

Multimodal Construction of Emotional Meaning in Animated Storytelling: Language and Visual Representation in *Inside Out*

Anisa Putri Cahyani¹, Ghaitsa Tsurayya²

Animation Department

Institut Seni Indonesia Surakarta, Indonesia

Email: anisaputri@ds.isi-ska.ac.id & tsurarchive@gmail.com

Abstract

Animated films are multimodal texts that construct meaning through the combination of language, visuals, movement, music, and sound effects. Among contemporary animated films, *Inside Out* (Pixar Animation Studios, 2015) yields a compelling representation of emotional responses by personifying Joy, Sadness, Anger, Fear, and Disgust as characters that govern emotional experiences. This study examines how emotional meaning is constructed through the interaction of verbal and visual modes in *Inside Out*. Employing a qualitative descriptive design, the research adopts Multimodal Discourse Analysis (MDA) as its conceptual framework to examine selected scenes from the film. The analysis focuses on dialogue, colour symbolism, character design, facial expressions, body movement, camera framing, and visual composition as linked semiotic resources. Data were collected from the film's screenplay and corresponding visual sequences to explore how linguistic and visual elements collaboratively communicate emotional experiences. The findings disclose that emotional meanings emerge through the combination of multiple semiotic resources rather than through spoken language alone. Visual elements such as colour palettes, gesture, gaze, and spatial arrangement reinforce the emotional value conveyed in dialogue, enabling audiences to discern complex psychological states more effectively. Furthermore, the film demonstrates how multimodal storytelling facilitates emotional involvement by combining verbal narration with symbolic visual representation. The study adds to multimodal discourse studies by showing how animated films function as sophisticated multimodal texts and point out the importance of multimodal literacy in understanding contemporary audiovisual communication.

Keywords: *animated film; emotions; Inside Out; multimodal discourse analysis; visual representation*

INTRODUCTION

Communication today goes beyond words. Modern media combine language, images, colour, movement, gesture, music, and sound to create meaning. As a result, scholars now see communication as multimodal, not just language-based. To understand messages in today's media, we need to look at how these elements work together. Multimodal Discourse Analysis (MDA) is now a key approach for studying communication in areas like advertising, social media, education, film, and animation.

Kress and van Leeuwen (2006) say that communication combines different ways of making meaning to show social reality. Language is only one part; images, colour, layout, gaze, typography, movement, and sound are also important. These modes work together to create clear messages. This approach challenges traditional discourse analysis, which focuses mostly on language, by showing that meaning comes from many modes shaped by society and culture.

Animated films are especially useful for studying multimodal communication. Unlike live-action films, animation lets creators design every visual detail, such as how characters look, colours, movement, lighting, and camera angles. Each part of the visual world is planned to help tell the story and guide viewers' understanding. This makes animated films a strong example of how visual and verbal elements work together to create meaning.

Pixar Animation Studios is known for animated films that blend strong visual storytelling with emotional stories. One of its most praised films, *Inside Out* (2015), shows human emotions in a new way by turning five main emotions—Joy, Sadness, Anger, Fear, and Disgust—into characters inside the mind of an eleven-year-old girl named Riley Andersen. Instead of showing emotions as abstract ideas, the film makes them into unique characters with their own colours, facial expressions, body language, voices, and personalities. This creative approach helps audiences understand complex emotions through different modes working together.

The narrative follows Riley's emotional adjustment after her family relocates from Minnesota to San Francisco. The transition creates psychological conflict that challenges the dominance of Joy, who initially believes happiness should govern every experience. Throughout the story, Sadness gradually assumes a more significant role in Riley's emotional development, demonstrating that emotional well-being depends not on suppressing negative emotions but on integrating them into healthy psychological functioning. This narrative reflects contemporary psychological understandings of

emotional regulation while simultaneously presenting these concepts through multimodal storytelling techniques.

