



MULTIMODAL ANALYSIS OF VERBAL AND VISUAL ELEMENTS IN *THE DAY YOU BEGIN*

Zarah Khodijah Yasmin¹, Rahmadsyah Rangkuti², Dian Marisha Putri³, Tasnim Lubis⁴,
Yulianus Harefa⁵

Universitas Sumatera Utara^{1, 2, 3, 4, 5}

zahrayasmin100803@gmail.com

Received: (October 2025)

Accepted: (November 2025)

Published : (December 2025)

Abstract

This study examined the verbal and visual elements in the children's picture book *The Day You Begin* (2018) by Jacqueline Woodson, illustrated by Rafael López, and explored how these intermodal connections conveyed the theme of diversity. A qualitative design was employed, utilising 14 data units, each comprising one verbal excerpt and its corresponding illustration, purposively selected from 28 narrative pages. The data were gathered through reading, identifying pages that portrayed diversity, selecting units, transcribing verbal texts, describing illustrations, and pairing them for analysis. The analysis drew on Halliday and Matthiessen's (2014) Systemic Functional Linguistics, Kress and van Leeuwen's (2006) Visual Grammar, and Painter et al.'s (2013) theory of intermodal coupling. Through data condensation, display, and conclusion drawing, the study examined ideational, interpersonal, and textual meanings. The findings showed that verbal and visual elements consistently expressed ideas of difference, belonging, and identity. Their intermodal couplings were largely convergent, reinforcing the theme of diversity. The study concluded that the integration of verbal and visual meanings represented experiences of diversity and identity through metafunctional realizations, while their convergent couplings emphasized empathy, inclusion, and acceptance of differences.

Keywords: intermodal coupling, multimodal analysis, picture book, verbal-visual elements

INTRODUCTION

Children today engage with stories through both language and visuals, reflecting how meaning in contemporary communication is shaped by multimodal interaction. Research indicates that children rely heavily on visual cues to interpret emotions, relationships, and events. Arslan-Ari and Ari (2022) found that preschoolers paid more attention to illustrations than to text in digital picture books, while Sun, Roberts, and Bus (2021) highlight that colors, facial expressions, and layout significantly help children grasp emotional and situational meanings. This tendency demonstrates that visuals are not mere decorative elements but essential semiotic resources that mediate learning and social understanding, particularly in picture books that convey values like diversity and inclusion.

Multimodal analysis provides a framework to understand this phenomenon. It explains how meaning is constructed not solely through verbal language but through the interplay of multiple modes such as image, space, and typography. Rooted in Halliday and Matthiessen's (2014) Systemic Functional Linguistics (SFL), which conceptualizes language through ideational, interpersonal, and textual metafunctions, multimodal theory extends these principles to the visual domain. Kress and van Leeuwen (2006) adapted these metafunctions into visual grammar, identifying representational, interactive, and compositional meanings in images. Painter, Martin, and Unsworth (2013) later expanded this relationship through their theory of intermodal coupling, which analyzes how verbal and visual elements cooperate, reinforce, or diverge from each other to create layered meaning in picture books.

In children's picture books, verbal and visual modes are interdependent. Nodelman (1988) argues that words and pictures work together to shape interpretation, each capable of supporting or challenging the other. Arizpe and Styles (2015) further emphasize that picture books help children not only develop literacy but also understand

emotions, cultural identity, and values. From a multimodal perspective, meaning arises from how these modes combine, creating synergies that help readers navigate complex social concepts such as difference, empathy, and belonging. Kress (2010) and Jewitt (2014) note that multimodality is not simply the coexistence of modes but a process of meaning integration, in which each mode contributes unique affordances that shape interpretation.

