CHAPTER 2
LITERATURE REVIEW

This chapter presents a brief explanation of some theories that support the study. The theories are related to the meanings behind English coursebook covers viewed from a Grammar of Visual Design and Multimodal Discourse Analysis perspective.

2.1 Multimodal Discourse Analysis

In the 1990s, a significant paradigm shift occurred in the study of discourse, giving rise to Multimodal Discourse Analysis. This approach has gained considerable prominence within the fields of linguistics and communication studies. Multimodal discourse analysis offers a robust framework for examining and comprehending how various signs and symbols cooperate in the communication process. Notably, Kress & van Leeuwen (2021) played a pivotal role in catalyzing this transformation. Their emphasis extended beyond traditional linguistic elements to encompass the visual, spatial, and even auditory dimensions of communication. This perspective suggests that communication is a multi-dimensional endeavor, where diverse modes, including language, imagery, spatial arrangements, and auditory cues, collaboratively shape the meaning embedded in discourse.

Over time, multimodal discourse analysis has evolved to encompass a fundamental insight: meaning does not exist in isolation. Instead, it is deeply intertwined with the cultural, societal, and historical contexts in which it is situated (Jewitt, 2009). This realization underscores that when analyzing phenomena like coursebook covers through the multimodal lens, we attain a nuanced understanding of how semiotic modes, in conjunction with written text, reflect and convey cultural ideas and values associated with English language education.

Furthermore, the versatility of multimodal discourse analysis extends to the examination of various communication forms. For instance, when employed

in the domains of advertising and media, multimodal discourse analysis has demonstrated its invaluable utility in deconstructing how visual compositions, typographical choices, and layout designs exert influence on our perceptions and decision-making processes (Jewitt et al., 2016). Applying this analytical framework within the context of coursebook covers presents a fresh perspective on how book cover designers strategically utilize semiotic modes to engage and inform prospective readers.

The choice of the multimodal discourse analysis approach over semiotic analysis as the preferred analytical method for investigating semiotic modes in book covers was deliberate. This decision stems from the recognition that multimodal discourse analysis offers a more comprehensive framework for comprehending the intricate interplay of elements within book covers (Shewarni & Mohammad, 2021). In this context, the semiotic modes, which encompass images, text, colors, and layout, are inherently interconnected. This analysis excels in this context by allowing the researcher to delve into the dynamic interactions between these semiotic modes, offering a holistic perspective on meaning formation.

While semiotic analysis primarily concentrates on dissecting signs and symbols in media (Bobrie, 2018), multimodal discourse analysis stands out by facilitating the combination of diverse communication modalities into a single analytical framework. This convergence aligns perfectly with the research's objective, which is to unravel how a book's content and theme are conveyed through a synergy of semiotic modes. Hence, multimodal discourse analysis equips the researcher with the tools to explore the complex web of semiotic modes that collectively shape the essence of book covers.

In summary, multimodal discourse analysis stands as an invaluable framework for the research pursuits. It offers the researcher a comprehensive lens through which can grasp the intricate interplay between semiotic modes within the English coursebook cover. Through this analytical framework, gaining profound insights into the purposeful choices made by book cover

designers in shaping these materials and their deep-seated connections to cultural, social, and ideologies.

