

A Multimodal Discourse Analysis of the Disney Movie *Mulan*

Yunjun Tian, Nannan Ma

Shenyang Aerospace University, Shenyang, Liaoning, 116001, China

Abstract

Multimedia discourse refers to the discourse that produces the overall meaning through the interaction of two or more communication modes such as image, language and sound. With the rapid development of multimodal discourse analysis in China, it has gradually become a research hotspot. This paper takes the Disney live action movie *Mulan* as the research object to explore the collaborative relationship between images and texts, to expand the research field of multimodal discourse analysis, and to verify its practicability and feasibility in the study of film discourse.

Keywords

Film discourse; Multimodal discourse analysis; *Mulan*.

1. Introduction

With the rapid development of science and technology, different discourse modes begin to appear, among which the information displayed by visual mode has inspired many researchers to explore its implied meaning. Since the 1990s, multimodal discourse analysis has been widely used in linguistics and communication studies. Van Leeuwen emphasizes that “Multimodal discourse is a kind of text that conveys information through various communication modes such as text, sound, image and animation. Multimodal discourse analysis refers to the analysis of different symbolic patterns in a communicative context, and the study of how they are combined to convey meaning” (2005). Film is a typical multimodal discourse, which integrates a variety of art forms and symbolic systems. It constructs meaning and conveys information through a variety of symbolic systems.

In September 2020, the much anticipated Disney live action version of *Mulan* was released around the world. This is the second time that Chinese women's images have swept the world since the release of *Mulan* in 1998, which has aroused extensive discussion about Chinese culture and national spirit all over the world. In the special period when the global epidemic continues to worsen, the global popularity of this cross-cultural film has a certain positive significance for the overseas spread of Chinese stories and the extensive exchange of Chinese culture.

This paper analyzes how different symbol systems cooperate with each other to construct the implied meaning in the film, and explores the social information in the film from a new perspective through the analysis of the verbal mode and visual mode of the film.

2. Multimodality and Multimodal Discourse

2.1. Multimodality and Multimodal Discourse Analysis

Mills (2006) pointed out that people use their senses to interact with the outside world through visual, auditory or tactile means. When people need to experience the world with their eyes, they use visual mode; when people need to listen to the world with their ears, they need auditory mode. When different modes are combined to express meaning in a text, multimodality is formed. Van Leeuwen (2005) suggested that multimodality means the combination of different semiotic modes, such as language and music in a “communicative

artifact or event". Multimodality focuses on the interaction and relationship among language, image, color, gesture, music and other communicative modes. It is more appropriate to regard multimodal perspective as a kind of analytical structure, which is constructed to enable those analysts who are so inclined to explore a wider range of phenomena rather than just verbal and linguistic symbol system.

Through the exploration of the definition of mode, medium and multimodal discourse, we know that multimodal discourse analysis is a comprehensive analysis of multiple symbolic modes in text or communication activities, which expands the scope of discourse analysis. MDA extends the research focus from language itself to various symbol resources, such as gesture, image, music and color. This new discourse analysis aims to explore how these different modalities work together to construct a successful communicative activity, and how people use these to convey various kinds of information and implied meanings in daily life.

2.2. Kress and Van Leeuwen's Visual Grammar

Kress and van Leeuwen (1996) used Halliday's concept of meta-function in systemic functional grammar to explain the theory of visual grammar. Visual grammar theory provides a new multimodal method for the analysis of visual images. They believe that images and other visual modes, such as language mode, can achieve these three meta functions. Language grammar determines the rules of organizing words, clauses and Examples into a complete text. Similarly, visual grammar also describes different complex visual statements composed of characters, places and things. Visual symbols have the functions of reflecting events and experiences in the real world, showing different interpersonal relationships, and forming different visual modes. The core of visual grammar is representational meaning, interactive meaning and compositional meaning.

3. Visual Analysis of The Film Mulan

3.1. Representational Meaning

According to whether there is "vector" interaction between the participants, "representation meaning" can be divided into "narrative process" and "conceptual process". Among them, "narrative process" with "vector" interaction represents a dynamic process, while "conceptual process" has no "vector" and is mostly a static process. Figure 1 and 2 are narrative process. Figure 3, 4 and 5 are conceptual process.



Figure 1. Narrative process



Figure 2. Narrative process

When participants are connected by "vectors", images represent purposeful interactions between participants. This can be divided into "action process" and "reaction process". Figure 1 is action process, the actor is the witch, the goal is Mulan. After a fight, the witch subdued Mulan. She sees through Mulan's lie of hiding her true identity, and forces her to tell the truth. It's their first meeting and their first conflict. Figure 2 is reaction process, the reactor is Mulan and the witch, the phenomenon is that Mulan kept a distance from the witch with a sword, and they just made eye contact. The disparity of strength, so Mulan can only keep a distance from the witch, at this time reflects the powerful witch.



