

Analysis of Illustrations in English Picture Books from the Perspective of the Grammar of Visual Design: A Case Study of the *I Am* Series

Xinyue Jiao*, Yuanqi Zhao

School of Foreign Languages, Liaocheng University, Liaocheng, Shandong, China

*Corresponding Author

Abstract: Using the grammar of visual design as a framework and Richard Scarry's *I Am* series as a case study, this paper analyzes the narrative characteristics and educational value of the illustrations in this series from three dimensions: representational meaning, interactive meaning, and compositional meaning. The study finds that this series primarily constructs narrative threads through representational storytelling. It fosters an equal and healing reading atmosphere by offer gaze, eye-level perspectives, and high sensory modality. The use of Centre-margin arrangement and unframed framing creates a visual system tailored to children. The results indicate that the visual design of *I Am* series aligns with children's cognitive and aesthetic development, achieving a fusion of narrative, emotional, and educational functions, and can serve as a reference for picture book creation, reading, and teaching.

Keywords: The Grammar of Visual Design; English Picture Books; *I Am* Series; Image-Text Interaction; Multimodal Discourse

1. Introduction

Against the backdrop of deepening globalization and reforms in basic education, English picture books, leveraging the unique advantage of text-image synergy in storytelling, have become core resources for English teaching, children's cognitive development, and cross-cultural education. The educational potential of picture books is being increasingly recognized, and they have become an excellent entry point for teaching various subjects [1]. From middle school through college, more and more teachers are also beginning to recognize the value of picture books in the classroom. Incorporating picture books into classroom instruction not only

enriches students' learning experiences but also fosters emotional and cognitive connections among learners of all ages. Their educational value and theoretical research have garnered widespread attention in academic circles. During independent or parent-child reading, picture books also offer adults reading, aesthetic, therapeutic, and educational value [2]. Meanwhile, the introduction of multimodal discourse analysis and the grammar of visual design has provided a new perspective for picture book research to break away from the traditional text-centered paradigm and shift toward the construction of meaning through text-image interaction [3,4]. Consequently, the grammar of visual design has become an important tool for interpreting the narrative logic, emotional transmission, and educational functions of English picture books. Based on the grammar of visual design, this paper explores the visual narrative characteristics of classic English picture books, revealing the design principles through which text and images collaboratively construct meaning and cater to children's cognitive and emotional needs, ultimately serving the creation, reading, and teaching of picture books.

This study focuses on Richard Scarry's *I Am* series (comprising four books: *I Am a Bunny*, *I Am a Mouse*, *I Am a Puppy*, and *I Am a Kitten*). The author has significant influence in the industry and possesses unique artistic value: he has won the Edgar Allan Poe Special Award in the United States and the Lifetime Achievement Award from the International Society of Illustrators, and his creative talent and standing in the industry are highly recognized by professionals worldwide [5]. Among his works, the Golden Children's Book Series is renowned worldwide, and the *I Am* series, a key component of this series, carries forward its core creative philosophy. Characterized by anthropomorphic animal characters and panoramic compositions,

the series features exquisite illustrations tailored to children's needs, and its unique artistic expression holds significant research value. These elements collectively foster cognitive development and emotional resonance in young readers, bridging imagination with reality through consistent visual storytelling. Each volume maintains thematic coherence while introducing nuanced variations in character dynamics and environmental detail. The deliberate simplification of forms—without sacrificing expressive depth—enhances accessibility for early learners. Furthermore, the series' color palette evolves deliberately across titles, reinforcing narrative progression and developmental milestones. Its pedagogical efficacy is validated by observed improvements in visual literacy and narrative comprehension among preschool audiences.

2. The Grammar of Visual Design

Building on Halliday's Systemic Functional Grammar theory, Kress and van Leeuwen (2006) proposed the theory, the grammar of visual design. In *Reading Images: The Grammar of Visual Design*, the author clearly points out that images are not merely a subsidiary form of text, but rather a vehicle of meaning equal in importance to language, possessing their own "grammatical rules" and capable of forming "visual statements" that convey specific information, values, and ideologies [6,7].

The analytical framework for the grammar of visual design revolves around three main dimensions: representational function, interactive function, and compositional function [6,7]. The representational function focuses on the concrete depiction of participants within the image; its core lies in using visual elements to recreate or construct real-world scenes and to typify the participants within those scenes. The interactive function serves as the intermediary dimension connecting the image and the audience; its core lies in using visual elements to establish relationships between the image's participants and the audience, guiding the audience to develop specific emotional attitudes and cognitive tendencies. The compositional function focuses on the layout logic of visual elements and the unity of information transmission. Its core lies in using design elements such as position, borders, and proportions to allocate the importance and informational value of information and

determine the viewer's visual focus.