Several scholars have investigated multimodal meaning-making in children's literature. Yefymenko (2024) examined how verbal and visual resources in picture books represent interpersonal relationships, demonstrating how gaze, framing, and linguistic patterns together build empathy. Jamilah, Ismail, and Faizah (2024) examined visual representations in *My Next Words*, finding that image composition, distance, and perspective significantly shape how children interpret meaning. Their study underscores the need for detailed visual analysis in child-centred materials. Rajah and Mei (2023) studied Malaysian picture books and found that textual elements alone could not fully construct character identity without the support of visual imagery. Haris, Febrianti, and Yannuar (2023) explored how intersemiotic complementarity enhances meaning in Indonesian children's comic books, revealing strong convergence between textual and visual elements that reinforce emotional and narrative coherence. Martínez Lirola (2022) explored same-sex parent families in *Daddy, Papa, and Me* and *Stella Brings the Family*, revealing that compositional and interpersonal strategies in text and image collaboratively express values of love, equality, and acceptance. Tsapiv (2022) highlighted the importance of visual parameters, like color, saturation, layout, and point of view, in generating psychological proximity and emotional connection with characters. Similarly, Damayanti, Moecharam, and Asyifa (2021) analyzed the picture book *Just Ask* using Unsworth's multimodal framework, identifying how text and image cooperate ideationally to portray diversity, though their analysis was limited to ideational meaning and excluded interpersonal and textual metafunctions. Oktarina, Hari, and Ambarwati (2020) emphasize that picture books effectively motivate children to read by combining engaging visuals and accessible language. Meanwhile, Fitriana and Wirza (2020) investigated visual elements in the textbook *Pathway to English* and confirmed that visual features play a vital role in supporting ideational meaning, though they noted the need for better integration between visuals and text. These studies collectively demonstrate that multimodal discourse analysis offers an effective approach to understanding how children's literature constructs social values, yet many have not comprehensively examined how metafunctions intersect across modes.

Building on this foundation, the present study analyzes how verbal and visual elements interact in Jacqueline Woodson and Rafael López's *The Day You Begin* (2018). This picture book was chosen because it embodies the theme of diversity through both narrative and illustration. The protagonist's feeling of difference, due to race, culture, or appearance, evolves into connection and self-acceptance, making it an ideal object for investigating how multimodal storytelling promotes inclusion. For instance, the narration "Maybe it will be your skin, your clothes, or the curl of your hair" is visually reinforced by illustrations of four children with distinct physical traits, varying in skin tone, hair texture, and attire, demonstrating how verbal and visual elements jointly construct diversity as both natural and meaningful.

This study aims to examine the verbal and visual resources in *The Day You Begin* and analyze how their intermodal couplings construct the theme of diversity. It integrates Halliday and Matthiessen's (2014) linguistic metafunctions for the verbal mode, Kress and van Leeuwen's (2006) visual grammar for the visual mode, and Painter et al.'s (2013) framework of intermodal coupling to investigate how ideational, interpersonal, and textual meanings co-occur across modes. The research specifically seeks to identify the verbal and visual elements present in the picture book and to explain how their interaction contributes to shaping the message of diversity.

The scope of this study is limited to pages containing both verbal and visual data from *The Day You Begin* (2018). Each selected page is analyzed for its metafunctional realizations and intermodal coupling, whether convergent (reinforcing meanings) or divergent (contrasting meanings). This delimitation ensures a focused interpretation of how multimodal strategies work together to convey social values through both linguistic and visual means.

The study contributes to two main aspects. Theoretically, it expands multimodal discourse analysis by combining Systemic Functional Linguistics, Visual Grammar, and intermodal coupling into a single analytical framework for children's literature. This synthesis provides a systematic way to explore how meaning is collaboratively constructed across modes, offering insights that bridge linguistic and visual semiotics. Practically, the research offers valuable implications for educators, authors, illustrators, and parents. It demonstrates how picture books can be used as inclusive educational tools that nurture empathy, respect, and cultural awareness among children. By revealing how *The Day You Begin* integrates verbal and visual storytelling to celebrate diversity, this study underscores the transformative potential of multimodal narratives in shaping young readers' understanding of identity and belonging.

METHOD

This study employed a qualitative research design, which is appropriate for examining how meaning is constructed through verbal and visual modes in children's picture books. As Creswell (2018) explains, qualitative research seeks to understand social phenomena in their natural context and interpret meaning through participants' perspectives. This approach allows an in-depth analysis of *The Day You Begin* (2018) by Jacqueline Woodson and Rafael López, focusing on how words and images interact to convey messages of diversity, belonging, and inclusion. The data consisted of 14 selected units from the book's 28 narrative pages, each containing a verbal excerpt and a corresponding illustration. These data were chosen purposively to represent significant expressions of cultural identity, emotional experience, and difference. The book was examined holistically to understand how linguistic and visual features combine to communicate social values to young readers, rather than to generalize the findings.

