

Original Paper

Multimodal Cues in the Wordless English Picturebook and Their Mechanisms for EFL Learners' Narrative Comprehension: A Case Study of *The Snowman*

Jun Chen

Faculty of Foreign Languages, Huai'an University, Huai'an, China

Received: August 29, 2026

Accepted: September 22, 2026

Online Published: October 7, 2026

doi:10.22158/jetss.v8n4p1

URL: <http://dx.doi.org/10.22158/jetss.v8n4p1>

Abstract

*As multimodal texts in the aspects of art and pedagogy, wordless English picture books build a comprehensive context of stories by means of images, colors and layouts, without the help of textual language. They act as an excellent teaching resource in EFL classes, which are able to make up for the deficiency of vocabulary and discourse understanding skills of foreign language learners and help them build up English narrative logic. Most studies are concentrated on the analysis of textual meaning or text-image relationship in illustrated books, and there are very few studies which systematically analyze the narrative cooperation of multimodal elements in wordless picture books and their effects in assisting EFL learners to understand narratives. This study, based on Visual Grammar, conducts the qualitative discourse analysis on the classic wordless picture book *The Snowman*, analyzing how images, colors and layouts cooperatively help EFL learners sort out plot development, identify emotional tones and form English narrative thought. As a result, it is found that image sequences make it possible for learners to understand complete discourse without any vocabulary, and color and layout modalities help to establish the atmosphere of the story and narrative cognitive schemas.*

Keywords

English picturebooks, multimodal cues, narrative understanding, EFL reading teaching

1. Introduction

1.1 Research Background: The Current Use of Multimodal Features in EFL Reading Teaching

English picture books are multimodal texts that combine linguistic and visual features and play a significant role in EFL reading teaching. In comparison with traditional textbooks, picture books allow learners to have a lot of contextual clues by using multimodal features including images, colors and

layouts. It not only fits the children's way of cognition but also can alleviate reading anxiety of EFL readers. Thus, it enhances their reading involvement and comprehension. With the advancements in digital and multimedia technology, para-linguistic signs including graphic design, color usage, and formatting have not just remained supplementary devices for communicating meaning but have become meaning-making tools alongside linguistics signs. In the context of multimodal discourse analysis theory, meaning is produced by the interaction between different semiotic signs. It gives a theoretical background that supports using wordless picture books to facilitate EFL learners to comprehend discourse through visuals only (Zhou & Fan, 2014). For this reason, analyzing multimodal clues in picture books is quite significant for improving the teaching methods.

However, present EFL reading teaching is still inclined to emphasize text rather than pictures. Most teachers over-rely on interpreting lexical features and syntax of texts in teaching but ignore the constructive function of non-linguistic modalities like images and colors in narration.

This one-dimensional method of teaching hampers the students from making full use of the visual clues in the picture books to arrange their plot sequence and understand the emotions of the characters in the story, and thus, it constrains the development of their high-level comprehension skills. The studies conducted up to now mainly concern the textual clues with no thorough investigation into the way that multimodal clues might benefit learners of foreign languages from compensating the inadequacy of vocabulary. Even though scholars have argued that multimodal input helps students create unique meaning which cannot be created when seeing alone (Zhou & Fan, 2014), there is still not enough empirical evidence and guidance as to how to guide learners to do so in picture book reading activities. Based on this, the current study tries to identify the supportive means by which multimodal cues, including pictures, colors, and layout design, in English picturebooks aid in narrative comprehension. Adopting the classic English wordless picturebook as the object of research, this study adopts a qualitative multimodal discourse analysis method to investigate how the three meanings under the visual grammar framework collaborate in shaping readers' cognition processes. This study is intended to identify the unique roles played by multimodal cues in making up for the deficiency of language, facilitating emotional involvement and constructing logic reasoning in order to offer practical pedagogical approaches in EFL picturebook reading instruction. It is hoped that the study could assist learners to move from being passive receivers of textual messages to active constructors of multimodal messages and improve English narrative understanding ability comprehensively in this way.

1.2 Problem Statement: Lack of Focus in Existing Research on the Independent Role of the Visual Mode in Narrative

Despite the increasing use of English picture books for EFL reading education, the current literature mainly focuses on the analysis of the linguistic aspects of either plain text or the teaching model at the macro level without paying much attention to the constructive role of the visual mode in independently performing narrative roles at the micro level, especially in situations where the book contains minimal textual content. In most cases, the visual elements of English picture books have been treated as a less

important part of the book as the primary function of the visual mode is to support the representation of meaning through verbal language (Qiu, 2020). Such a one-sided understanding of the multimodal feature of picture books poses challenges for the teacher in guiding students to use visual elements to understand plot development and character emotions.