Inside Out stands out for its strong use of visual symbolism. Each emotion is shown with a main colour that matches common cultural meanings. Joy is bright yellow for optimism and warmth, Sadness is blue for melancholy, Anger is red for frustration, Fear is purple for anxiety, and Disgust is green for aversion. These colours work together with dialogue, facial expressions, body language, gestures, and camera angles to show emotions. As a result, audiences often understand how characters feel before anything is said.

Researchers have studied Inside Out from many angles, such as psychology, emotional intelligence, education, child development, and film studies. Psychological studies show how the film helps people understand emotional regulation and mental health, while educational research points out its value for teaching children about emotions. Film scholars have examined its story, symbolism, and audience responses. However, fewer studies focus on how emotional meaning is built through the use of different modes together, using Multimodal Discourse Analysis. Most current research looks at the story or psychology, but pays less attention to how language, visuals, gesture, colour, and layout work together.

This research aims to fill that gap by looking at how emotional meanings are created through both words and visuals in Inside Out. Instead of studying dialogue or images separately, this study looks at how different elements work together to build emotional experiences. It focuses on how dialogue, character design, facial expressions, body movement, gaze, colour, visual layout, and camera angles all help show emotions in selected scenes.

The theoretical foundation of this study is Multimodal Discourse Analysis, particularly the framework proposed by Kress and van Leeuwen (2006), which conceptualises communication as the interaction of multiple semiotic resources. Halliday's Systemic Functional Linguistics further informs the analysis by explaining how language fulfils ideational, interpersonal, and textual functions that interact with visual communication. Together, these perspectives provide an analytical framework for examining how verbal and visual resources jointly construct emotional meaning within animated storytelling.

This study uses a qualitative descriptive method because it aims to understand meanings, not numbers, in communication. It examines selected scenes from *Inside Out* to identify patterns in how emotions are shown across different modes. The analysis looks at both the dialogue and visual features like colour, gesture, facial expression, framing, composition, and movement. By combining language and visual analysis, the study shows how emotional meaning comes from the way these different modes work together.

The findings of this study are expected to help in theory and practice. In theory, the study shows how Multimodal Discourse Analysis can be used to understand animated films by showing how words and visuals work together to create emotional meaning. In practice, the research helps people become more aware of how meaning in today's media goes beyond just spoken language. This understanding is especially important in digital communication, where visuals play a big part in how we interpret and connect with messages.

METHOD

A qualitative descriptive research design was employed to examine how emotional meaning is constructed through verbal and visual resources in Pixar Animation Studios' *Inside Out* (2015). This approach facilitated the interpretation of meanings embedded in multimodal texts, prioritising qualitative analysis over the quantification of variables. The study specifically described the interaction between linguistic and visual modes that collectively represent emotions throughout the film.

The analytical framework was Multimodal Discourse Analysis (MDA), drawing primarily on Kress and van Leeuwen's (2006) theory of multimodality. This framework posits that meaning is generated through the interaction of multiple semiotic resources, such as language, images, colour, gesture, movement, spatial organisation, and visual composition. Given that animated films intentionally integrate these communicative resources, MDA offers an effective approach for examining the simultaneous representation of emotions across different modes.

Halliday's (1994) Systemic Functional Linguistics (SFL) also informed the analysis, with particular attention to the three metafunctions of language: ideational, interpersonal, and textual meanings. These linguistic functions were examined in conjunction with visual elements to investigate how dialogue interacts with character design, facial expression, gaze, camera perspective, and colour symbolism in

constructing emotional experiences. Instead of analysing verbal and visual elements independently, the study treated them as complementary meaning-making resources within a unified communicative event.