Figure 3. Structural analysis



Figure 4. Non-structural analysis



Figure 5. Explicit symbol

There is no vector in the concept representation of visual grammar, including classification process, analysis process and symbolic process. "Analysis process" is the presentation of the participants' part and whole on the image, which is divided into "structural analysis" and "non structural analysis". Figure 3 is structural analysis, the carrier is Mulan, the possessive attributes are army uniform and sword. According to the special dress characteristics, Mulan can be identified as a soldier's identity characteristics. In addition, Mulan's embarrassment can be seen from her messy clothes. Figure 4 is non structural analysis, It only shows "attribute", but not the whole of "carrier". From the picture, even if we can only see the skirt, we can know that this is part of the witch, and this person is the witch. This kind of frame enhances the sense of mystery."Symbolic process" means "what the participant is" and "what the participant symbolizes", which can be divided into "explicit symbol" and "implied symbol". In the "explicit symbol", there are at least two participants, the identity of one carrier is clear, and the other attribute often represents some symbolic meaning. This symbolic meaning is combined with the participants with a clear identity through association. Figure 5 is explicit symbol, the carrier is Mulan, the attribute is Phoenix. Mulan's silhouette is located on the far left side of the picture, while the Phoenix is located in the center of the picture. With the gorgeous colors, Phoenix is very prominent. Phoenix is the symbol of Chinese traditional culture, which means luxury, greatness, progress and peace. Two unrelated participants appear in the picture at the same time, in order to arouse the audience's thinking and establish a connection between the two participants.

3.2. Interactive Meaning

Through the four dimensions of "contact", "distance", "modality" and "attitude", this paper explores the complex and subtle interaction between image participants and image viewer.



Figure 6. The four dimensions

"Contact" reflects a kind of imaginary contact between the image participants and the viewer. If there is interaction between the participants in the image and the viewer's eyes, it means that the characters in the image "ask" for some information from the viewer. This kind of image is classified as "ask for" image, such as Figure 6. In the picture, Mulan stared at the camera without saying a word or expression. Pass the sadness of the witch's death to the viewers in front of the camera. If the image participants do not look at the viewer, they will become the object to be

watched, and what they convey is the meaning of "providing" information. This category of images is the "provide" category. Almost all of the films are from the third perspective, and most of the images are of this kind, so we will not list them one by one.



Figure 7. Close-up



Figure 8. Close range



Figure 9. Medium close range



Figure 10. Medium range

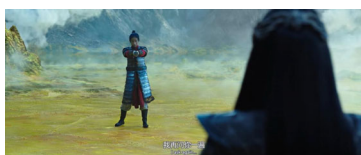


Figure 11. Long range

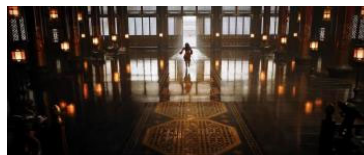


Figure 12. Large range

In the practice of social interaction in reality, the intimate relationship between people is usually expressed by the distance between the two sides. In the visual grammar, we understand the virtual social distance generated by images, and think that the virtual social relationship between the image viewer and the participants is mainly determined by the camera's viewfinder, which is manifested in seven degrees from intimacy to alienation: close-up, close range, medium close range, medium range, medium range, long range and large range. These seven lens distances are respectively reflected in the Figure 7, 8, 9, 10, 11, 6 and 12. From this, we can see the distance between Mulan and the witch from far to near, from estrangement to intimacy. Until the end, the witch died bravely for Mulan, and their feelings got the highest sublimation.



Figure 13. High grades



Figure 14. Medium grades

"Modality" is related to the authenticity of images, which can be divided into three categories: high, medium and low. Color saturation, contextualization, depth, illumination and brightness are the indicators of modality classification. The author will select the effective index "contextualization" for analysis according to the actual situation and the characteristics of the film. Based on the observation of the research samples, the author compares the correspondence between naked eye and visual reproduction, and divides "contextualization" into high and medium grades. Taking Figure 13 as an example, the background of the picture is

clear and distinguishable, and can accurately reflect people's environment. This kind of image tends to be consistent with reality. The image shown has a feeling of immersing the viewer, which is classified as "high contextualization". Although the background is blurred, its environment can be roughly distinguished according to the general outline of the background. This kind of image is divided into "medium contextualization", as shown in Figure 14.



Figure 15. Face-to-face relationship



Figure 16. Strong position



Figure 17. Head-up lens

"Attitude" is to express the meaning of "intervention" and "power" through five "Perspectives" in the horizontal and vertical directions. In the horizontal direction, it is mainly divided into "front angle" and "oblique angle" to judge the degree of "intervention" image. "Positive" means to watch the participants in the image from the perspective of intervention, forming a sense of involvement that "the viewer is placed in the image scene". The "slant" reflects the detached relationship between the viewer and the image participants, and conveys a spectator attitude of "I" rather than "you". For example, in Figure 15, the Witch and the viewer form a face-to-face relationship, while in Figure 16, the viewer can only see Mulan's side face, which makes it difficult to enter her world. In the vertical direction, it can be divided into "looking up", "looking up" and "looking down" to judge the attitude of image creators or viewers towards image participants. In Figure 16, the witch is in a high position, overlooking the camera, which makes the image viewer in a weak position. At this time, the image created by the witch is strong or respected. In Figure 6, Mulan holding the Witch and looking up at the camera shows that the image viewer is in a strong position. In Figure 17, Mulan, head up lens, reflects the equal relationship between the viewer and the image participants.