2.2 Grammar of Visual Design

Proposed by Kress and van Leeuwen, the Grammar of Visual Design framework builds upon the insights of the former semiotic school, originally designed to elucidate linguistic texts. Halliday (2004) conceptualizes language as a semiotic mode that embodies three metafunctions: the ideational, the interpersonal, and the textual metafunctions. In alignment with Halliday's theory, Kress & van Leeuwen (2021) employ distinct terminology to describe the same concepts: they use "representational" for "ideational," "interactive" for "interpersonal," and "compositional" for "textual."
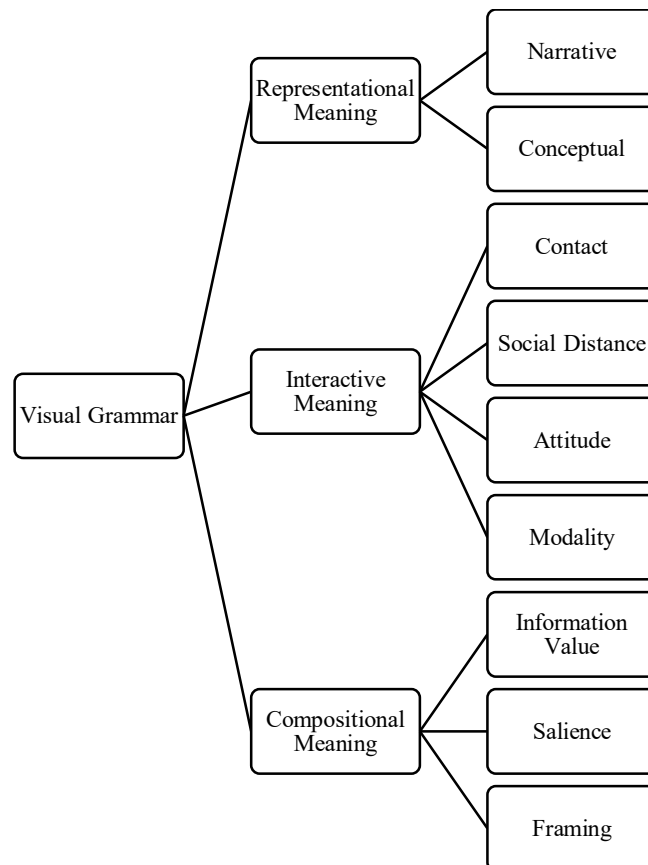
Figure 2.1 Grammar of Visual Design (Adopted from Kress & van Leeuwen, 2021)

The concept of representational meaning within the Visual Grammar framework pertains to how images depict the relationships among the depicted participants in a visual composition. This concept can be further dissected into two categories: narrative representation and conceptual representation. Within narrative representation, subcategories such include action processes, reaction processes, speech processes, and mental processes. On the other hand, narrative representation primarily deals with unfolding actions, events, processes of change, and transitory spatial arrangements. In contrast, conceptual representation presents participants in terms of their more generalized and relatively stable and timeless essence. This essence can be understood in the context of classes, structures, or meanings. Conceptual representation further breaks down into classificational processes, analytical processes, and symbolic processes.

The concept of interactive meaning within the Grammar of Visual Design framework primarily concerns itself with the social relations existing among interactants and the evaluative stances that participants take towards each other and the world as represented within the text. This facet of meaning relies on four crucial elements: contact, social distance, attitude, and modality. Firstly, "contact" delves into the demand and offer relationship between the represented participant and the viewer. Secondly, "social distance" encapsulates the social relationships at play among the producer, the viewer, and the represented participant. This is often realized through the size of the frame used in the visual composition. Thirdly, "attitude" classifies images as either subjective or objective. Interestingly, this classification is typically considered neutral but can be influenced by various angles and perspectives. Lastly, "modality," a concept stemming from linguistics, deals with the truth value or credibility of linguistically expressed statements about the world. In this context, it explores the role of modality markers, including aspects such as color saturation, color differentiation, and color modulation, among others.

Compositional meaning, a key facet within the Grammar of Visual Design framework, plays a vital role in interconnecting the representational and

interactive meanings of images. It operates through three closely linked systems, each guided by distinct principles of composition: information value, salience, and framing. Information value scrutinizes the positioning of elements and their associated values, considering spatial arrangements such as left and right, which signify given and new information, as well as top and bottom, representing ideal and real information. Additionally, the placement of elements in the center versus the margins is contemplated, taking cultural context into account. Salience, on the other hand, delves into the establishment of a hierarchy of importance among elements through variations in attributes like size, sharpness of focus, and tonal contrast. These variations serve to guide the viewer's attention and perception of significance. Lastly, framing explores the intricate relationship between the degree of connectedness and the individuality or distinctiveness of elements within the composition, shedding light on their contextual meanings. However, it is essential to acknowledge that this theory does not exempt itself from the potential pitfalls of over-interpretation or misinterpretation, considering the inherent subjectivity of the researcher and the possibility of encountering unconventional attributes bearing decisive meanings.