Data Source

The primary data source was the animated feature film *Inside Out*, produced by Pixar Animation Studios and released by Walt Disney Pictures in 2015. The film, with a running time of approximately 95 minutes, narrates the emotional development of Riley Andersen through the personification of five primary emotions: Joy, Sadness, Anger, Fear, and Disgust. Verbal data included dialogues, monologues, and significant verbal expressions spoken by the characters. These linguistic elements were obtained from the official screenplay and verified through repeated viewing of the film to ensure consistency with the audiovisual text. Visual data included character appearance, facial expressions, body movement, gaze, colour schemes, camera framing, composition, lighting, and spatial arrangement observed directly from selected scenes. study concentrated on scenes that prominently depict emotional conflict and transformation. Purposive sampling was used to select sequences that clearly demonstrate the interaction between verbal and visual resources. Five key scenes were identified for detailed analysis:

1. Riley's first day at her new school.
2. Joy and Sadness' journey through Long-Term Memory.
3. Bing Bong's sacrifice.
4. Joy's realisation of Sadness' emotional significance.
5. Riley's reunion with her. These scenes were selected as they represent significant emotional turning points in the narrative and contain multimodal features directly relevant to the research objectives.

Data Collection

Data collection was conducted in several stages. Initially, the film was viewed multiple times to develop a comprehensive understanding of the storyline, emotional progression, and multimodal communication strategies employed throughout the narrative.

Subsequently, the screenplay was examined to identify dialogues directly related to emotional expression. Relevant verbal data were transcribed and aligned with corresponding visual sequences to ensure contextual accuracy.

In the next stage, screenshots were captured from selected scenes to facilitate detailed visual analysis. Each screenshot documented specific visual features, such as character posture, gaze direction, facial expression, colour dominance, spatial positioning, and camera angle.

Finally, all verbal and visual data were organised into analytical tables. Each table included the dialogue excerpt, scene description, observed visual features, multimodal interpretation, and preliminary analytical notes. This organisation enabled systematic comparison across scenes and facilitated thematic analysis.

Data analysis

Data analysis followed the interactive model proposed by Miles, Huberman, and Saldaña (2014), which consists of data condensation, data display, and conclusion drawing. During data condensation, multimodal features that directly contributed to emotional representation were identified, while scenes not relevant to the research objectives were excluded. The selected data were then categorised according to linguistic and visual components.

The second stage involved displaying the data in descriptive analytical tables. Verbal expressions were analysed using Systemic Functional Linguistics, while visual elements were interpreted through Kress and van Leeuwen's visual grammar, which includes representational, interactive, and compositional meanings.

The final stage involved interpreting how these semiotic resources interacted to construct emotional meaning. The analysis emphasised the interdependence of verbal and visual elements in shaping audience interpretation, rather than evaluating them in isolation.

Analytical Framework

The study examined emotional representation through five major multimodal components:

- **Verbal representation**, including dialogue, lexical choice, interpersonal expressions, and emotional statements.
- **Visual representation**, including character appearance, costume, facial expression, body posture, and gesture.
- **Colour symbolism**, emphasising how dominant colours signify different emotional states.
- **Spatial composition**, including framing, salience, perspective, and information value within each scene.

- **Movement and cinematic techniques**, including camera angle, shot distance, character movement, and transitions that reinforce emotional meaning.

These analytical categories were integrated throughout the discussion and were not treated as isolated variables.

Trustworthiness of the Study

To enhance the credibility of the findings, the researcher engaged in prolonged interaction with the data through repeated viewings of the film and continuous comparison between screenplay dialogue and visual sequences. Triangulation was achieved by integrating linguistic analysis, visual analysis, and theoretical interpretation within the Multimodal Discourse Analysis framework.

Furthermore, interpretations were consistently supported by observable evidence from the selected scenes, rather than by subjective assumptions. The analytical procedures were systematically documented to ensure transparency and to enable future researchers to replicate similar multimodal investigations.

RESULTS AND DISCUSSION

The analysis demonstrates that *Inside Out* constructs emotional meaning through the integration of spoken language and visual elements. Rather than relying solely on dialogue, the film employs colour, facial expressions, body movement, framing, and spatial layout to convey emotions. Five key scenes were selected to illustrate the interplay of these multimodal elements.

