attempts, failures, and realization.

Third, projection relations contributed to interpersonal and psychological meaning. Projection–idea allowed the stories to represent imagination, reflection, desire, emotional realization, and problem-solving. This was most visible in *Run Away to the Stars*, where Juno's imaginative journey is central to the narrative. It also appeared in *A Cry by the Lake*, where projection–idea helped represent Mike's emotional bond with Spike and his reflection on friendship. In *Which Route is Faster?,* the projection–idea marked a moment of problem-solving when the character generated a new idea after losing the map. These relations show that digital picturebooks not only narrate external events; they also represent characters' inner experiences through visual–verbal coordination.

Table 3. Functional Contribution of Multimodal Patterns to Meaning-Making

Pattern	Meaning-making function	Example from corpus
Equal–complementary relations	Distribute narrative meaning across image and text	Family picnic, rocket-building, route journey
Image-dominant relations	Foreground emotion, imagination, or spatial comparison	Spike's departure, Juno's space fantasy, route map
Text-dominant relations	Provide factual or explanatory information	Tea-making process, factual information about Patenggang Lake, and the Javan gibbon
Elaboration	Clarify setting, action, character, or contextual information	Story openings, factual closure, and activity contexts
Extension	Add related information or shift narrative focus	Sock collection, galaxy exploration, divergent routes
Enhancement	Construct time, space, cause, and consequence	Hearing mysterious sounds, a failed rocket launch, and route obstacles
Projection–idea	Represent thought, imagination, reflection, and problem-solving	Juno's dream, Mike's reflection, route solution

Pattern	Meaning-making function	Example from corpus
Projection–locution	Represent or reinterpret vocal expression	Spike's sound is reinterpreted as singing

The findings suggest that the selected Let's Read storybooks strategically employ multimodal coordination. *A Cry by the Lake* uses image–text relations to construct exploration and emotional attachment. *Run Away to the Stars* uses visual–verbal patterns to foreground imagination and internal motivation. *Which Route is Faster?* uses causal and spatial relations to support problem-solving and didactic meaning. These differences show that image–text relations are shaped by narrative purpose. The meaning of each story is not produced by images or words separately, but by the patterned interaction between them. The findings demonstrate that the selected digital picturebooks are semiotically organized through patterned visual–verbal coordination. Equal–complementary relations dominate because the stories rely on interdependence between illustration and written text. Expansion relations support narrative clarity, causal logic, spatial orientation, and event development, while projection relations represent imagination, reflection, emotion, and thought. These patterns show that digital picturebooks on the Let's Read platform can be understood as integrated multimodal narratives in which images and words work together to guide young readers through story meaning.

4. Discussion

The findings of this study demonstrate that the selected Let's Read digital picturebooks construct narrative meaning primarily through patterned visual–verbal interdependence. Across A Cry by the Lake, Run Away to the Stars. Which Route is Faster?, equal–complementary image–text relations appeared as the dominant pattern, indicating that images and written language generally worked together rather than functioning separately or hierarchically. This finding supports the view that picture books are not simply verbal stories accompanied by illustrations, but multimodal texts in which meaning is distributed across visual, verbal, and design resources (Serafini et al., 2018; Serafini et al., 2020; Callow, 2020; Kelly & Kachorsky, 2022). In the analyzed corpus, written text commonly provided explicit narrative sequencing, while illustrations supplied emotional cues, spatial orientation, character action, setting details, and contextual information. This interdependence is especially important in children's digital literature because young readers rely on both visual and verbal cues to infer relationships, follow narrative development, and understand meanings that are not fully verbalized. Thus, the dominance of equal–complementary relations suggests that the selected digital storybooks are designed as integrated multimodal narratives rather than as text-led stories with decorative images.