3.3. Compositional Meaning



Figure 18. The up-down composition



Figure 19. The left-right composition

"Information value" means that image elements realize their respective roles in the whole picture through their positions in the whole picture composition. The positions of information values form different reading paths: "from top to bottom", "from left to right", "from center to edge". Figure 18 is the up-down composition, the information at the top is general and ideal, while the information at the bottom is real or specific details. From the upper part of the picture, we can't see what Mulan is doing, but we can get the extensive information that she is bowing. From the gesture of Mulan in the lower part, we can see that Mulan is bowing to the emperor,

so we can infer that she is squatting at this time. In Figure 19, the information on the left side is new and the information on the right side is given. From the right part, we can clearly see the witch's appearance, but the fuzzy and unknown content on the left occupies a large part. This is the left-right composition. Through this composition, it shows the arrival of the witch. And Figure 6 is the center edge composition. In Figure 6, the information in the center, Mulan holding the witch, is important and core, while the space in the margin is subordinate to the core information. When reading the image, the viewer first captures the information point of the central protagonist, and then goes to the surrounding title information. From this composition, the audience can be attracted by the things in the most central position, so as to achieve the effect of immersive and thought-provoking.



Figure 20. From the perspective of witches

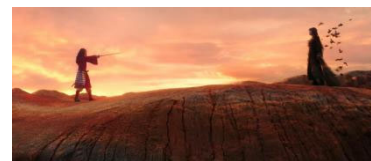


Figure 21. Separation

"Information value" examines the impact of different locations on information value, but there is also a principle of highlighting important information in different forms in images, that is, the "saliency" principle. The marks of "saliency": relative size, sharpness of focus, tonal contrast, color contrast, visual position, perspective angle and special cultural symbols. According to the characteristics of the film, the author uses the important index of visual position to analyze. From Figure 20, we can see Mulan from the perspective of witches. In this perspective, even if the witch repeatedly asked, Mulan was still indomitable. We seem to be in the world of witches, to observe the small and brave woman in the distance. This kind of visual position highlights the implied meaning that the creator wants to convey.

"Viewfinder" refers to whether the elements in an image have lines or frames to separate or connect them. There are six clear classification frameworks: aggregation, separation, integration, overlap, rhyme, contrast. In terms of operation, the first two frames are selected according to the characteristics of the film. Figure 21 is separation. Figure 5 is integration. "Separation" means that the elements are separated by a certain space rather than a dividing line, as shown in Figure 21. There is a certain area between the Witch and Mulan, which separates the two elements. The two wrestling together, forced apart, and produced a long distance, let the viewer feel as if they are separated from each other. "Integration" means that different elements occupy the same space, as shown in Figure 5, Mulan and Phoenix have nothing to do with the two elements are combined, emphasizing the integration of the two, making the viewer have association, guess the relevance of the two.

"Information value" means that image elements realize their respective roles in the whole picture through their positions in the whole picture composition. The positions of information values form different reading paths: "from top to bottom", "from left to right", "from center to edge". Figure 18 is the up-down composition, the information at the top is general and ideal, while the information at the bottom is real or specific details. From the upper part of the picture, we can't see what Mulan is doing, but we can get the extensive information that she is bowing. From the gesture of Mulan in the lower part, we can see that Mulan is bowing to the emperor, so we can infer that she is squatting at this time. In Figure 19, the information on the left side is new and the information on the right side is given. From the right part, we can clearly see the witch's appearance, but the fuzzy and unknown content on the left occupies a large part. This

is the left-right composition. Through this composition, it shows the arrival of the witch. And Figure 6 is the center edge composition. In Figure 6, the information in the center, Mulan holding the witch, is important and core, while the space in the margin is subordinate to the core information. When reading the image, the viewer first captures the information point of the central protagonist, and then goes to the surrounding title information. From this composition, the audience can be attracted by the things in the most central position, so as to achieve the effect of immersive and thought-provoking.

4. Conclusion

The Disney's film *Mulan* is to discover the spiritual characteristics of "surpassing oneself", and to narrate and adapt the story around it. This paper analyzes the twenty lines and twenty-one pictures of the three encounters of the Witch and Mulan in the film from the perspective of image in visual grammar and language in functional grammar, and discusses how the implied meaning of the film is so manufactured and expressed.

References

- [1] Kress, G, and Van Leeuwen T. *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold, 2001.
- [2] Halliday, M. A. K. *An introduction to functional grammar*. London: Oxford University Press, 2000.
- [3] Halliday, M. A. K, and Matthiessen, C. M. I. M. "An introduction to functional grammar." *Language*, 65(4), 862, (1994).
- [4] Kress, G, and Van Leeuwen T. *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold, 2001.