

A Study of Multimodal Metaphor in the Film *Deep Sea* from the Cognitive Perspective

Zixuan Zhong

School of Foreign Languages, University of Shanghai for Science and Technology, 200093

Received: 20 Aug 2026; Received in revised form: 17 Sep 2026; Accepted: 20 Sep 2026; Available online: 27 Sep 2026

©2026 The Author(s). Published by Infogain Publication. This is an open-access article under the CC BY license

(<https://creativecommons.org/licenses/by/4.0/>).

Abstract— Grounded in Conceptual Metaphor Theory (CMT) and the multimodal metaphor framework, this study examines 58 valid multimodal metaphors identified in the 2023 Chinese animated film *Deep Sea*. Adopting a mixed-methods approach of quantitative statistics and qualitative case analysis, it categorizes the metaphors, describes their distributional features, and explores how pictorial, linguistic, and sonic modalities collaborate to construct metaphorical meaning. The findings show that ontological metaphors account for the largest proportion (52%), followed by structural and orientational metaphors (24% each). Pictorial modality serves as the fundamental carrier, while pictorial-linguistic and pictorial-linguistic-sonic combinations are the dominant realization patterns. The synergy of multiple modalities visualizes the protagonist's inner psychological state and enhances audience empathy. This study extends the application scope of multimodal metaphor research to animated film discourse and provides cognitive insights for animation production and mental health communication.



Keywords— cognitive linguistics; multimodal metaphor; animated film; *Deep Sea*; conceptual mapping

I. INTRODUCTION

Metaphor is widely acknowledged as a core cognitive mechanism shaping human thought and language, rather than a mere rhetorical device (Lakoff & Johnson, 1980). With the development of multimodal discourse analysis, metaphor research has extended from purely linguistic texts to multimodal media such as animated films, which integrate visual, linguistic and sonic resources to construct narratives and convey emotions (Forceville & Urios-Aparisi, 2009).

The 2023 Chinese animated film *Deep Sea* has drawn extensive public attention for its innovative visual style and profound thematic exploration of adolescent depression. However, existing studies on this film mainly focus on its visual artistry and color design (Dong, 2023; Jiang, 2023), with scarce systematic investigation into its multimodal metaphorical mechanisms from a cognitive perspective. To address this research gap, this study explores the classification, distribution and meaning construction of multimodal metaphors in *Deep Sea*, aiming to reveal how multiple semiotic modalities work

collaboratively to convey abstract psychological themes. This paper is structured into five sections: introduction, theoretical foundation, methodology, results and discussion, and conclusion.

II. THEORETICAL FOUNDATION

2.1 Conceptual Metaphor Theory

Proposed by Lakoff and Johnson (1980), Conceptual Metaphor Theory (CMT) posits that human understanding of abstract concepts is essentially grounded in the mapping of concrete, embodied experiences from a source domain to an abstract target domain. That is, people comprehend unfamiliar, intangible concepts by projecting structural features of familiar, tangible concepts onto them. Based on their cognitive functions, conceptual metaphors are categorized into three core types: ontological, structural, and orientational metaphors.

Ontological metaphors treat abstract concepts such as emotions, events and psychological states as tangible entities or substances, enabling people to quantify, identify

and comprehend intangible experiences. Structural metaphors organize abstract concepts using the structured logical framework of concrete concepts, establishing systematic correspondences between the two domains. Orientational metaphors project spatial orientations derived from bodily experiences (e.g., up-down, in-out) onto abstract concepts such as emotional states and social status, forming metaphorical expressions rooted in universal physical experience.

2.2 Multimodal Metaphor Framework

Building on CMT, Forceville and Urios-Aparisi (2009) extend metaphor research to multimodal contexts, defining multimodal metaphor as a metaphor whose source and target domains are predominantly or exclusively represented through two or more semiotic modes. They identify nine modes involved in multimodal communication, among which three are most salient in film discourse: pictorial (visual images, character design, scene composition), linguistic (dialogues, subtitles, narration), and sonic (background music, sound effects).

Unlike monomodal linguistic metaphors, multimodal metaphors rely on the synergy of multiple sensory channels to construct meaning, enhancing the vividness and emotional impact of metaphorical expressions. This study adopts this three-mode analytical framework to examine the multimodal metaphorical system in *Deep Sea*, focusing on how the three modalities interact to complete cross-domain mapping.

III. METHODOLOGY

3.1 Research Questions

This study addresses three interrelated research questions:

1. What are the primary types of multimodal metaphors in the animated film *Deep Sea*?
2. What is the distribution pattern of these metaphor types and their modal realizations?
3. How do pictorial, linguistic, and sonic modalities interact to construct metaphorical meaning from a cognitive perspective?