Although equal–complementary relations were dominant, the occurrence of unequal image–text relations indicates that modal dominance shifted with narrative function. Text-dominant relations appeared when factual, explanatory, or contextual information needed to be communicated more explicitly, such as in the tea-making explanation or factual information about Patenggang Lake and the Javan gibbon in A Cry by the Lake. By contrast, image-dominant relations appeared at moments requiring affective, imaginative, or spatial emphasis. In A Cry by the Lake, images became dominant in emotionally charged scenes involving Mike's encounter with Spike and Spike's departure. In Run Away to the Stars, image dominance occurred during Juno's imaginative rocket scenes and space fantasy. In Which Route is Faster?, the visual map became dominant because route comparison, direction, and spatial reasoning could be communicated more effectively through an image than through verbal description. This pattern confirms that unequal relations should not be interpreted as weaknesses in multimodal design. Instead, they reveal functional specialization: verbal language is foregrounded when explanation is needed, while images are foregrounded when emotion, imagination, or spatial configuration requires stronger visual representation. This finding aligns with picturebook research showing that images can guide interpretation, intensify affect, and communicate

narrative meanings that exceed verbal narration (Pantaleo, 2018; Papen, 2020; Moya-Guijarro & Martínez Mateo, 2022).

The logico-semantic analysis further shows that expansion relations functioned as the main organizing principle across the corpus. Elaboration, extension, and enhancement helped structure narrative meaning by clarifying events, expanding information, and specifying the circumstances of time, place, cause, and consequence. Elaboration was used to introduce settings, characters, activities, and factual closure; extension added related details or shifted the narrative focus; and enhancement supported temporal, spatial, and causal development. This finding supports Martinec and Salway's (2005) argument that image–text relations can be understood through functional semantic relations, not merely through visual resemblance or thematic association. It also reinforces later multimodal studies that emphasize the need to examine how modes work together to organize meaning rather than analyzing image and text in isolation (Pagano et al., 2018; Hu & Qiu, 2020; Gu & Catalano, 2022; Huang, 2022). In the present study, expansion relations were central to narrative coherence because they enabled readers to understand where events occurred, why actions took place, how characters moved, and how one event led to another. This was particularly visible in Which Route is Faster?, where causal enhancement structured the children's journey through obstacles, delays, and eventual arrival, showing that route efficiency depends not only on distance but also on decisions and consequences.

Projection relations were less frequent than expansion relations, but their narrative role was significant. Projection–an idea appeared in scenes involving imagination, reflection, emotional realization, desire, or problem-solving. In Run Away to the Stars, projection–idea was especially prominent because the narrative is built around Juno's dreams, imaginative escape, and eventual realization. In A Cry by the Lake, the projection–idea supported the representation of Mike's emotional bond with Spike and his reflection on friendship. In Which Route is Faster?, the projection idea appeared only once. However, it marked an important moment in problem-solving when the character came up with a new idea after losing the map. These findings suggest that projection relations are used selectively to shift the narrative from external action to internal experience. This is important because children's picture books often invite readers not only to follow events but also to infer characters' thoughts, emotions, motivations, and perspectives through visual–verbal coordination (Johnson, 2017; Papen & Peach, 2021; Kelly & Kachorsky, 2022). Projection–the idea therefore contributes to interpersonal and psychological meaning by allowing readers to access characters' inner worlds through the interaction of illustration and written language.

The three storybooks also reveal that multimodal patterns are shaped by narrative purpose. A Cry by the Lake combines elaboration, spatial enhancement, causal enhancement, and projection to construct a story of exploration, friendship, and emotional reflection. Run Away to the Stars relies more heavily on the idea of projection because its central theme concerns imagination, dreams, and internal motivation. Which Route is Faster? depends heavily on causal enhancement and exemplification because its narrative is organized around route selection, obstacles, and problem-solving. These differences demonstrate that image–text relations and logico-semantic patterns are not randomly distributed across the corpus. Rather, they are selected according to the thematic and narrative demands of each story. This supports the argument that multimodal meaning-making in picture books depends on patterned relationships across narrative sequences, not simply on the presence of attractive images or accessible written text (Callow, 2020; Serafini & Reid, 2024). The functional organization of visual and verbal modes, therefore, becomes central to how young readers are guided through plot development, character motivation, emotional meaning, and narrative resolution.