1. Emotional Conflict on Riley's First Day at School

Riley's first day at her new school marks the onset of emotional instability following her move from Minnesota to San Francisco. Initially, Joy attempts to maintain Riley's positivity, but Sadness increasingly influences her responses. This scene demonstrates how the combination of verbal and visual elements can effectively represent emotional conflict.

Table 1. Multimodal Analysis of Riley's First Day at School

Multimodal Element	Observation	Interpretation
Dialogue	"I miss home."	Riley verbally expresses homesickness, indicating emotional vulnerability.

Facial Expression	Trembling lips, teary eyes, lowered eyebrows	Visual realisation of sadness before explicit crying occurs.
Color	Dominance of pale blue and grey classroom colours	Reinforces loneliness and emotional discomfort.
Body Movement	Riley lowers her head and pauses while speaking	Indicates hesitation and emotional suppression.
Camera Angle	Medium close-up focusing on Riley's face	Directs audience attention toward her emotional state.
Meaning Construction	Verbal and visual modes complement each other to portray homesickness and psychological transition.	

The interplay between dialogue and visuals enables viewers to perceive Riley's emotional state before any explicit verbal expression of distress. Colour and facial expressions further clarify and intensify the conveyed emotions. (Anker & Flach, 2010)

2. Joy and Sadness' Journey Through Long-Term Memory

As Joy and Sadness traverse Long-Term Memory, the visual appearance of their environment shifts significantly. These evolving landscapes reflect Riley's emotional fluctuations and underscore the narrative function of spatial design.

Table 2. Multimodal Analysis of Long-Term Memory Scene

Multimodal Element	Observation	Interpretation
Dialogue	Joy repeatedly insists that Riley must stay happy.	Joy attempts to maintain emotional control.
Character Movement	Joy walks energetically while Sadness moves slowly.	Contrasting personalities are visually emphasised.
Color Palette	Bright yellow surrounding Joy; cool blue surrounding Sadness.	Colour differentiates emotional functions.
Spatial Composition	Endless memory shelves dominate the background.	Represents the complexity of human memory.

Camera Framing	Wide shot showing characters surrounded by memories.	Highlights the vastness of Riley's psychological world.
Meaning Construction	Emotional conflict is expressed through movement, colour, and spatial organisation simultaneously.	

The contrast between Joy and Sadness illustrates how multimodal resources visually represent psychological processes that are challenging to convey through dialogue alone.

3. Bing Bong's Sacrifice

A particularly significant scene occurs when Bing Bong sacrifices himself to enable Joy's escape from the Memory Dump.

Table 3. Multimodal Analysis of Bing Bong's Sacrifice

Multimodal Element	Observation	Interpretation
Dialogue	"Take her to the moon for me."	Represents Bing Bong's unconditional love and farewell.
Facial Expression	Joy cries while Bing Bong smiles gently.	Contrasting expressions intensify emotional impact.
Gesture	Bing Bong jumps off the wagon.	Symbolises self-sacrifice.
Lighting	Soft fading light gradually disappears.	Suggests loss and disappearance from memory.
Camera Shot	Long shot followed by close-up on Joy.	Allows audiences to witness sacrifice before emphasising emotional reaction.
Meaning Construction	Emotional meaning is produced through synchronised dialogue, gesture, lighting, and cinematography.	

The emotional intensity of this scene relies primarily on visual storytelling rather than verbal exposition (Corliss, 2015).

4. Joy Realises the Importance of Sadness

This scene serves as a narrative turning point, as Joy recognises that sadness is not a deficiency but an essential component of emotional healing.