3.2 Data Collection

The research corpus is the full-length 2023 Chinese animated film *Deep Sea*, with a runtime of approximately 112 minutes. The researcher watched the film repeatedly and carefully to identify potential multimodal metaphors, following the three identification criteria proposed by Forceville (2009): first, the source and target domains belong to distinct categorical domains; second, the two domains form an “A is B” metaphorical structure that enables feature mapping from source to target; third, the

metaphor is manifested through at least two semiotic modes.

After screening and verification, a total of 58 valid multimodal metaphors were identified. Screenshots of typical metaphorical scenes were captured, and corresponding dialogues and soundtrack information were transcribed and recorded for further in-depth analysis.

3.3 Analytical Procedures

This study combines quantitative and qualitative analytical approaches. Quantitatively, the 58 identified metaphors are classified into three categories based on the CMT classification standard, and the frequency of each subtype and their modal realization patterns are counted and summarized. Qualitatively, representative cases from each metaphor category are selected for in-depth analysis, focusing on how each modality contributes to cross-domain mapping, and how the interaction of multiple modalities completes the meaning construction of metaphors and exerts narrative and expressive functions.

IV. RESULTS AND DISCUSSION

4.1 Classification and Distribution of Multimodal Metaphors

Based on the three-way classification of CMT, the 58 identified multimodal metaphors are divided into three categories: ontological metaphors, structural metaphors, and orientational metaphors. The specific distribution is elaborated below.

Multimodal ontological metaphors are the most frequent type, accounting for 52% of the total, followed by structural and orientational metaphors, each making up 24%. The predominance of ontological metaphors aligns with the core thematic focus of *Deep Sea*: visualizing abstract psychological concepts such as depression and spiritual redemption. Among ontological metaphors, person metaphor ranks first (21%), as the film extensively uses anthropomorphic marine creatures to reflect real-world social groups and human traits. Entity metaphor (14%) and container metaphor (8%) also account for considerable proportions, serving to materialize invisible emotions and construct spatial cognitive frameworks respectively. For structural metaphors, journey metaphor (10%) is the most prominent, as the entire narrative is built upon the framework of a sea voyage. For orientational metaphors, down metaphor (10%) and up metaphor (9%) are widely used to imply changes in consciousness and emotional states.

Regarding modal realizations, statistical results show that pictorial modality is present in the source domain of all 58 multimodal metaphors, indicating that visual images

serve as the fundamental carrier for metaphorical meaning in animated films. Two-mode and three-mode combinations are the dominant realization forms: pictorial-linguistic (PL) and pictorial-linguistic-sonic (PLS) combinations each account for the largest proportion, while pictorial-sonic (PS) combinations are also common. Single pictorial mode accounts for the smallest proportion, as visual information alone may cause ambiguous interpretation without the anchoring of linguistic or sonic modalities. Specifically, linguistic modality plays a role in clarifying metaphorical reference and eliminating ambiguity, while sonic modality mainly functions to enhance emotional atmosphere and strengthen the perceptual impact of metaphors.

4.2 Meaning Construction Through Multimodal Interaction

4.2.1 Multimodal Ontological Metaphors

Ontological metaphors transform intangible psychological concepts and abstract states into perceivable entities, which is the primary way the film visualizes the protagonist's inner world. Three typical subtypes are analyzed below.



Picture 1



Picture 2

The most representative person metaphor is CATFISH IS A PICKY HUMAN CUSTOMER. In the film, the catfish guests at the Deep-Sea Hotel are designed with human-like facial features and postures: they hold mobile phones, show disdainful expressions, and make complaints about the hotel's service, which constructs the source domain of "rude, picky human customers" through pictorial modality. Linguistic modality reinforces this mapping through the character's line, "Then I have to leave a negative review", which explicitly associates the catfish's behavior with that of dissatisfied human consumers. Through the synergy of pictorial and linguistic modalities, this metaphor not only shapes vivid character images in the fantasy deep-sea world, but also implicitly reflects the snobbish and indifferent social groups in reality, laying a cognitive foundation for understanding the protagonist Shen Xiu's sense of alienation.