These findings contribute to the study of digital children's literature by showing the analytical value of Martinec and Salway's functional framework for examining page-based digital picturebooks. Much existing research on digital picturebooks has emphasized reader engagement, literacy outcomes, classroom use, interactivity, or affective responses. While these perspectives are important, they do not always explain how meaning is structurally organized within the texts themselves. By treating each page as a multimodal unit

and examining both status relations and logico-semantic relations, this study provides a more precise account of how image and text jointly produce narrative meaning. This is particularly relevant for open-access digital platforms such as Let's Read, where storybooks often circulate widely across multilingual and multicultural reading contexts. The findings suggest that even in relatively static digital picturebooks, where animation or interactive features may be limited, multimodal meaning-making remains complex. Narrative coherence, emotional depth, spatial reasoning, and reflective meaning are produced through the functional coordination of images and words.

The study also has practical implications for children's literature analysis, digital publishing, and literacy education. For researchers, the findings show that functional multimodal analysis can reveal how narrative meaning is built through relationships among modes, rather than by interpreting images or text in isolation. For digital picturebook designers and publishers, the results suggest that effective visual–verbal design requires purposeful coordination: images should not merely repeat written text, and written language should not simply label visual content. Instead, each mode can be designed to contribute distinct but complementary meanings. For literacy educators, the findings imply that digital picture books can be used to develop children's multimodal reading skills. Teachers can guide students to notice how images show emotions, how words sequence events, how maps or spatial layouts support reasoning, and how visual cues reveal characters' thoughts or motivations. Such instructional practices can help children move beyond literal reading toward more inferential and multimodal interpretation.

Several limitations should be acknowledged. First, this study analyzed only three adventure-themed storybooks from the Let's Read platform; therefore, the findings cannot be generalized to all digital picturebooks or children's literature platforms. Second, the analysis focused on co-present written text and illustrations within page-based narrative units, excluding audio, animation, interactivity, translation, and platform navigation features. Third, the study examined textual and semiotic organization rather than children's actual reading responses, comprehension, or classroom interaction with the storybooks. Future research could extend this analysis by comparing picture books across genres, languages, age levels, and digital platforms. Further studies may also combine functional multimodal analysis with reader-response data to examine whether the identified image–text patterns shape children's comprehension, inference-making, emotional engagement, or multimodal literacy development. In addition, comparative studies between print and digital versions of the same storybooks would help clarify whether digital presentation influences the organization and interpretation of image–text relations.

The central contribution of this study is to demonstrate that the selected Let's Read digital picturebooks are semiotically organized through systematic visual–verbal coordination. Equal–complementary relations establish modal interdependence, expansion relations organize narrative coherence and causal logic, and projection relations provide access to imagination, reflection, emotion, and problem-solving. These patterns show that digital picturebooks should be read and analyzed as integrated multimodal narratives in which images and words co-construct meaning. By applying a functional multimodal framework, the study moves beyond general claims about the value of picture books. It offers a more detailed explanation of how digital children's stories organize narrative meaning through image–text relations.

5. Conclusion

The analysis of the three Let's Read digital picturebooks shows that narrative meaning is constructed mainly through the interdependence of visual and verbal modes. Equal–complementary relations were the most dominant pattern, indicating that images and written text generally work together rather than separately or hierarchically. The text often provides explicit narrative direction, while the images enrich the story through emotional expressions, spatial details, character actions, and contextual cues. This confirms that digital picturebooks function as integrated multimodal texts in which meaning emerges from the coordination of modes. The logico-semantic analysis further reveals that expansion relations, especially elaboration and causal enhancement, play a central role in organizing narrative coherence. Elaboration helps

clarify characters, settings, and events, while enhancement supports the development of time, place, cause, and consequence across the story. Projection–idea relations, although less frequent, are significant in scenes involving imagination, reflection, emotional realization, and problem-solving. These patterns show that image–text relations are not randomly distributed, but are shaped by the narrative function of each story. The findings indicate that the selected digital picturebooks support meaning-making by guiding young readers through narrative events, helping them infer emotions, understand causal relationships, and interpret characters' thoughts and motivations. The use of Martinec and Salway's functional framework enables a more systematic explanation of how image and text work together in digital children's literature. Rather than focusing on reader response or learning outcomes, this research highlights the semiotic design of digital picturebooks and demonstrates how visual–verbal coordination contributes to narrative clarity, emotional depth, and communicative effectiveness.