Data collection and analysis followed the interactive qualitative model developed by Miles, Huberman, and Saldaña (2014), consisting of data collection, condensation, display, and conclusion drawing. The process began with repeated reading of the book, identification of relevant pages, transcription of verbal text, and detailed description of visual elements, including gaze, color, and spatial composition. Data condensation involved selecting and categorizing verbal and visual cues related to key themes, while data display was conducted through narrative analysis that demonstrated the intermodal relationships between text and image. The final stage involved interpreting and verifying how intermodal couplings construct meanings of diversity and inclusion, guided by the multimodal frameworks of Halliday and Matthiessen (2014), Kress and van Leeuwen (2006), and Painter et al. (2013). This systematic process ensured analytical depth, coherence, and reproducibility of the study's findings.

FINDINGS AND DISCUSSION

Verbal and Visual Elements in *The Day You Begin*

To identify how the theme of diversity is constructed in *The Day You Begin*, this study examines the verbal and visual modes through intermodal coupling analysis. Each set of data is labeled with the same number but distinguished by the letter a (verbal text) and b (visual illustration). The verbal data (a) focus on linguistic realization, such as ideational, interpersonal, and textual metafunctions, while the visual data (b) highlight how illustrations convey representational, interactive, and compositional metafunctions following Kress and van Leeuwen's (2006) framework. The analysis explores how both modes work together (convergent coupling) to express similar experiential and emotional meanings, particularly in relation to difference, exclusion, and acceptance.

Table 1 presents the overall intermodal coupling types identified across the selected data. All pairs demonstrate convergent coupling, showing that both verbal and visual elements consistently reinforce one another in representing aspects of diversity, from feelings of exclusion and hesitation to eventual empowerment and unity.

Table 1. Intermodal Coupling Types and Diversity Themes

Data	Type of Coupling	Theme Represented
1a & 1b	Convergent	Feeling different and hesitant
2a & 2b	Convergent	Awareness of physical difference and individual identity
3a & 3b	Convergent	Emotional withdrawal and lack of understanding from others
4a & 4b	Convergent	Cultural identity and self-expression as beauty and pride
5a & 5b	Convergent	Exclusion and contrasting life experiences
6a & 6b	Convergent	Internal experience vs external expectations
7a & 7b	Convergent	Cultural difference as a source of social exclusion
8a & 8b	Convergent	Misunderstanding of cultural practices and food differences
9a & 9b	Convergent	Social exclusion in group settings and the struggle for acceptance
10a & 10b	Convergent	Emotional alienation and feelings of being on the social periphery
11a & 11b	Convergent	Self-acceptance, courage, and readiness to embrace individuality
12a & 12b	Convergent	Empowerment through storytelling
13a & 13b	Convergent	Connection through shared identity and mutual understanding
14a & 14b	Convergent	Unity in diversity and resolution

As shown in Table 1, every data pair (verbal and visual) demonstrates a convergent relationship between modes, where linguistic expressions and visual compositions align to construct unified meanings about diversity. This consistent intermodal alignment reinforces the picture book's didactic purpose, to help young readers understand difference not as separation but as connection.

In the following section, Data 7 (7a and 7b) is selected as a representative example to illustrate how verbal and visual elements collaborate to depict cultural difference as a source of social exclusion. This particular data is chosen because it clearly encapsulates the emotional and social dimensions of diversity that recur throughout the narrative. The verbal clause introduces the idea of cultural exclusion through evaluative language, while the corresponding illustration visualizes the protagonist's isolation in a social setting. Together, they exemplify how *The Day You Begin* constructs diversity both textually and visually through emotional and spatial contrast.

1. Verbal Element

a. Ideational Metafunction

Ideational metafunction is concerned with how language represents real and imagined world experiences, answering questions about what is happening, who is involved, how, when, and where. In SFL, this consists of two layers of meaning: experiential meaning, which focuses on the processes, participants, and circumstances that make up the representation of experience, and logical meaning, which organizes the logical relationships between clauses (parataxis/hypotaxis).