In particular, the existing body of literature fails to provide any systematic studies of the workings of the three dimensions of visual grammar, representational meaning, interactive meaning and compositional meaning, in the understanding of visual narratives in picturebooks without words. While it is true that the images in picturebooks are fixed and offer little help in terms of guidance to readers' interpretation or pose limited constraints on visual codes, there are certain rules followed by the eye when looking at the pictures (Qi, 2022). However, there are not enough detailed and qualitative studies investigating the processes involved in the change of interpersonal distance between the reader and the text facilitated by the composition, color and layout differences and the creation of non-linear narrative logic using double-page spreads. For example, there is not enough knowledge about the way in which the subtle multimodal interaction affects the interpretation of picturebooks without words – the way images create a sense of emotional tension through a change from warm to cold color tone, or how they show the state of mind of the characters through visual perspective change.

Thus, this research will further expand the exploration in this field by no longer regarding images as supplements or replications of the texts, but as an autonomous system that can form meanings on its own. Through applying the theory of visual grammar, the current study tries to explore how multimodal elements like images, color, and layout combine together on the representational, interactive, and compositional levels in order to help readers get rid of linguistic obstacles and construct a coherent narrative structure. In addition to addressing the problem that exists in current research about the underestimation of the visual mode's independent significance, the introduction of visual grammar will offer a more sophisticated theoretical foundation and practical approach to optimize the teaching strategy for reading picture books in EFL learning contexts, and highlight the need for teachers to help students transform from passively acquiring textual information into actively integrating multimodal information.

2. Theoretical Framework and Research Methodology

2.1 Theory of Visual Grammar and its Three-Dimensionality of Meanings

This research follows the theoretical framework put forward by Kress & van Leeuwen (2006). This theory originated from systemic functional linguistics and holds that images, like languages, are semiotic resources that have certain grammatical structures and social meanings. In the case of multimodal narratives of English picture books, visual aspects do not merely serve as the subsidiary parts of texts but help construct the meanings according to certain coding rules. In a special type of picture book, wordless picture book, visual aspects become the only carriers of meanings and play the ultimate role in narratives. According to this theory, the process of meaning-making of images involves three kinds of meanings, namely representational, interactive, and compositional meanings.

Representational meaning deals with the representation of reality and interpersonal relationship by images, which corresponds to the story plotting and setting in narration. When interpreting wordless picturebooks, the aspect takes the form of images conveying narrative meaning independently. Images present action lines concretely through the depiction of posture and indicate changes in time and space by means of scene change, thus overcoming the problem of limited vocabulary and understanding the whole story correctly. This form of narrative system based on visual sequence turns the abstract plot logic into visual experience and relieves the learner's mental burden.

Interactive meaning is about the interaction between images and readers, including the modulating of the emotional attitude and psychological distance. The application of color and perspective choice and composition of images is done in wordless picturebooks within this aspect. Warm colors usually give rise to a cozy and harmonious feeling and reduce psychological distance between readers and characters, while the other way around is done by cool colors or perspective change. Through this visual interpretation, learners are able to understand the story emotionally and get deep insights into characters' psychological world and story atmosphere.

Composition means the arrangement of components within a frame and their structure in pages that control the reader's movement and create the logic of the story. In the case of wordless picture books, the arrangement of pages, double-page spreads, and sequence of information presented on the pages are all evidence of the regulating role of composition. Good composition may direct the reader to perceive the information in the way the writer wants and create a coherent story of separate frames. By analyzing these three factors thoroughly, the current research aims to reveal how multimodal elements cooperate to enable EFL learners to move from superficial to deep understanding of narratives.

2.2 Research Design and Corpus Selection: A Multimodal Study of the Classic Picture Book without Words The Snowman

In order to analyze how visual semiotic elements function independently to create meaning in narratives, a qualitative multimodal discourse analysis research design will be used in this paper. The classic wordless picture book *The Snowman* written by British author Raymond Briggs will be the main object of analysis. Consisting only of visual sequences and having no texts at all, *The Snowman* is a book with a great potential in visual storytelling and thus can be considered an excellent choice to analyze representational, interactional, and compositional meanings in visual grammar theory (Zhou & Fan, 2014). The reasons why this picture book was chosen are the coherence of the plot, typicality of the imagery, and the dependence on the visual means for the development of the storyline.