Table 4. Multimodal Analysis of Joy's Realisation

Multimodal Element	Observation	Interpretation
Dialogue	"Sadness...Mom and Dad...they came to help because of Sadness."	Joy recognises the positive social function of sadness.
Facial Expression	Joy's smile gradually disappears into thoughtful silence.	Indicates cognitive realisation.
Color Transition	Bright yellow becomes softer and less dominant.	Symbolises emotional maturity.
Camera Focus	Close-up on Joy holding the memory sphere.	Highlights internal reflection.
Symbol	Blue-yellow memory orb	Represents emotional integration.
Meaning Construction	Visual symbolism communicates emotional growth beyond spoken dialogue.	

The memory orb containing mixed colours functions as a central symbol in the film, signifying the coexistence of multiple emotions rather than their opposition (Gayego et al., 2020, p. 1823).

5. Riley Reunites with Her Parents

In the final scene, Riley accepts both her positive and negative emotions, demonstrating her ability to integrate diverse emotional experiences.

Table 5. Multimodal Analysis of Family Reconciliation

Multimodal Element	Observation	Interpretation
Dialogue	Riley admits that she misses Minnesota.	Honest emotional expression restores family communication.
Body Language	Riley hugs both parents tightly.	Physical affection reinforces emotional reconciliation.
Facial Expression	Parents and Riley cry together while smiling.	Mixed emotions indicate emotional authenticity.
Color	Warm lighting dominates the family home.	Symbolises safety and emotional comfort.
Composition	Three characters positioned centrally in one frame.	Represents family unity.
Meaning Construction	Language, gesture, and visual composition jointly represent emotional healing.	

The film's use of language, gestures, and warm visual elements suggests that emotional recovery is facilitated by accepting sadness rather than suppressing it (Hasiana et al., 2024).

Overall Discussion

Across all five scenes, emotional meaning is constructed through the integration of multiple communicative modes. Verbal language conveys characters' thoughts and feelings directly, while visual elements such as colour, facial expressions, movement, camera angles, and spatial arrangement reinforce and expand upon these messages. This analysis supports Kress and van Leeuwen's (2006) assertion that communication is inherently multimodal rather than exclusively linguistic.

The analysis further indicates that colour functions as a significant symbolic system within the film. Each emotion is associated with a distinct colour, enabling

viewers to interpret characters' emotional states rapidly. Yellow represents joy and optimism, blue signifies sadness and empathy, red denotes anger, green corresponds to disgust, and purple indicates fear. These colour associations remain consistent throughout the narrative, facilitating audience comprehension of emotions even in the absence of explicit verbal cues (Gayego et al., 2020).

Facial expressions and body movements also serve as powerful means of conveying emotion in the film. Characters frequently exhibit emotional shifts through subtle changes in eye movement, posture, gestures, and overall movement before any verbal articulation. (D. et al., 2016, pp. 340-348) Cinematic techniques such as close-ups, framing, and varied camera angles direct audience attention to significant emotional moments and enhance viewer engagement with the characters (Emotions Conveyed By Camera Angles: Techniques For Filmmakers To Manipulate Feelings, 2025).

In summary, the findings indicate that *Inside Out* constructs emotional meaning through the dynamic integration of verbal and visual modalities. The film exemplifies multimodal storytelling by combining dialogue, imagery, movement, and cinematic techniques to convey complex emotional experiences. This underscores the effectiveness of animated films as multimodal texts.

CONCLUSION AND SUGGESTION

This study examined how emotional meaning is constructed through the interaction of verbal and visual modes in Pixar Animation Studios' *Inside Out* by applying the framework of Multimodal Discourse Analysis (MDA). The findings demonstrate that emotional representation in the film is not conveyed through dialogue alone but emerges from the integration of multiple semiotic resources, including language, color symbolism, facial expressions, body movement, spatial composition, and cinematic techniques. These resources work collaboratively to create coherent emotional narratives that enable audiences to understand the characters' psychological experiences more effectively.