3. Analysis of the Visual Meaning of the *I Am* Series Picture Books from the Grammar of Visual Design

3.1 Representational Meaning

Representational meaning is divided into narrative representation and conceptual representation, with the fundamental distinction between the two being the presence or absence of identifiable vector relationships within the image. Narrative representation can be further subdivided into action processes, reaction processes, speech and mental processes. When there are two participants in a narrative representation—one as the agent and the other as the target—the target is the participant toward whom the vector points, constituting an action process [8]. If the vector consists solely of gaze and eye direction, it constitutes a reaction process. Speech and mental processes are characterized by character-related text narration or inner monologue, which, through the combination of images and text, present the characters' behaviors, feelings, and thoughts. Conceptual representation does not involve vectors; if narrative representation is the interpretation of phenomena, then conceptual representation is the reflection of their essence.

Based on a statistical analysis of the illustrations in the four English picture books of the *I Am* series, the series exhibits distinct characteristics in terms of representation, with narrative representation dominating [9]. Action processes are fully displayed; these children's picture books drive the narrative through clear action vectors, intuitively presenting the characters' typical behaviors, which aligns with children's cognitive characteristics regarding dynamic images. This process is consistently depicted throughout all the works. Through their gaze and posture, the figures convey states such as "observation," "gazing," and "stillness," expressing emotions like relaxation, curiosity, and joy, thereby enhancing the emotional impact of the images. Speech and mental processes serve as the core vehicles of narrative in picture book. Through narration that complements the visuals, the books directly articulate characters' actions and inner thoughts, achieving a synergistic expression of meaning between illustrations and text. This ensures that characters' speech and mental processes run throughout the

narrative, becoming crucial means of driving the story and conveying emotions. This approach differs markedly from the presentation styles of some purely visual picture books or educational illustrations, better aligning with the need for emotional and psychological expression in children's picture books [10]. Overall, the *I Am* series of picture books uses the sequence of actions as its framework, enriches emotional expression through the process of reaction, and clarifies the narrative through the integration of text and images that depict verbal and mental processes, thereby forming a model of visual meaning construction that is highly suited to children.

3.2 Interactive Meaning

Interactive meaning encompasses four dimensions: contact, social distance, attitude, and modality, which are used to explain how images construct social relationships, emotional distance, and interaction patterns between the viewer and the characters in the illustrations [8]. Contact can be categorized into two types: demand and offer. In the demand type, the character in the illustration looks directly at the reader, actively initiating dialogue and emotional interaction; in the offer type, there is no eye contact with the reader, and the content is presented from an objective perspective. Across the entire series of picture books, the offer type predominates. From picking flowers in spring and bird watching in summer to playing with leaves in autumn and admiring the snow in winter, throughout *I Am a Bunny*—with the exception of the first and last pages—the bunny's gaze remains fixed on the seasonal scenery, various natural activities, and In *I Am a Mouse* and *I Am a Kitten*, the illustrations throughout the books feature animal characters gazing at their companions, the scenery, and their own actions. This objective, non-interactive approach minimizes a sense of forced interaction, advancing the story in a gentle manner that aligns with the cognitive pace of young children. Of course, the picture books also incorporate a small number of demand interactions as embellishments. Many books in this series use direct eye contact on their opening and closing pages: the bunny on the first and last pages of *I Am a Bunny* and the puppy on the opening page of *I Am a Puppy* both look directly at the reader. In the bath scene from *I Am a Puppy*, the puppy also conveys its emotions by looking directly

into the camera. These scattered demand scenes not only enrich the layers of interaction but also bridge the psychological gap between the characters and the reader. The two approaches are organically integrated, significantly enhancing the readability of the picture books. Social distance is determined by the framing range, which directly reflects the psychological closeness between the reader and the characters. There are three options: close-up, medium shot, and long shot. This series of picture books centers on medium shots, with long shots and close-ups complementing them: Medium shots serve as the dominant framing style—such as scenes of a kitten interacting with its owner, a chipmunk greeting a mouse, or a puppy carrying a twig. These shots present a full subject in a moderate setting and establish an objective and equal viewing relationship which clearly reveals the characters' actions and is appropriate for the cognitive development of young children. Long shots are mostly for big scenes. Like a bunny in the changing seasons. Or a puppy in the snow or at the park. Such shots put characters in such huge spaces that they maximize the presentation of the environment and give a broad perspective, so the readers can better appreciate the beauty of nature. Close-ups are used for detail and emotional moments, like the disgusted look on the puppy's face when taking a bath, or the kitten grooming its fur. These intimate shots draw the readers in and create more emotional resonance. The three perspectives together are appropriate to the tone of children's picture books: mid-range shots keep the narrative thread, long shots open up the space, and close-ups convey emotion at key moments, creating a visual relationship that is warm, safe, and free from power imbalances.