This research uses a rigorous approach in its analysis. Firstly, the pages of the picturebook are analyzed page by page in detail to discover and label multimodal clues such as gestures and postures of the characters, changes in the scene color, perspectives of composition, and arrangement of space in the pages. Secondly, interpretations are made regarding the three aspects of visual grammar. In the aspect of representational meaning, the analysis revolves around the way in which the pictures reproduce the interactional sequence of actions of the snowman and the boy and communicate complete narrative

meaning by themselves. In the aspect of interactive meaning, it examines the mood set by the transitions between warm and cool colors and perspective zoom-ins and analyzes the way in which they induce the feelings of empathy in readers. In the aspect of compositional meaning, it explores the way in which the double-page spread and gaze flow construct the pace of reading and form a consistent narrative logic. With this analysis, this study tries to uncover the way non-verbal signs help EFL learners overcome linguistic barriers in multimodal environments and go beyond mere information processing to emotional involvement.

In conclusion, this study proves that the approach of visual grammar is appropriate to analyze wordless picture books. In addition, the findings set a basis for further studies in order to design appropriate teaching methods. The findings reveal that teachers should help students deconstruct visuals so that they are able to build their story schemata.

3. Analysis of the Constructive Function of Multimodality within Narrative Understanding

3.1 Representational Meaning: Self-standing and Enhancing of the Logical Flow of the Plot Through Visual Information

With regard to the representational meaning of the images in English picture books, representational meaning refers to the way in which images illustrate the events, actions, and relationships among characters objectively. For EFL students, wordless picture books eliminate the aid of the verbal mode completely, and thus plot information and logical relations depend totally upon the order of the images. At this stage, the image details take the role of narration through visual signs, posing a higher requirement for learners' capability of observing, inferring, and integrating visual information.

Based on the classic wordless picture book *The Snowman*, the entire artwork consists of visual images that narrate the development of its plot with the aid of the visual narrative ability. In the depiction of interactions between the boy and the snowman, visual images convey all the narrative information through the depiction of the gestures of characters and transitions of scenes. For example, through the depiction of several actions such as pushing a snowball, making the snowman, and putting on a scarf to the snowman in frames (see Figure 1), the readers can understand the sequences of events as well as their causal relations without any words. It makes EFL learners completely escape from the restriction of vocabulary size and comprehend the development of the plot through visual sense. According to the literature, when the function of synergism of visual grammar is not involved in multimodal construction, the compensation by other contextual modes will be inadequate (Du, 2026). In the case of wordless picture books, visual grammar itself forms a complete context.



Figure 1. The Boy Building the Snowman (Briggs, 1978)

Furthermore, on the representational meaning level, images are able to serve as tools in constructing narrative coherence. When it comes to wordless picture books, logical relations between frames, such as time sequence, space transition, and causality, are fully conveyed visually via images. Based on the changing of facial expressions of characters, the location of objects, and the representation of light and shadow, learners are able to deduce implicit causality and emotions. Through this visual reasoning process, both the interest in reading and the ability to perceive and interpret visual information are improved. Consequently, in EFL reading classes, it is necessary for teachers to assist learners in focusing on the image details and the relation among images in order to build a narrative structure by using visual information on the representational meaning level independently.

3.2 Interactive Meaning: Emotional Atmosphere Presentation Using Color Tone and Changing Perspectives

In the aspect of interactive meaning in the visual grammar scheme, color tone and perspective become significant tools for adjusting the emotional involvement and decreasing the distance. In case of EFL learners, the presentation of emotional information in the picturebook that does not include words takes place solely via visual channel. Using the principles of color psychology, the images may create an atmosphere using perspectives, thus helping learners feel the emotions of characters.

For instance, with regard to the classic wordless picture book *The Snowman*, any change of color is a direct response to the emotional ups and downs of the story. At the very start of the story, the color used is primarily that of cold colors, such as white and light blue, which creates the quiet and cold ambience of a winter morning (refer to Figure 2). Not only is this color scheme appropriate for the realistic setting of the story, but it is also suitable for setting the first interactive meaning mood of the story of being detached yet calm. As the story continues, when the boy and the snowman go into the house and engage in interaction, the images slowly include the use of warm colors, such as that of the fire from the fireplace, indoor light, and the red scarf of the snowman. The incorporation of this color creates a strong contrast

in visual, which indicates the presence of warmth and joy. Color is a very important element in the expression of visual meaning in picture books for kids.