This framework fundamentally governs the functioning of visual communication and elucidates the intricate interconnections between all three metafunctions. Furthermore, each of these metafunctions constructs a comprehensive structure comprising multiple elements that serve as invaluable tools for images analysis.

Table 2.1 *Halliday's Language Metafunctions and Visual Grammar*

Language Metafunctions (Halliday, 2004)	Visual Grammar (Kress & van Leeuwen, 2021)	Elements of Visual Patterns
Ideational	Representational	Narrative Conceptual
Interpersonal	Interactive	Contact Social Distance Attitude Modality

Textual	Compositional	Information Value Salience Framing

2.2.1   Representational Meaning

Representational meaning within semiotic systems serves as a means to depict objects while considering the cultural context and their relationships with entities beyond the system of representation. This aligns with the internal connections between participants who represent objects and engage in actions within specific situational circumstances. The participants in this context can be categorized into two distinct types: represented participants, which encompass individuals, objects, and locations portrayed in an image, and interactive participants, referring to individuals who engage in communication with one another. The internal relationships present within visual images can be further classified into two distinct categories:

a.   Narrative Process

1) Action Process

The action process within visual grammar is portrayed by participants either originating from the vector partially or wholly, or the vector itself taking form as specific types of vectors. Within this process, two distinct kinds of participants come into play: the actor and the goal. The actor is the represented participant who emerges from the vector or constitutes the vector. In contrast, the goal represents a participant at which the vector is directed, essentially signifying the objective of the action.

2) Reactional Process

The reactional process is observed through emanations from the vector, typically represented by the eye lines of the participants featured. In this process, participants are not categorized as actors but rather as reactors, and they are not characterized as goals but rather as phenomena.

3) Speech Process & Mental Process

Speech and mental processes involve the use of oblique protrusions like thought balloons or dialogue balloons to establish a connection between individuals (or even animals) and the inner mental processes or speech depicted in comic strips. The participant emerging from these balloons is referred to as the sensor (in the context of a thought balloon) or the speaker (in the context of a dialogue balloon).

4) Conversion Process

The conversion process represents a cyclical process where all the participants involved serve as relays, and agency is somewhat vaguely signified. This type of process is known as the conversion process and is particularly employed in the representation of natural events. For example, it finds application in illustrating food chain diagrams or diagrammatic representations of the hydrological cycle. Interestingly, it can also be used to depict human interactions, wherein such interactions are represented akin to natural processes.

5) Geometrical Process

The geometrical process presents a unique communication model characterized by the absence of participants. Instead, it involves a vector that signifies directionality, often represented by an 'infinity' sign.

6) Circumstances

Circumstances are another aspect of the narrative process, and they involve secondary participants. They serve as a means to connect specific participants, especially in terms of setting or context.

b. Conceptual Process

The conceptual process provides a more stable and generalized representation of participants, categorizing them based on types, structure, and meaning. This process can be further classified into three sub-categories: classification process, analytical process, and symbolic process.

1) Classification Process

The classification process typically involves the use of taxonomy, diagrams, or tree structures. However, in some instances, it does not necessarily depict a family or diagrammatic structure. Instead, it can compare two or more participants based on criteria like age or physical appearance. In this process, the represented participants are connected through a taxonomy-like relationship. Several participants take on subordinate roles, while at least one other is referred to as superordinate. The superordinate participant is usually explicitly represented in the image.

2) Analytical Process

Within the analytical process, the focus is on revealing the part-whole structure of the participant. This process involves two distinctive participant roles: Carriers embody the entire participant, representing it as a holistic entity. In contrast, possessive attributes highlight the specific parts or components that constitute the whole participant. These attributes provide insight into the relationship between the whole and its individual elements

3) Symbolic Process

The symbolic process involves participants that symbolize specific meanings or purposes. In this context, the participant representing the nature or significance in the symbolic relation is referred to as the carrier, while the participant embodying the meaning or identity itself is termed an attribute.