The analysis of the selected scenes reveals that verbal language provides explicit emotional information, while visual elements reinforce, complement, and sometimes extend the meanings expressed through dialogue. The consistent use of

color to distinguish different emotions, together with expressive character design and carefully composed visual framing, strengthens the audience's interpretation of Riley's emotional development. Significant narrative moments, such as Riley's first day at school, Bing Bong's sacrifice, Joy's realization of the importance of Sadness, and Riley's reconciliation with her parents, illustrate how multimodal resources collectively communicate emotional complexity and character growth.

The findings support the theoretical perspective of Kress and van Leeuwen (2006), which argues that meaning is constructed through the interaction of multiple semiotic modes rather than through language in isolation. The study also demonstrates the relevance of Multimodal Discourse Analysis for investigating animated films, where visual design is intentionally integrated with verbal communication to produce rich and meaningful storytelling. Consequently, *Inside Out* can be understood as an effective example of multimodal communication that promotes emotional literacy by illustrating that emotions function interdependently rather than independently.

Despite these contributions, this study has several limitations. The analysis focuses on a limited number of scenes from a single animated film, which may not fully represent multimodal practices across different genres or animation studios. In addition, the study primarily examines verbal and visual resources without conducting a detailed analysis of auditory elements such as background music, sound effects, and voice quality, which also contribute to emotional meaning.

Future research is therefore encouraged to investigate a broader range of animated films and incorporate additional multimodal resources, particularly audio and musical elements, to provide a more comprehensive understanding of emotional meaning-making in audiovisual media. Comparative studies involving different animation studios, cultural contexts, or audience interpretations would also enrich the development of multimodal discourse studies. Furthermore, future research may explore the pedagogical application of animated films in developing multimodal literacy, emotional intelligence, and critical media literacy within educational settings.

REFERENCES

- Anker, E., & Flach, F. F. 2010. *Representing emotions in visual narratives*. New York: Routledge.
- Bateman, J. A. 2014. *Text and Image: A Critical Introduction to the Visual/Verbal Divide*. London: Routledge.
- Corliss, R. 2015. Inside Out Review: Pixar's emotional masterpiece. *Time Magazine*. Available at <https://time.com/>. (Accessed 2 August 2026).
- Forceville, C. 2020. Visual and multimodal communication: Applying multimodal metaphor theory. *Review of Cognitive Linguistics*, 18(2), 325–347.
- Gayego, J. C., Sarmiento, M. P., & Cruz, A. B. 2020. Emotional representation in Pixar's *Inside Out*. A multimodal discourse analysis. *International Journal of Innovation, Creativity and Change*, 14(8), 1818–1832.
- Halliday, M. A. K. 1994. *An Introduction to Functional Grammar*, 2nd ed. London: Edward Arnold.
- Hasiana, I., Pratiwi, D., & Nugroho, A. 2024. Emotional literacy through animated films: A multimodal perspective on *Inside Out*. *Journal of Language and Literature Studies*, 24(1), 67–82.
- Jewitt, C. (ed.). 2014. *The Routledge Handbook of Multimodal Analysis*, 2nd ed. London: Routledge.
- Kress, G., & van Leeuwen, T. 2006. *Reading Images: The Grammar of Visual Design*, 2nd ed. London: Routledge.
- Machin, D. 2016. *Introduction to Multimodal Analysis*, 2nd ed. London: Bloomsbury Academic.
- Miles, M. B., Huberman, A. M., & Saldaña, J. 2014. *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed. Thousand Oaks, CA: Sage Publications.
- O'Halloran, K. L. 2004. *Multimodal Discourse Analysis: Systemic Functional Perspectives*. London: Continuum.
- Pixar Animation Studios. 2015. *Inside Out*. Emeryville, CA: Pixar Animation Studios and Walt Disney Pictures.
- Unveil the emotional depth of *Inside Out*. Available at <https://movies.disney.com/inside-out>. (Accessed 2 August 2026).
- van Leeuwen, T. 2005. *Introducing Social Semiotics*. London: Routledge.