6. Acknowledgement

The authors would like to express their sincere appreciation to colleagues who provided constructive feedback, academic suggestions, and support during the preparation and refinement of this manuscript.

7. Declaration of AI Use

The authors used OpenAI's ChatGPT during the preparation of this manuscript. The tool was used to assist with language refinement, academic phrasing, and the improvement of the flow of selected sections. These tools were used solely for editorial and language support purposes. They did not contribute to the research design, corpus selection, data coding, multimodal analysis, interpretation of findings, or formulation of conclusions. The authors reviewed and approved all tool-assisted revisions and remain fully responsible for the accuracy, originality, analytical decisions, and scholarly integrity of the final manuscript.

8. References

Aguilera, E., Kachorsky, D., Gee, E., & Serafini, F. (2016). Expanding analytical perspectives on children's Picturebook apps. *Literacy Research Theory Method and Practice*, *65*(1), 421–435. https://doi.org/10.1177/2381336916661516

Bai, J., Zhang, H., Chen, Q., Cheng, X., & Zhou, Y. (2022). Technical supports and emotional design in digital picture books for children: A review. *Procedia Computer Science, 201*, 174–180. https://doi.org/10.1016/j.procs.2022.03.025

Barthes, R. (1977). Rhetoric of the image. In *Image, music, text* (S. Heath, Trans., pp. 32–51). Hill and Wang. (Original work published 1964).

Birks, M., Chapman, Y., & Francis, K. (2008). Memoing in qualitative research: Probing data and processes. *Journal of research in nursing*, *13*(1), 68-75. https://doi.org/10.1177/1744987107081254

Bus, A. G., & Anstadt, R. (2021). Toward digital picture books for a new generation of emergent readers. *AERA Open, 7*(1), 1–15. https://doi.org/10.1177/23328584211063874

Callow, J. (2020). Visual and verbal intersections in picture books: Multimodal assessment for middle years students. *Language and Education, 34*(2), 115–134. https://doi.org/10.1080/09500782.2019.1689996

Dahlström, H. (2022). Students as digital multimodal text designers: A study of resources, affordances, and experiences. *British Journal of Educational Technology*, *53*(2). https://doi.org/10.1111/bjet.13171

Dahlström, H., & Damber, U. (2020). Meanings Made in Students' Multimodal Digital Stories: Resources, Popular Culture, and Values. *Designs for Learning*, *12*(1). https://doi.org/10.16993/dfl.145

Egert, F., Cordes, A.-K., & Hartig, F. (2022). Can e-books foster children's language? Meta-analysis on the effectiveness of e-book interventions in early childhood education and care. *Educational Research Review, 37*, 100472. https://doi.org/10.1016/j.edurev.2022.100472

Fundelius, E., Wade, T., Robbins, A., Wang, S., McConomy, M. A., & Fumero, K. (2023). Universal Design Principles for Multimodal Representation in Literacy Activities for Preschoolers. *Inclusive Practices*, *2*(1). https://doi.org/10.1177/27324745221140380

Grange, H., & Lian, O. S. (2022). "Doors started to appear:" A methodological framework for analyzing visuo-verbal data drawing on Roland Barthes's classification of text-image relations. *International Journal of Qualitative Methods, 21*. https://doi.org/10.1177/16094069221084433

Gu, X., & Catalano, T. (2022). Representing transition experiences: A multimodal critical discourse analysis of young immigrants in children's literature. *Linguistics and Education*, *71*, 101083. https://doi.org/10.1016/j.linged.2022.101083