2.2.2 Interactive Meaning

The interactive dimension of images encompasses what is commonly referred to as "non-verbal communication," a shared "language" used by both producers and viewers. According to Kress & van Leeuwen (2021), this dimension involves two types of participants: represented participants (which include people, places, and objects depicted in images) and interactive participants (the individuals who communicate with each other through

images, namely, the producers and viewers of these images). It consists of three key types of relations: the relations between represented participants, the relations between interactive participants and represented participants (highlighting the interactive participants' attitudes toward the represented participants), and the relations between interactive participants themselves (emphasizing the actions or behaviors of interactive participants towards each other through images). The interactive meaning within this context is actualized through four critical factors: contact, social distance, attitude, and modality.

a. Contact

The visual configuration serves two interconnected functions. Firstly, an image can take the form of a direct address, establishing contact through direct gaze or gestures from the represented participants. In such instances, we categorize this type of image as a 'demand.' Secondly, an image may choose to address the viewer indirectly, presenting the represented participants to the viewer as sources of information or objects for contemplation. This indirect approach treats them impersonally, akin to specimens in a display case, and is referred to as an 'offer.'

b. Social Distance

The selection of distance in an image conveys distinct relationships between the represented participants and the viewers, shaping how close or distant viewers feel from those depicted. In our daily interactions, societal norms and regulations dictate the appropriate distance we should maintain from one another, influencing our interpersonal dynamics. In an image, the size of the frame is consistently defined in relation to the human body. According to Kress & van Leeuwen (2021), different shot types signify varying degrees of proximity. For instance, a close shot (close personal distance) displays the head and shoulders of the subject, while a very close shot (intimate distance) captures anything closer than that. A medium close shot (far personal distance) typically cuts off the subject at the waist, and a medium shot (close social distance) does so at the knees.

The medium long shot (far social distance) presents the full figure, while the long shot (public distance) depicts the human figure occupying approximately half the frame's height. Anything 'wider' than that falls into the category of a very long shot. In summary, the depicted participants' distance from the viewer is a tangible representation of social interaction within a specific context.

c. Attitude

Attitude in visual communication pertains to the relationships between the represented participants and the viewer, and this system of perspective, often socially determined, is typically realized through the choice of angle. Selecting an angle or a 'point of view' implies the potential for expressing subjective attitudes toward the depicted participants, whether they are human or otherwise. The horizontal angle can assume either a frontal or an oblique point of view, each signifying varying degrees of involvement or detachment. When a represented participant is shown from a frontal angle, it suggests that particular attention should be paid to them, while an oblique angle usually implies a different perspective. Vertical perspective, on the other hand, relies on the camera's height and conveys different power dynamics. A high angle tends to make the subject appear small and insignificant, diminishing their individuality. Conversely, a low angle conveys an imposing and awe-inspiring presence, evoking a sense of superiority, exaltation, and triumph. An eye-level angle indicates an equal relationship and conveys a neutral and objective documentation of the scene.

d. Modality

The concept of 'modality,' borrowed from linguistics, pertains to the truth value or credibility of statements about the world, linguistically expressed. Specific degrees of modality are measured through various modality markers. These markers encompass aspects such as color saturation, differentiation, modulation, depth, illumination, and brightness.

2.2.3 Compositional Meaning

The compositional metafunction serves as a powerful lens through which one can discern and decode the intricate meanings woven into a visual composition. This metafunction enables us to dissect the entirety of an image presentation, meticulously examining how the placement of semiotic modes orchestrates the narrative and conveys profound messages.

The compositional metafunction prompts us to reflect on the intentional decisions involved in organizing semiotic modes within a composition. It underscores that each facet, whether it is the placement of images or the structure of text, plays a role in shaping the overarching meaning conveyed by the composition (Kress & van Leeuwen, 2021). Essentially, the composition transcends being a mere assemblage of elements; it serves as a canvas where meaning is intricately fashioned through the language of visual design.