Picture 1



Picture 2

The representative container metaphor is VISUAL FIELD IS A CONTAINER, embodied in the scene where Shen Xiu first wakes up in the fantasy world. When Shen Xiu floats on a small yellow swimming ring, her field of vision is limited to the scope surrounded by the ring, which takes the swimming ring as the boundary of a container and constructs the source domain of "closed container space" through pictorial modality. As the camera pulls back, Shen Xiu's field of vision expands, and the full view of the fantasy deep-sea world gradually unfolds, visually presenting the "expansion of the container". Sonically, the melodious and grand background music rises as the vision expands, synchronizing with the visual information to enhance the immersive feeling of "entering a new world". This container metaphor constructs the audience's cognitive perspective consistent with the protagonist, allowing the audience to experience the transition from reality to fantasy world through Shen Xiu's perspective.



Picture 1



Picture 2

The most functionally significant entity metaphor is DEPRESSION IS A STICKY RED MONSTER, which visualizes the abstract concept of depression into a tangible entity called the "Downcast Wretch". Pictorially, the Downcast Wretch is presented as a mass of sticky red substance that constantly chases Shen Xiu and wraps her tightly when it catches her. The red color carries the symbolic meaning of danger and suffocation, and the enveloping form intuitively presents the oppressive feeling of depression, turning invisible psychological pain into a visible external threat. Sonically, every appearance of the Downcast Wretch is accompanied by rapid, tense and depressing background music and sharp sound effects, which strengthen the sense of crisis and helplessness brought by depression through auditory stimulation. Linguistic modality names this entity "the Downcast Wretch" through character dialogues, explicitly anchoring the metaphorical target domain to depressive emotion. The tripartite collaboration of pictorial, sonic and linguistic modes completes the mapping from "tangible monster entity" to "abstract depression", providing the general audience with an intuitive cognitive approach to understand depression.

4.2.2 Multimodal Structural Metaphors

Structural metaphors use the structured framework of concrete things to organize abstract concepts, helping the audience understand complex psychological processes and

narrative logic through familiar experiential structures. Two typical subtypes are analyzed below.



Picture 1



Picture 2

The core structural metaphor throughout the film is SPIRITUAL REDEMPTION IS A SEA VOYAGE, which constructs the protagonist's process of psychological healing as a sea voyage to the Eye of the Deep Sea. Pictorially, a large number of voyage scenes form the complete source domain of "journey": the Deep-Sea Hotel ship sails through the *Deep Sea*, encounters storms and attacks by the Downcast Wretch, and finally heads for the Eye of the Deep Sea, corresponding to the stages of escaping from trauma, encountering inner resistance, and approaching spiritual liberation respectively. Sonically, the soundtrack changes with the voyage process: passionate music during smooth sailing, tense and urgent music when encountering obstacles, and soothing music when approaching the destination, which strengthens the rhythm of the journey through auditory cues. Linguistic modality explicitly points out the "Eye of the Deep Sea" as the final destination through character dialogues, and anchors its metaphorical meaning of "untying the heart knot and ending the nightmare". Through the structural mapping of journey to redemption, the film transforms the abstract and complex psychological healing process into a familiar voyage experience, greatly reducing the cognitive difficulty of understanding abstract psychological themes and enhancing the audience's emotional resonance.



Picture 1



Picture 2

The business metaphor NAN HE'S INNER WORLD IS A FAILING HOTEL BUSINESS reveals the spiritual state of Nan He, the owner of the Deep-Sea Hotel. Pictorially, the dilapidated hotel facilities, Nan He's busy but frustrated operation state, and customers' disgusted expressions together present the image of a failed businessman, constructing the source domain of "unsuccessful business operation". Linguistic modality reinforces this structure through dialogues: customers complain about the terrible dishes, and Nan He tries every means to make money to pay off debts. These business details are mapped to Nan He's psychological level: the poorly managed hotel corresponds to his barren inner world, and his fawning attitude towards customers corresponds to his disguise of true self in real life.

Sonically, the sad and low background music sets off the bleakness of the business situation, and also implies the depression and helplessness hidden under Nan He's optimistic appearance. This metaphor not only enriches the character depth of Nan He, but also reveals that he is also a person trapped in spiritual predicament like Shen Xiu, making the mutual salvation between the two characters more logically convincing.

4.2.3 Multimodal Orientational Metaphors

Orientational metaphors are rooted in human bodily spatial experience, projecting spatial orientations such as up-down and in-out onto abstract psychological and social concepts. Such metaphors are implicit but widely distributed, playing an important role in shaping emotional atmosphere.