Haiyan, L. (2018). Image-Text relations in Junior High School EFL textbooks in China: a mixed-methods study. *Journal of Language Teaching and Research*, *9*(6), 1177. https://doi.org/10.17507/jltr.0906.07

Hu, Y., & Qiu, Q. (2020). A study on verbal and image relations in multimodal texts from the perspective of intersemiotic complementarity. *Canadian Social Science, 16*(10), 50–56. https://doi.org/10.3968/11938

Huang, S.-Y. (2022). Image-text relations and interjections in animated language-learning materials. *Language Education & Assessment, 5*(1), 1–17. https://doi.org/10.29140/lea.v5n1.685

Johnson, S. J. (2017). Multimodality and footing in peer correction in reading picture books. *Linguistics and Education, 41*, 20–34. https://doi.org/10.1016/j.linged.2017.07.004

Jones, R. H. (2012). Multimodal discourse analysis. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley-Blackwell.

Kachorsky, D. (2022). Digital children's literature: Current understandings and future directions. In K. Coats, D. Stevenson, & V. S. Yenika-Agbaw (Eds.), *A companion to children's literature* (pp. 258–270). Wiley. https://doi.org/10.1002/9781119038276.ch21

Kachorsky, D., Moses, L., Serafini, F., & Hoelting, M. (2017). Meaning making with picturebooks: Young children's use of semiotic resources. *Literacy Research and Instruction, 56*(3), 231–249. https://doi.org/10.1080/19388071.2017.1304595

Kelly, L. B., & Kachorsky, D. (2022). Text complexity and picturebooks: Learning from multimodal analysis and children's discussion. *Reading & Writing Quarterly, 38*(1), 33–50. https://doi.org/10.1080/10573569.2021.1907636

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

Kucirkova, N. (2019). Children's reading with digital books: Past, moving quickly to the future. *Child Development Perspectives, 13*(4), 208–214. https://doi.org/10.1111/cdep.12339

Marissa, D. N., & Hamid, S. (2022). Bridging out-of-school digital literacy through multimodal composition for EFL students with developing proficiency. *JALT CALL Journal*, *18*(2). https://doi.org/10.29140/jaltcall.v18n2.539

Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication, 4*(3), 339–374. https://doi.org/10.1177/1470357205055928

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

McGlashan, M. (2021). Linguistic and visual trends in the representation of two-mum and two-dad couples in children's picture books. In A. J. Moya-Guijarro & E. Ventola (Eds.), *A multimodal approach to challenging gender stereotypes in children's picture books* (pp. 268–305). Routledge. https://doi.org/10.4324/9781003145875-14

Moses, L., Rylak, D., Kachorsky, D., & Serafini, F. (2020). Investigating young readers' use of visual, textual, and design resources in contemporary picturebooks. *Journal of Language & Literacy Education*, *16*(2).

Moya-Guijarro, A. J., & Martínez Mateo, R. (2022). Challenging masculine gender stereotypes in children's picture books: A social semiotic and multimodal analysis. *Atlantis, 44*(1), 164–185. https://doi.org/10.28914/Atlantis-2022-44.1.10

Moya-Guijarro, A. J., & Ventola, E. (Eds.). (2022). *A multimodal approach to challenging gender stereotypes in children's picture books*. Routledge.

Müller, M. (2021). MyPad as a reflection of multimodal action in elementary school children's foreign language learning. *Journal of Language and Linguistic Studies*, *17*(2). https://doi.org/10.52462/jlls.46

Müller-Brauers, C., Miosga, C., Fischer, S., Maus, A., & Potthast, I. (2020). Narrative potential of picture-book apps: A media- and interaction-oriented study. *Frontiers in Psychology, 11*, Article 593482. https://doi.org/10.3389/fpsyg.2020.593482

O'Halloran, K. L. . . (2008). Systemic functional-multimodal discourse analysis (SF-MDA): constructing ideational meaning using language and visual imagery. *Visual Communication*, *7*(4), 443–475. https://doi.org/10.1177/1470357208096210