The compositional metafunction stands as a crucial analytical instrument, allowing us to delve into the holistic nature of an image presentation. By meticulously scrutinizing the placement of semiotic modes, we gain access to the intricate web of meanings woven into a composition. Whether it involves purposeful image positioning, skillful framing techniques, or strategic element alignment, each decision made by the book cover designer plays a pivotal role in shaping the narrative and influencing the viewer's interpretation. Thus, the compositional metafunction equips us with the capability to unveil the subtle meanings hidden within the visual language of communication.

a. Information Value

Information value in the context of visual composition pertains to how the placement or arrangement of specific semiotic modes within an image assigns particular values or meanings to those elements. These values or meanings can undergo shifts depending on the specific zones or areas within the composition where these elements are situated. In essence, information value influences the overall composition of visual elements,

which can range from a centered arrangement to a polarized one, impacting how viewers perceive and interpret the image.

1) Centered

Within this composition framework, a central element takes its position at the very center of the composition. Surrounding this central element are other elements, referred to as non-central elements, which are strategically placed in distinct zones, including the circular zone, triptych zone, center-margin zone, and mediator zone. This arrangement and distribution of elements within these defined zones contribute to the overall visual composition and subsequently influence how viewers engage with and interpret the image.

- Circular

The non-central elements in this composition are strategically positioned both above and below, as well as to the sides of the central element. Circular structures are employed to create a nuanced and gradual distinction between the central and marginal areas of the composition. This arrangement allows for the expansion or addition of essential information, represented by the central element, with supplementary information conveyed by the non-central elements. This expansion occurs gradually and in a graded manner, contributing to the overall depth and complexity of the visual narrative.

- Triptych

A triptych composition involves the placement of non-central elements either on the right and left or top and bottom of the central element. The term "triptych" originates from pictures presented on three panels. Within a triptych composition, two main structures are observed: Simple and Symmetrical Margin-Centre-Margin Structure; in this arrangement, the non-central elements are positioned symmetrically on either side of the central element. This structure creates a balanced and harmonious visual presentation.

Polarized Structure, the central element functions as a mediator between elements representing different concepts, such as Given and New or Ideal and Real. This structure conveys a sense of mediation and interconnectedness between distinct pieces of information. In essence, a triptych composition serves to convey interconnected information through its arrangement, whether through symmetry or mediation.

- Center-Margin

The central element within this composition serves as the essence or nucleus of the conveyed information. In contrast, the non-central elements are positioned on the margins, either on the left, right, top, or bottom. These non-central elements are intrinsically linked to the central element, and their meaning is contingent upon or related to the central nucleus. Furthermore, the non-central elements are often identical or nearly identical, contributing to the creation of symmetry within the composition. This symmetry adds a sense of balance and visual harmony to the overall presentation, reinforcing the interconnectedness of these elements.

- Mediator-Polarized Elements

The concept of a "Mediator" refers to the central element within a composition that acts as a bridge between two or more polarized non-central elements. These non-central elements can represent different polarities, such as Given and New or Ideal and Real, similar to the Triptych structure. Essentially, the central element serves as the connecting point that imparts a specific meaning to the relationship between the polarized non-central elements. It plays a crucial role in contextualizing and conveying the significance of this connection.

2) Polarized

In this type of composition, there is no central element that serves as the primary focus. Instead, the non-central elements are distributed

across various zones: given-new, no horizontal polarization, ideal-real, and no vertical polarization.

- Given-New

    In this type of composition, the non-central elements are positioned on both the left and right sides, with some elements located in the margin while others are placed in the center. Unlike symmetrical compositions, these elements are not identical or nearly identical. Instead, they serve distinct purposes within the composition. The element positioned on the left signifies "Given," representing information that the viewer is already familiar with or has prior knowledge of. Conversely, the element placed on the right signifies "New," containing fresh and noteworthy information that demands the viewer's attention and focus. This arrangement creates a visual contrast between familiar and novel elements within the composition.

- No Horizontal Polarization

    There is *no horizontal polarization* (on the left or right side of the composition). There are no Given or New elements. It means there is no contradiction between the information conveyed by the elements placed on the horizontal axis.