In this analysis, the clauses selected from *The Day You Begin* all feature experiential meaning because they represent the subjective experience of the child character as the focal point of diversity and social alienation.

Data 7a

Verbal Text: " *There will be times when the lunch your mother packed for you is too strange or too unfamiliar for others to love as you do.*"

- 1) Process: Relational process, "*is too strange or too unfamiliar*" describes an evaluative state attributed to "*the lunch*".
- 2) Participants: Carrier, "*the lunch your mother packed for you*" → symbol of cultural identity and family background
Attribute: "*too strange or too unfamiliar*" → negative evaluation assigned to the lunch by outsiders
- 3) Circumstance: implicit context (school/social setting during lunch)

Ideationally, the clause constructs difference through a relational process that links the participant "*the lunch your mother packed for you*" with the attributes "too strange or too unfamiliar." The lunch functions as a marker of cultural and familial identity, while the attributes encode judgment from an external perspective. By presenting culturally specific food as "*strange*" or "*unfamiliar*," the text captures how cultural diversity can become a source of exclusion in social contexts. This highlights the child's vulnerability when personal and cultural identity is dismissed as unacceptable or odd.

b. Interpersonal Metafunction

Interpersonal metafunction concerns how language enacts social relationships, expressing attitudes, negotiating roles, and positioning the reader emotionally toward the narrative. This analysis focuses on **mood** (whether clauses are declarative, interrogative, imperative) and **modality** (expressing degrees of certainty, obligation, or evaluation). The selected clauses from *The Day You Begin* reveal how the book guides young readers' empathy and emotional alignment with the protagonist's experiences of difference and belonging.

Data 7a

Verbal Text: "*There will be times when the lunch your mother packed for you is too strange or too unfamiliar for others to love as you do.*"

- 1) Mood: Declarative mood
- 2) Modality: High modality, realized through "*too*" → signaling extremity and rejection

Interpersonally, the clause conveys exclusion by using a declarative mood that presents the evaluation as truth. The intensifier "*too*" raises the degree of strangeness and unfamiliarity, leaving no space for neutrality. This evaluative stance encodes the collective judgment of "*others*," positioning the protagonist in a marginalized position. By embedding negative appraisal within the language, the narrator highlights the social stigma children may face when their cultural practices, represented here by food, are not accepted by their peers.

c. Textual Metafunction

Textual metafunction ensures that the text is organized cohesively and coherently, enabling it to function effectively in its communicative context. In Hallidayan terms, this metafunction attends to how a message is structured thematically (**Theme–Rheme**), how it manages given and new information (**Information Structure**), and how **cohesion** is maintained through linguistic devices such as lexical repetition, parallelism, and conjunction. In *The Day You Begin*, the textual metafunction plays a crucial role in shaping the flow of the narrative, positioning the reader to process both the emotional content and the underlying social message of diversity. Through repeated framing devices, patterned clause structures, and rhythmic listing, the text establishes a narrative rhythm that both anchors the story and guides the reader’s attention toward key moments of introspection and affirmation.

Data 7a

Verbal text: “*There will be times when the lunch your mother packed for you is too strange or too unfamiliar for others to love as you do.*”

- 1) Theme: “*the lunch your mother packed for you*” → foregrounding personal and cultural identity
- 2) Rheme: “*is too strange or too unfamiliar*” → introducing exclusionary evaluation
- 3) Information structure: Moves from given (personal belonging) to new (negative social perception)
- 4) Cohesion: Connected to the larger frame “*There will be times when...*” → recurring refrain of exclusion

Textually, the clause opens with the theme “*the lunch your mother packed for you,*” which foregrounds the child’s cultural and familial identity as the point of departure. The rheme “*is too strange or too unfamiliar*” then supplies the evaluative judgment, highlighting exclusion as the key new information. This given–new progression dramatizes the contrast between self-identity and social perception. Cohesively, the clause ties into the narrative’s recurring structure “*There will be times when...*” which prepares the reader for different situations of exclusion. Through this textual arrangement, the experience of cultural difference is foregrounded as a central aspect of the diversity theme.