Visual attitude reflects the viewer's subjective stance toward the subjects in the illustrations. The entire series maintains a consistent visual attitude, rarely employing high angle or low angle. From the perspective of the grammar of visual design, the four picture books in this series demonstrate a high degree of consistency in the use of perspectives. All illustrations throughout the books adopt an eye-level perspective, with no high-angle or low-angle shots applied, which visually eliminates the power disparity between readers and illustrated characters. A low-angle shot implies that the illustrated characters hold a superior status over the viewers, while a high-angle shot places the

viewers in a dominant position. In contrast, the eye-level perspective signifies equal status and mutual respect between the two sides. Whether depicting animal characters such as the bunny, mouse, kitten or puppy, all figures maintain the same horizontal line of sight as young readers. Although the books switch between frontal engaging perspectives and lateral observational perspectives to meet narrative needs, the equal visual relationship remains unchanged. This consistent visual design abandons the condescending adult perspective commonly seen in traditional children's books. Treating young readers with equality and gentleness, it allows children to immerse themselves in the stories as companions, and ultimately creates a soothing and harmonious interactive reading experience. Modality measures the degree of realism and the level of detail in an image, categorized into three levels: high, medium, and low. The entire picture book series features high sensory modality, with soft, realistic tones and colors that closely match the natural appearance of objects. Elements such as the texture of animal fur, the veins of plants, the tactile quality of objects, and environmental details are depicted with exquisite precision, while texture, depth, and lighting are rendered naturally. For example, in *I Am a Bunny*, the scene where the little bunny watches frogs in the pond features exceptionally accurate color reproduction of the pond water, lily pads, water lilies, and frogs. Details such as the veins on the lily pads, the texture of the frogs' skin, and the contours of the lilies are exquisitely rendered. The scene adopts a three-layer structure of foreground, midground and background, which truly presents the ecological state of real ponds. Likewise, the bathtub scene in *I Am a Puppy* features typical high-sensory modalities. This painting adopts warm and realistic color matching. It finely depicts the texture of the puppy's fur, the soft shape of bubbles and the surface texture of bath supplies, with gentle and natural light and shadow changes. The application of high-sensory presentation restores real natural scenes and daily life scenarios, which enhances the authenticity and immersion of picture book illustrations. This design not only meets readers' visual aesthetic demands, but also allows children to perceive nature and experience daily life through realistic images. It further realizes the cognitive educational value of picture books [11].

3.3 Compositional Meaning

Compositional meaning is achieved through the organization of visual elements to facilitate collaborative meaning-making among different symbolic resources. It primarily encompasses three dimensions: informational value, salience, and composition. Informational value refers to the differences in the information conveyed by visual elements in different positions; salience refers to the varying degrees of visual prominence achieved by visual elements through their position, size, color, and clarity; and composition defines the boundaries of an image through dividing lines, borders, and other means, thereby influencing the interactive relationship among the image, the reader, and the text.

This series of picture books exhibits consistent compositional characteristics, with an overall design that aligns with children's reading habits and aesthetic preferences. In terms of information presentation, the series primarily employs centered layouts with white space and left-right divisions, while top-bottom divisions are rarely used. In the vast majority of illustrations, the main character is placed at the center of the page as the carrier of the core message, while scene elements such as flowers, trees, and snowflakes are distributed around the periphery. This creates a clear hierarchical structure, preventing secondary elements from distracting children's visual attention. In terms of prominence, the series excels overall, using a variety of techniques to reinforce visual focus on the main character. The protagonist's colors contrast with the background, and the character is typically centered, allowing children to quickly identify them. In terms of composition, the series primarily employs borderless, continuous layouts, with very few instances of borders used to divide the frame, thereby presenting a complete and cohesive effect. This borderless design breaks down visual boundaries, making the image seem to extend right before the reader's eyes and creating an immersive, realistic experience.

Text is also naturally integrated into the edges of the illustrations, blending seamlessly with the images to create a harmonious visual narrative that avoids distracting children while they read. Overall, the compositional design of the *I Am* picture book series centers on children's reading needs. Through clear information layout, emphasis on the main subject, and immersive

compositional techniques, the series constructs an efficient visual narrative system tailored to children, ensuring the clear conveyance of narrative logic while aligning with children's cognitive and aesthetic characteristics.

4. Conclusion

With the vigorous development of multimodal discourse analysis, English picture books have been widely applied in English teaching and early childhood literacy education. In this context, illustrations are no longer simple supplements to texts. Instead, they have become core symbols for constructing meanings, conveying emotions and spreading cultures. This study adopts the grammar of visual design put forward by Kress and van Leeuwen as the core analytical tool. It selects the four-volume English picture book series *I Am* created by Richard Scarry as the research object. From the three functional dimensions of representational meaning, interactive meaning and compositional meaning, this paper conducts a comprehensive analysis of the symbolic features, narrative logic and educational value of the images in the picture book series.