Picture 1



Picture 2

The in metaphor SAFETY IS INSIDE A CLOSED SPACE is based on the spatial experience of "inside = safe, outside = dangerous", and extends to the social and psychological level. In the cruise ship scene at the beginning of the film, pictorial modality presents the cruise ship as a closed, warm and bright space, where families gather happily and interact closely, forming the source domain of "inside the safe group". In contrast, the sea outside the ship is dark and unpredictable, symbolizing danger and isolation. Linguistic modality presents the lively social interaction inside the ship through the noisy conversations and laughter of other passengers, highlighting the sense of security brought by group integration. However, Shen Xiu is always excluded from the group, wandering on the edge of the "inside" space. This orientational metaphor connects spatial experience with psychological security silently, and uses implicit spatial coding to shape the protagonist's lonely and isolated psychological state.



Picture 1



Picture 2

The down metaphor UNCONSCIOUSNESS AND DESPAIR ARE GOING DOWN is consistent with the universal cognitive experience summarized in CMT that "down = low emotion, unconsciousness". When Shen Xiu falls into the sea and sinks into the deep-sea vortex, the

picture tone turns dark and blurred, and her body keeps falling, which constructs the source domain of “sinking down” through pictorial modality, corresponding to her gradual loss of consciousness and entry into the fantasy world. When she floats back to the sea surface, the picture becomes brighter, corresponding to her partial recovery of consciousness. Sonically, the music becomes deep, slow and depressing during the sinking process, and becomes clear and bright when floating up, forming an auditory echo with the spatial orientation. This metaphor relies on the audience’s universal embodied experience, enabling them to naturally perceive the change of the protagonist’s consciousness state without explicit explanation.

4.3 General Discussion

Overall, the multimodal metaphor system in *Deep Sea* presents a clear hierarchical structure, with different types of metaphors undertaking differentiated narrative and expressive functions. Ontological metaphors, as the most frequent type, undertake the core function of visualizing abstract psychological concepts and building the basic fantasy world view. Structural metaphors provide an organizational framework for the narrative and psychological growth line, making the abstract redemption process easy to understand. Orientational metaphors, as the underlying cognitive schema, implicitly shape emotional atmosphere and character states.

In terms of modal interaction, the three modalities form a complementary and collaborative cognitive system. Pictorial modality provides intuitive source domain images and is the core carrier of metaphor construction. Linguistic modality plays an anchoring role, clarifying metaphorical reference and avoiding semantic ambiguity, especially for metaphors with relatively implicit visual meaning. Sonic modality is responsible for emotional enhancement, strengthening the emotional connotation of metaphor through auditory stimulation and triggering the audience’s emotional resonance. The synergy of multiple modalities not only improves the efficiency of meaning transmission, but also enhances the aesthetic and artistic effect of the film. The findings confirm that CMT has strong explanatory power for multimodal metaphor phenomena in animated films, and the cognitive mechanism of cross-domain mapping is still applicable in multimodal contexts.

V. CONCLUSION

This study systematically analyzes the multimodal metaphors in the animated film *Deep Sea* from a cognitive linguistic perspective. The main findings are as follows: first, ontological metaphors account for the largest proportion among the three types of multimodal metaphors, serving as the main tool to visualize adolescent

depression; second, pictorial modality is the foundational carrier of all metaphors, and multi-modal combinations are the dominant realization forms; third, the synergy of pictorial, linguistic and sonic modalities completes cross-domain mapping and constructs metaphorical meaning together.

Theoretically, this study enriches the empirical research of multimodal metaphor in animated film discourse and expands the application scope of CMT. Practically, it provides cognitive reference for domestic animated films to express abstract psychological themes through multimodal design. This study has limitations in the subjectivity of metaphor identification, and future research can expand the corpus and adopt more standardized analytical frameworks to improve research reliability.

REFERENCES

- [1] Baldry, A., & Thibault, P. J. (2006). Multimodal transcription and text analysis: A multimedia semiotics approach. *Equinox*.
- [2] Forceville, C. (1996). *Pictorial metaphor in advertising*. Routledge.
- [3] Forceville, C., & Urios-Aparisi, E. (Eds.). (2009). *Multimodal metaphor*. Mouton de Gruyter.
- [4] Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- [5] Dong, L. P. (2023). A study of metaphors in domestic animated film *Deep Sea*. *Home Theater Technology*, (4), 77-80.
- [6] Feng, D. Z. (2011). The construction and classification of multimodal metaphor: A systemic functional perspective. *Foreign Language Research*, (1), 24-29.
- [7] Jiang, C. (2023). Research on the influence of color design on spatial shaping of animation: Taking animated film *Deep Sea* as an example. *Appreciation of Art*, (35), 144-147.
- [8] Zhao, X. F. (2011). New development of conceptual metaphor research: Multimodal metaphor research. *Foreign Language Research*, (1), 1-10.