2. Visual Element

a. Representational Metafunction

The representational metafunction, as formulated by Kress and van Leeuwen (2006), concerns how visual images depict participants, objects, and events. There are two types of representational metafunction, such as **narrative representations**, which depict dynamic actions through vectors (lines of movement or gaze that indicate directionality and agency), and **conceptual representations**, which depict participants in a static, timeless, classificatory or symbolic state without vectors.

In *The Day You Begin*, illustrations strategically employ both narrative and conceptual representations to visually communicate the emotional and social experiences associated with diversity, hesitation, belonging, and exclusion.

Data 7b

Visual illustration:



Figure 1. Representational Analysis of Data 7b
Source: *The Day You Begin* (2018)

- 1) Type: Narrative representation
- 2) Vector: Protagonist's downward body gesture (defensive posture) + gaze vector toward the peer group; children's converging gazes back at her.
- 3) Participants: Protagonist as Actor (isolated yet observant), peer group as collective Counter-participants (observers).

The representational meaning here is realized through reciprocal vectors of gaze: the protagonist looks toward the group while the group simultaneously gazes at her, centering her as the focal participant. Her hunched posture and embrace of the lunchbox visually encode vulnerability and defensiveness, while the classmates' clustered gazes emphasize her difference. The interaction of these vectors narrates exclusion as a socially enacted process: she is both the object of attention and the subject experiencing discomfort.

b. Interactive Metafunction

The interactive metafunction, as described by Kress & van Leeuwen (2006), governs the relationship between the image, the represented participants, and the viewer. Through **contact (gaze)**, peserta dalam gambar *menatap pemirsa* (demand) atau *tidak menatap* (offer), **social distance**, jarak sosial yang tercipta lewat framing (close-up means intimacy; long-shot means impersonal distance) and **point of view (angle)**, high angle means powerlessness; low angle means dominance. These parameters contribute significantly to constructing meaning around inclusion, exclusion, empathy, and empowerment, all central to *The Day You Begin*.

Data 7b

Visual illustration:



Figure 2. Interactive Analysis of Data 7b
Source: *The Day You Begin* (2018)

- 1) Contact (Gaze): Offer (the protagonist's gaze is averted, and the other children direct their gaze toward each other or her lunch, not the viewer)
- 2) Social Distance: Medium shot, includes both protagonist and group, capturing spatial and social separation.
- 3) Angle: Eye-level, neutral relation, positioning the viewer as an equal observer.

The interactive structure positions the viewer as a witness rather than a participant. Because the protagonist and peers gaze at each other rather than the audience, the image excludes the viewer from the social exchange, intensifying its exclusivity. The medium shot frames both sides of the divide, making the distance between group and individual clear. At eye level, viewers are aligned empathetically with the protagonist, able to sense her vulnerability while still recognizing the social dynamics at play.

c. Compositional Metafunction

The compositional metafunction ensures that visual elements are organized cohesively to guide viewers in interpreting meaning. Kress & van Leeuwen (2006) identify three key principles that shape this metafunction, such as information value, salience, and framing. Information value to see how placement encodes meaning (e.g., left/right means Given-New, top/bottom means Ideal-Real, and centre/margin). Salience to see that visually attracts attention through size, color, focus, contrast, etc. Framing to see how visual boundaries (lines, spaces, overlaps) indicate connectedness or separation.

Data 7b

Visual illustration:

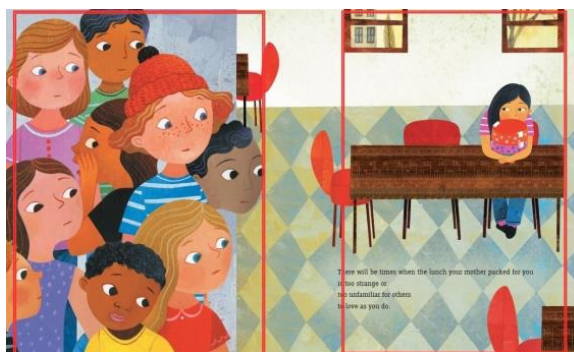


Figure 3. Compositional Analysis of Data 7b
Source: *The Day You Begin* (2018)

- 1) Information Value: Left = peer group (collective norm, shared gaze), Right = protagonist (individual marked as “different”).
- 2) Saliency: Protagonist’s defensive gesture (hunched body, clutching lunchbox) draws attention despite her isolation; children’s numerous faces and gazes amplify contrast.
- 3) Framing: Empty space of the tiled floor + wall functions as a compositional boundary, highlighting separation.