This series presents two types of representational meanings, namely narrative representation and conceptual representation. Among these two types, narrative representation plays a leading role. Specifically, action sequences visually show the movements of characters and form the main narrative clue to push the story development. Reaction sequences use the characters' eye contact and facial expressions to deliver various emotions, including curiosity, joy and relaxation. Speech and mental processes integrate visual and text resources to directly display the characters' inner ideas and spoken content. The perfect combination of the three elements creates a vivid and lively visual narrative effect. This effect fits the cognitive features of young children, who are fond of dynamic visual content and understand the world mainly through direct sensory experience.

In terms of interaction, this picture book achieves a reasonable balance of "demand" and "offer" types. It tells the story from the perspective of an observer and removes readers' psychological barriers, so as to enhance the emotional bond between readers and story characters. The flexible use of close-up, medium and long shots produces diverse social distance

effects with different degrees of closeness. The whole work adopts a consistent eye-level perspective which builds an equal and harmonious viewing atmosphere. It also features high sensory modality and a realistic artistic style. Such creation methods restore real natural and daily scenarios, which can improve children's immersive experience and help them identify items and understand natural phenomena.

In terms of compositional expression, the picture book follows the basic rules of children's visual cognition. It applies center-margin and left-right layout arrangements to arrange information in a clear hierarchy. Color and position settings are adopted to highlight core subjects. The unframed design avoids fragmented visual effects and realizes the seamless integration of text and images. This layout maintains the logical integrity of the narrative and effectively prevents children's visual attention from being scattered.

Overall, the *I Am* series combines narrative, emotional, aesthetic and cognitive educational functions in an integrated manner. Its well-developed visual language system fully reflects the creators' thorough understanding of children's psychology and their proficient grasp of children's aesthetic and cognitive rules. The creative practice of this series provides valuable references for the visual design and narrative creation of English picture books and children's picture books in China. In the field of English education, the analytical findings of this study can also help English teachers gain a deeper understanding of the connotations conveyed through picture book imagery. This allows them to move beyond the traditional reading approach that emphasizes text over images. As a result, teachers can carry out more engaging and more in-depth multimodal instruction using picture books. In addition, this study further confirms that the grammar of visual design can serve as an effective tool for interpreting typical multimodal texts, including picture books. In this way, it provides a clear research direction for the visual analysis of foreign language teaching materials.

References

- [1] Wu H M. Practical Exploration of Teaching English Writing in Primary Schools Using Picture Books. *Teaching and Administration*, 2025, (23): 60-63.
- [2] Li F R, Gao X. An Analysis of the Value of Picture Books to Adults: Topic Mining and

- Sentiment Analysis Based on Picture Book Short Reviews. *Library Journal*, 2025, 44(07): 122-129.
- [3] Hu P P, Wang X Y. A Study of Textbook Illustrations Based on Visual Grammar Theory and Image-Text Position Theory. *Basic Education Research*, 2025, (01): 62-66+70.
- [4] Zhang Y. English Textbook Illustrations from the Perspective of Visual Grammar: A Case Study of Yilin Edition High School English Compulsory Book 1. *Asia-Pacific Education*, 2023, (09): 118-122.
- [5] Wang X. Using Picture Book Images to Enhance Thinking Skills: Teaching and Reflection on the Picture Book *I am a Bunny*. *Jiangsu Education*, 2019, (09): 58-62.
- [6] Kress G R, van Leeuwen T. *Reading Images: The Grammar of Visual Design* (2nd ed.). Routledge, 2006.
- [7] Baker M. *Reading Images: The Grammar of Visual Design*, 3rd ed., by Gunter Kress and Theo van Leeuwen. *Visual Communication Quarterly*, 2021, 28(4): 253-254.
- [8] Sun W L, Sun X H. Gender Representation in Images of PEP High School English Textbooks from the Perspective of Visual Design Grammar. *Teaching and Administration*, 2023, (06): 38-42.
- [9] He T, Cao X Y. Exploring a Multimodal Meaning Analysis Framework for English Children's Picture Books. *Foreign Languages in China*, 2025, 22(06): 60-68.
- [10] Gui Y. A Multimodal Analysis of Image-Text Interaction in English Children's Picture Books. *English Square*, 2020, (26): 60-62.
- [11] Chen H, Huang K. The Value, Text Selection, and Methodological Strategies of Children's Picture Book Reading. *Publishing Panorama*, 2024, (23): 4-10.