- Ideal-Real

    In this composition, the non-central elements are positioned in both the top and bottom areas, which can be either in the margin or in the center. It's important to note that these elements are not identical or nearly identical. The element placed at the top signifies "Ideal," conveying the idealized or generalized essence of the information. On the other hand, the element placed at the bottom represents "Real," providing specific and detailed information. This arrangement creates a dynamic interplay between the idealized and specific aspects of the information being conveyed within the composition.

- No Vertical Polarization

  In this type, there is *no vertical polarization* (on the top or bottom side of the composition). There is no ideal element or real element. It means there is no contradiction between the information conveyed by the elements placed on the vertical axis.

b. Salience

Salience pertains to the deliberate emphasis placed on specific visual elements within an image to capture the viewer's attention to varying degrees. This emphasis is achieved through several techniques, including the strategic placement of elements, variations in size, color contrasts, and sharpness. To enhance the salience of a particular element, an image producer may position it in the foreground rather than the background. They can also increase its size, heighten color contrast, or enhance sharpness. The degree of salience can range from maximum, where the element is intended to be highly attention-grabbing, to minimum, where it is less prominent but still contributes to the overall composition. This nuanced control over salience allows image producers to guide viewers' focus and highlight elements of significance within the visual narrative.

c. Framing

Framing in visual communication refers to how visual elements are either connected or disconnected within an image. When elements are connected, there are no frames, dividing lines, or empty spaces that separate them. Conversely, when elements are disconnected, they are separated by frames, empty spaces, or dividing lines. The choice of connecting or disconnecting elements conveys whether they are meant to be perceived as belonging together or as distinct and separate entities within the visual composition. Kress & van Leeuwen (2021) propose that the strength of framing directly influences how an element is perceived. Strong framing suggests that an element is presented as a separate unit of information, emphasizing its distinctiveness. On the other hand, when elements in the spatial composition are strongly connected, they are presented as

belonging together, forming a unified unit of information. The degree of connection or disconnection can vary, ranging from maximum disconnection, where elements are clearly isolated from each other, to maximum connection, where elements are tightly integrated, reinforcing their collective meaning within the visual narrative. This intentional use of framing allows image producers to convey specific relationships and associations between visual elements, guiding the viewer's interpretation and understanding of the composition.

2.2.4   The Meanings of Color

Color is a powerful element that can communicate a message to the reader. In a book cover design, color is not only a decorative choice, but also a strategic one. It can help explain and reinforce the message, content, or theme of a book. For example, red can signify danger, passion, or violence, while blue can suggest calmness, trust, or sadness. The number of colors used in a book cover can also affect the impression conveyed by a book. A monochromatic scheme can create a sense of harmony, simplicity, or elegance, while a multicolored scheme can evoke a feeling of diversity, complexity, or vibrancy. Therefore, choosing the right color or colors for a book cover is an important decision that can influence the reader's perception and interest in the book.

Table 2.2 *The Meanings of Color adopted from Lindsay (2008) and Cerrato (2012)*

Color	Meaning
Red	Passion, love, anger, danger
Orange	Energy, enthuasiasm, creativity
Yellow	Happiness, optimism, warmth
Green	Nature, growth, harmony
Blue	Calmness, trust, intelligence
Purple	Royalty, spirituality, mystery
Pink	Romance, feminity, sweetness
Black	Power, elegance, mystery
White	Purity, innocence, simplicity
Gray	Neutrality, balance, sophistication
Brown	Earth, stability, reliability

Gold	Wealth, success, luxury
Silver	Technology, innovation, elegance

2.3 Coursebook Cover as Communication Medium

The coursebook cover plays a pivotal role as the initial point of interaction between readers and the book. Beyond its role as a protective layer, it acts as a dynamic and communicative medium (EL-Sakran & Ankit, 2018). Within its various semiotic modes lies the power to convey essential information about the book, including its content, genre, tone, and overarching themes, all while intriguing and capturing the reader's interest. A well-crafted book cover possesses the ability to set a book apart from its competitors, ultimately aiding in the establishment of an author's distinct brand.