Compositionally, the spread encodes diversity through stark opposition. The left side is crowded with children, unified by shared gaze, while the right side isolates the protagonist at a table, physically distanced from her peers. The saliency of her protective posture against the background’s emptiness ensures that the viewer’s eye is drawn to her discomfort. Framing created by the table and floor tiles accentuates this spatial separation, symbolizing how cultural identity can make an individual visibly different in a shared social space.

Intermodal Coupling and the Construction of Diversity

This section examines how verbal and visual elements in *The Day You Begin* (2018) interact intermodally to communicate the theme of diversity. The analysis focuses on convergent couplings, where verbal and visual elements align across metafunctions (ideational, interpersonal, textual), reinforcing shared meanings. No clear examples of divergent coupling were found in the spreads, suggesting that the book consistently uses convergence to convey diversity as an emotionally coherent and accessible theme.

Data 7

Verbal Text: “*There will be times when the lunch your mother packed for you is too strange or too unfamiliar for others to love as you do.*”

Visual illustration: It can be seen through **Figure 1, 2, and 3**

1. Verbal–Visual Ideational/Representational: The verbal text highlights food as a marker of difference, while the visual shows peers’ gazes directed toward the protagonist and her lunchbox. Both modes encode exclusion through cultural identity.
2. Verbal–Visual Interpersonal/Interactive: The certainty in “*will be*” resonates with the protagonist’s defensive posture and the peers’ collective attention. Together, they construct empathy for her discomfort in being singled out.
3. Verbal–Visual Textual/Compositional: The verbal refrain signals the recurrence of exclusion, while the visual composition enacts it spatially: protagonist isolated on one side, peers unified on the other.

The coupling in Data 7 is convergent: the verbal text encodes cultural difference through food as a site of exclusion, while the visual concretizes this by staging the protagonist’s isolation within a peer group. Together, they emphasize diversity as a visible and socialized difference, not only what someone has (food, culture) but how others perceive it.

Overall, the findings reveal that *The Day You Begin* employs intermodal convergence to portray diversity as both a personal and collective experience. Verbal language articulates internal feelings of hesitation, exclusion, and acceptance, while visuals externalize these emotions through spatial composition, color, and gaze direction. Across

the data, the narrative progression, from exclusion to empowerment, demonstrates how harmony between words and images reinforces the theme of diversity as empathy, understanding, and shared humanity.

This consistent alignment between modes highlights the pedagogical function of the book: it teaches young readers to recognize difference as part of belonging, not as separation. By coupling linguistic evaluation with visual symbolism, *The Day You Begin* transforms everyday experiences of otherness into opportunities for connection, courage, and inclusivity.

CONCLUSION

This study confirms that *The Day You Begin* (2018) employs a cohesive multimodal strategy in which verbal and visual elements align across ideational, interpersonal, and textual metafunctions to represent diversity. Verbal elements communicate the protagonist's experiences of exclusion, hesitation, and empowerment through existential and relational processes, modality, and cohesive devices. Visual elements complement these meanings using gaze, body posture, social distance, and compositional framing, creating clear narratives of social and emotional realities. Analysis of the selected data shows consistent convergent intermodal couplings, where both modes reinforce shared meanings. This convergence highlights diversity as a lived and emotionally grounded experience rather than an abstract concept. A key advantage of this study is demonstrating how picture books can strategically combine text and image to promote empathy, social understanding, and inclusive values, making them effective pedagogical tools for literacy and social-emotional learning.

The study has certain limitations. It focuses on selected spreads from a single book, which may limit the generalizability of findings to other texts or cultural contexts. Additionally, it does not explore reader-response perspectives, leaving how children interpret multimodal meanings unexamined. Nevertheless, the findings offer practical applications: educators can use the book to facilitate discussions on inclusion and empathy; authors and illustrators can enhance meaning-making through careful text–image alignment; and parents can engage children in conversations about diversity, self-expression, and belonging. Future research may extend this approach to comparative studies, include audience reception analysis, or investigate divergent intermodal couplings in other picture books, providing further insight into how multimodal strategies communicate complex social themes.