Bright warm tones give the reader a positive emotional orientation and give warm emotional hints (Qiu, 2020) and thus help in creating a warm and harmonious interpersonal relationship between the reader and characters. In terms of EFL students, the transition between cool and warm colors gives them clear emotional cues that will allow them to go beyond the limits of language understanding and feel an emotional transition of the story from loneliness to companionship.



Figure 2. Opening Snow Scene of the Story (Briggs, 1978)

Perspective also is an important way of construction of interactive meanings. While showing the scene of the boy and the snowman flying together, the images use low-angle or wide-angle shots with the sky being wide and a full picture of the city visible. The use of such perspective makes the space feel larger and also gives subjective feelings of freedom and amazement to the readers, thus, making the learners identify with the characters' joy. However, at the end of the story when the snowman starts melting, the images move to close-up or eye-level shots that focus on the face of the boy with his sad looks and on the melting process itself. Thus, this closer perspective adds to the emotional effect and allows the readers to understand the reluctance to part better. It was found out that the images usually use a certain perspective in order to attract the readers into the story as either spectators or actors and evoke emotional response (Zhou & Fan, 2014).

The color tones and perspective shifts serve an important role in emotional representation in the multimodal narratives of picture books without words. Through the power of non-verbal visual signs, these aspects individually enable the expression and communication of complex emotions, thus facilitating EFL learners' understanding of the connotations associated with the story. In classroom application, it is necessary for teachers to assist students in identifying the visual aspects, and by analyzing the changes in colors and perspectives, foster an emotional awareness of the text.

3.3 Compositional Meaning: The Guidance of Narrative Rhythm through Page Layout and Double-Page Spread Design

Composition meaning refers to how the elements of an image are placed in the composition and how the image is organized on the page in such a way as to control the viewer's eye movement and establish narrative sense. For EFL learners, the layout of wordless picture books is not only visual in nature, but rather a crucial feature that determines the rhythm of reading and meaning making. With the right layout design, it becomes possible for readers to make a coherent story out of separate images.

Regarding the classic wordless picturebook *The Snowman*, the double-page spread format is a significant form of narrative rhythm control. At the moment when the book story comes to its climax, i.e., the flight of the boy and the snowman, the author uses a full double-page spread with a big picture (please refer to Figure 3). Thanks to such composition format that breaks the framework of the one-page format, the visual space is significantly expanded, creating the image of an endless sky. In terms of EFL learners, apart from enhancing the emotionality of the feeling of “freedom” and “wonder,” it also guides the readers to scan the image quickly, imitating the rapidness of flying by the natural prolongation of vision. In this way, the narrative rhythm is compressed and intensified. As related research emphasizes, images can be used to create meaningful narrative meanings through scene transition and synergy of visual and audio modalities. In this case, the double-page spread format, as a visual modality, manages to turn the static reading process into a dynamic one.

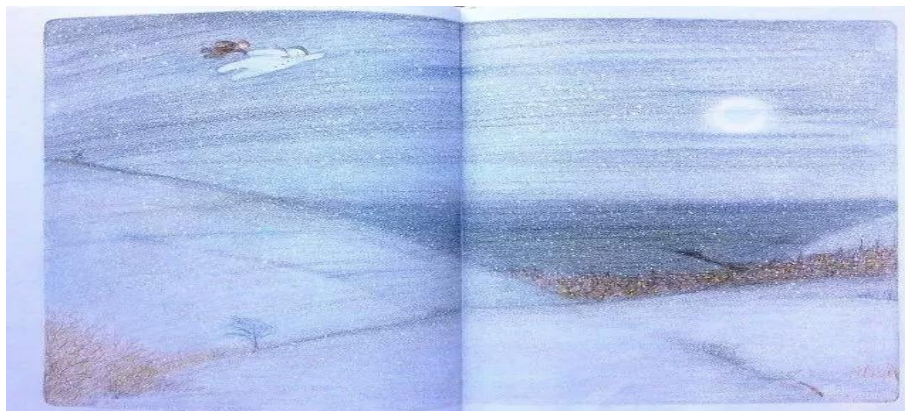


Figure 3. Flying Scene of the Boy and the Snowman (Briggs, 1978)

On the contrary, when the story nears its closure and the snowman starts to melt, the layout reverts back to the more common one-page layout, where attention is drawn lower down to the face of the boy in sorrow and the process of melting. Such constriction of composition and a closer point of view force readers to slow down while reading and pay attention to each and every slight detail. The presence of greater amount of white space or a smaller scale of the figures reflects an emotional loss and time suspended. Such contrasting in layout design marks out the stages of the narrative for the learners to build up a cognitive schema. Studies prove that compositional meaning of an image is usually greater

than textual meaning of a word since it directs readers' visual path and controls the pace of reading through page layout (Zhou & Fan, 2014). The absence of any verbal elements in wordless picturebooks makes such function even more pronounced.