Otto, C., Springstein, M., Anand, A., & Ewerth, R. (2020). Characterization and classification of semantic image-text relations. *International Journal of Multimedia Information Retrieval, 9*, 31–45. https://doi.org/10.1007/s13735-019-00187-6

Pagano, A. S., De Paula, F. F., & Ferreguetti, K. (2018). Verbal and verbal-visual logico-semantic relations in picturebooks: an English-Portuguese parallel corpus study. *Ilha Do Desterro, a Journal of English Language Literatures in English and Cultural Studies*, *71*(1), 53–76. https://doi.org/10.5007/2175-8026.2018v71n1p53

Pantaleo, S. (2018). Learning about and through picturebook artwork. *The Reading Teacher, 71*(5), 557–567. https://doi.org/10.1002/trtr.1653

Pantaleo, S. (2020). Slow looking: "Reading picture books takes time." *Literacy, 54*(1), 40–48. https://doi.org/10.1111/lit.12190

Papen, U. (2020). Using picture books to develop critical visual literacy in primary schools: Challenges of a dialogic approach. *Literacy, 54*(1), 3–10. https://doi.org/10.1111/lit.12197

Papen, U., & Peach, E. (2021). Picture books and critical literacy: Using multimodal interaction analysis to examine children's engagements with a picture book about war and child refugees. *The Australian Journal of Language and Literacy, 44*(1), 61–74. https://doi.org/10.1007/BF03652065

Santamaría-García, C. (2022). A semiotic and multimodal analysis of interactive relations in picture books that challenge female gender stereotypes. In A. J. Moya-Guijarro & E. Ventola (Eds.), *A multimodal approach to challenging gender stereotypes in children's picture books* (pp. 144–164). Routledge. https://doi.org/10.4324/9781003145875-10

Sargeant, B. (2015). What is an ebook? What is a book app? And why should we care? An analysis of contemporary digital picture books. *Children's Literature in Education, 46*(4), 454–466. https://doi.org/10.1007/s10583-015-9243-5

Serafini, F., & Reid, S. F. (2024). Analyzing picturebooks: Semiotic, literary, and artistic frameworks. *Visual Communication, 23*(2), 330–350. https://doi.org/10.1177/14703572211069623

Serafini, F., Kachorsky, D., & Aguilera, E. (2016). Picture books in the digital age. *The Reading Teacher, 69*(5), 509–512. https://doi.org/10.1002/trtr.1452

Serafini, F., Kachorsky, D., & Reid, S. F. (2018). Revisiting the multimodal nature of children's literature. *Language Arts, 95*(5), 311–321. https://doi.org/10.58680/la201829586

Serafini, F., Moses, L., Kachorsky, D., & Rylak, D. (2020). Incorporating multimodal literacies into classroom-based reading assessment. *The Reading Teacher, 74*(3), 285–296. https://doi.org/10.1002/trtr.1948

Shanahan, L. E. (2013). Composing "Kid-Friendly" Multimodal Text: When Conversations, Instruction, and Signs Come Together. *Written Communication*, *30*(2). https://doi.org/10.1177/0741088313480328

Twiner, A., Lucassen, M., & Tatlow-Golden, M. (2022). Supporting children's understanding of emotions through creative, dance-based movement: A pilot study. *Learning, Culture and Social Interaction, 37*, 100659. https://doi.org/10.1016/j.lcsi.2022.100659

Undheim, M. (2022). Children and teachers engaging together with digital technology in early childhood education and care institutions: A literature review. *European Early Childhood Education Research Journal, 30*(3), 472–489. https://doi.org/10.1080/1350293X.2021.1971730

Unsworth, L., & Macken-Horarik, M. (2015). Interpretive responses to images in picture books by primary and secondary school students: Exploring curriculum expectations of a "visual grammar." *English in Education, 49*(1), 56–79. https://doi.org/10.1111/eie.12047