REFERENCES

- Arizpe, E., & Styles, M. (2015). *Children reading picturebooks: Interpreting visual texts*. Routledge.
- Arslan-Ari, I., & Ari, F. (2021). The effect of visual cues in e-books on pre-K children's visual attention, word recognition, and comprehension: An eye tracking study. *Journal of Research on Technology in Education*, 54(5), 800–814. <https://doi.org/10.1080/15391523.2021.1938763>
- Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Damayanti, I., Moeharam, N., & Asyifa, F. (2021). Unfolding Layers of Meanings: Visual-Verbal Relations in Just Ask—A Children's Picture Book. *Indonesian Journal of Applied Linguistics*, 11(2), 371-380. doi:<https://doi.org/10.17509/ijal.v11i2.39195>
- Fitriana, W., & Wirza, Y. (2021, April). An analysis of multimodal text in EFL textbook of secondary school in Indonesia in assisting students' text understanding. In *Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)* (pp. 627-632). Atlantis Press. <https://doi.org/10.2991/assehr.k.210427.095>
- Goodreads Choice Awards. (2018). Goodreads Choice Awards 2018: Best Picture Books. Retrieved April 25, 2025, from <https://www.goodreads.com/choiceawards/best-picture-books-2018>

- Halliday, M. A. K., & Matthiessen, C. M. (2014). *Halliday's introduction to functional grammar*. Routledge.
- Haris, N., Febrianti, Y., & Yannuar, N. (2023). Exploring augmentation of meaning through intersemiotic complementarity in children comic book series. *Indonesian Journal of Applied Linguistics*, 12(3), 739-751. doi:<https://doi.org/10.17509/ijal.v12i3.39951>
- Jamilah, S., Ismail, N. M., & Faizah, C. (2024). Multimodal Analysis of an English Textbook Used for EFL Young Learners. *New Language Dimensions*, 5(1), 50-63. <https://doi.org/10.26740/nld.v5n1.p50-63>
- Jewitt, C. (2014). *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Routledge.
- Martínez Lirola, M. (2022). Compositional and interpersonal analysis of the picture book Stella brings the family: deconstructing affection. <https://doi.org/10.7764/onomazein.55.01>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Nodelman, P. (1988). *Words about pictures: The narrative art of children's picture books*. University of Georgia Press.
- Oktarina, P. S., Hari, N. P. L. S., & Ambarwati, N. M. W. (2020). The effectiveness of using picture book to motivate students especially young learners in reading. *Yavana Bhasha: Journal of English Language Education*, 3(1), 72-79.
- Painter, C., Martin, J. R., & Unsworth, L. (2013). *Reading visual narratives: Image analysis of children's picture books*. Equinox Publishing.
- Rajah, K. S., & Mei, C. C. Y. (2023). UNDERSTANDING THE LITERARY ELEMENT 'CHARACTER' IN MALAYSIAN PICTURE BOOKS: A MULTIMODAL ANALYSIS. *MOJES: Malaysian Online Journal of Educational Sciences*, 11(1), 13-24.
- Sun, H., Roberts, J., & Bus, A. G. (2021). The role of illustrations in children's story comprehension: A meta-analysis. *Educational Research Review*, 34, 100403. <https://doi.org/10.1016/j.edurev.2021.100403>
- Sun, H., Roberts, A. C., & Bus, A. (2022). Bilingual children's visual attention while reading digital picture books and story retelling. *Journal of Experimental Child Psychology*, 215, 105327. <https://doi.org/10.1016/j.jecp.2021.105327>
- Tsapiv, A. (2022). Multimodal imagery in picture books for children. *Cognition, Communication, Discourse*, (25), 80-88. <https://doi.org/10.26565/2218-2926-2022-25-06>
- Woodson, J., & López, R. (2018). *The Day You Begin*. Nancy Paulsen Books. <https://play.google.com/store/books/details?id=n2IBDwAAQBAJ>
- Yefymenko, V. (2024). Multimodal text-image synergy in representing interpersonal relations in picture books. *Cognition, Communication, Discourse*, (28), 102-109. <https://doi.org/10.26565/2218-2926-2024-28-07>