The interrelationships among the elements in an image affect the order of information acquisition as well. In *The Snowman*, the author has readers focus on the relevant information by using the position of the characters in the frame, their scale and proportions, as well as their gaze directions. For instance, the boy's gaze often directs at the important objects in the image or the direction of subsequent events, and this technique of visual guidance becomes the means of performing the referential and cohesive functions that text does. For those with limited vocabulary knowledge, it makes visual layout guidance fully free of language constraints and allows them to comprehend the key meaning and plot development via visual guidance. Thus, when teaching, teachers need to help students pay attention to changes in page composition and understand how such changes contribute to the adjustment of the rhythm of narration. Understanding the reasons for the choice of double-page spread and the layout will make it possible to feel more clearly the emotions and structure of the story, thus moving from passive reception of visual information to its active integration with multimodal resources and enhancing the interpretation of English narratives.

4. Conclusion and Implications for Teaching

This paper, based on the concept of visual grammar, has done a multimodal discourse analysis of the well-known wordless English children's picturebook *The Snowman*, focusing on the synergy of images, color, and layout in the process of narrative comprehension of EFL learners. It can be seen from the findings that multimodal cues have certain strengths in overcoming the linguistic constraints of EFL learners and helping to understand the story.

Specifics of images, through the ability to function independently from the text by means of autonomous narration, successfully overcome the language barrier. In total absence of words, the depiction of concrete actions and scene transitions ensures a comprehensive visual narration, helping learners comprehend the sequence of events without the need to decode them lexically. Color tones and perspective changes contribute importantly to the emotional rendering of interpersonal meaning. The change from warm to cool color tones and the choice of particular perspectives help not only create a unique atmosphere of the story but also narrow the psychological gap between readers and characters, enabling learners to understand and empathize with the emotional swings of the latter. Finally, the design of pages and the spread of two-page pages direct the flow of the narration in terms of compositional meaning. The opening of a two-page spread with large illustrations provides a wider visual scope and creates a dynamic flow of reading, whereas the closure of the layout forces readers to slow down. Such non-linear visual flow helps learners combine different images into an overall story structure.

On the basis of the above findings, the following pedagogical implications are made in this paper. First, teachers need to get rid of the old model of instruction that only focuses on text, and apply the new

method of multimodal discourse teaching, instructing students to consciously recognize and apply visual means. More specifically, instructional periods could be arranged so that students would pay their attention to non-verbal semiotic elements like the composition of images, colors, and perspective, and how they independently contribute to the meaning and development of the narrative. In other words, they could compare various editions of images or analyze changes of tone in one picture book and see how visual means change the mood and characterization in the story. Moreover, students need to be instructed to create a story using multimodal means by making drawings or collages based on their understanding of the narrative in order to foster their multiliteracies and integrated foreign language competence. The limitation of the present research is that the sample is limited as only one classic wordless picturebook was selected for the analysis. It would be beneficial to examine different picturebooks on diverse topics in future empirical research.

References

- Briggs, R. (1978). *The Snowman*. Random House Children's Books.
- Du, M. Y. (2026). Dissonance and reconstruction in film translation and dubbing of *A Land So Rich in Beauty* from the perspective of multimodal discourse analysis. *Jin Gu Wen Chuang*, (23), 93-95.
- Kress, G. & T. van Leeuwen. 2006. *Reading Images: The grammar of visual design*. London: Routledge.
- Qi, F. F. (2022). A study on visual narration of traditional art in original Chinese picturebooks from the perspective of multimodal discourse analysis. *Publishing Journal*, (3), 43-50.
- Qiu, Q. (2020). Image-text complementarity and meaning construction—A multimodal discourse study of children's picturebooks. *Journal of Jiangxi Normal University (Philosophy and Social Sciences Edition)*, (3), 129-136.
- Zhou, H. H., & Fan, J. H. (2014). English video appreciation teaching from the perspective of multimodal discourse analysis theory. *Journal of Beijing Jiaotong University (Social Sciences Edition)*, (1), 124-128.