The connection between a coursebook cover and its content should be seamless and strong. An effectively designed cover has the potential to vividly convey not only the content but also the tone of the book, thus drawing the attention of the right audience (Shafiei & Hoseinifar, 2022). Conversely, if a cover misrepresents the actual content of the book, it can result in disappointment and unfavorable reviews. Therefore, it is of paramount importance to ensure that the cover accurately mirrors the book's content, giving potential readers a true glimpse of what lies within.

The fundamental objective of a coursebook is to captivate the reader's attention. According to Leitão et al. (2018), an effectively designed cover plays a vital role in achieving this objective by making the book noticeable, whether on a shelf or in the digital platform, and by sparking the reader's curiosity. Moreover, a cover serves to establish the book's brand and create a visual identity that enhances its memorability, contributing to its overall appeal and recognition.

Semiotic modes employed on a book cover serve as potent tools for expressing the book's underlying theme and content (Muhassin et al., 2022). To illustrate, an elementary level coursebook may opt for vibrant colors and playful typography to evoke a sense of enthusiasm and liveliness. Conversely, a senior high school coursebook might lean towards subdued colors and formal

typography to cultivate an aura of simplicity and seriousness. The choice of graphics on the cover further aids in communicating the book's content and theme, contributing to a holistic visual representation.

The choice of semiotic modes on a coursebook cover goes beyond random selection; it frequently mirrors the cultural and social context of the readers. Each visual component, whether it's an image, symbol, or illustration, undergoes meticulous selection to resonate with the intended audience, cultivating a feeling of relevance and familiarity. For instance, coursebook covers tailored for diverse regions often incorporate culturally significant imagery and representations, ensuring that learners establish a profound connection with the material.

The coursebook cover functions as a medium of communication, the combination of diverse semiotic modes constructs a narrative that captivates, educates, and involves readers. The careful selection of visuals and the intended values all contribute to its communicative effectiveness. Understanding the coursebook cover as a communication medium provides profound insights into how different semiotic modes collaborate to convey the content and theme of the book.

2.4 Study of the Relevant Research

Before embarking on the current research endeavor, it is essential to shed light on relevant previous studies that have delved into the realm of Multimodal Discourse Analysis. These studies, which bear relevance to the present investigation, have made noteworthy contributions to our understanding of semiotic modes interact within diverse contexts:

In a study conducted by Jam et al. (2016), an insightful analysis was undertaken to uncover the hidden meanings embedded within the visual and verbal layers of the American English File textbook series. Employing Kress & van Leeuwen's social semiotics framework, the researchers delved into the semiotic resources within English Language Teaching (ELT) textbook series. Their primary objective was to unveil the conveyed meanings through these

diverse semiotic resources. Data for this study were collected using the documentary technique. The findings revealed that visual grammar within ELT textbooks employed various processes within the representational metafunction, diverse designs within the interactional metafunction, and a range of compositions within the compositional metafunction to effectively represent and interpret concepts and meanings of visuals and related verbal texts.

In another notable study conducted by Ahmad et al. (2019), the focus shifted to the analysis of best-selling fantasy novel covers. The study selected fantasy novel covers as the objects of analysis, employing Kress & van Leeuwen's Visual Grammar framework, specifically addressing representational and interactive meanings. The primary aim was to delve into the meanings conveyed through the visual representations on these covers. The research sought to uncover the nuanced meanings conveyed through the visual elements of fantasy novel covers, providing insights into the dynamics of this specific genre.

Muzdalifah et al. (2022) conducted a comprehensive visual grammar analysis to reveal gender representation within English textbooks designed for elementary school students. The study's objective was to uncover the gender representations conveyed through semiotic modes in these textbooks. Employing Kress & van Leeuwen's visual grammar theory, the research scrutinized the data to unearth how gender representations were portrayed. The study's findings revealed that there existed unequal gender representations, typically characterized by stereotypical portrayals. This portrayal of asymmetrical gender representation in the textbooks highlighted the influence of ideology and societal interests on the content and structure of the textbooks.

These previous studies serve as valuable benchmarks, offering insights into the multifaceted ways in which various semiotic modes intertwine within diverse contexts. They provide a foundation upon which the current research on English coursebook covers and their meaning can build, enriching the researcher's understanding of the complexities within these visual materials.