Multimodal Discourse Analysis of News Pictures

Mengyuan Bi

Shanxi Normal University, China

Abstract—**Multimodal discourse analysis is the analysis of different symbolic modes within a text, which breaks through many limitations of traditional discourse analysis to a great extent. This paper takes the visual grammar of Kress and Leeuwen as the theoretical framework, which gives a good explanation of the reproducing meaning, interactive meaning and composition meaning of image discourse, which is also suitable for the analysis of news picture discourse. This paper expounds how other symbolic resources interact with each other, so as to construct a complete text with linguistic symbols, and then convey more social interactive meaning. The results show that visual grammar is feasible and operational in the analysis of multimodal news texts. The background and text of news discourse can be effectively supplemented and explained, and it is of great significance to improve readers' pictures' reading ability.**

Index Terms—**multimodal discourse analysis, visual grammar, news pictures**

I. RESEARCH STATUS OF MULTIMODAL DISCOURSE ANALYSIS

With the emergence and development of multimedia technology, great changes have taken place in the way of human information exchange. In the 1990s, "multimodality" became a widely studied term in the field of semiotics and linguistics in the West. Scholars have come to realize that although language still plays an irreplaceable role in the process of meaning construction, however, in the field of discourse, it is only one of the symbolic systems. It is necessary to combine the unimodal study of language in discourse linguistics with other fields such as vision and hearing (Li, 2003). The emergence of the word "multimodality" emphasizes the importance of symbolic resources other than language. For example, the combination of language and sound, Kress and Van Leeuwen think, the significance of multimodal research is that it clearly points out that the interest in the study of modes other than words and speech is inevitable rather than accidental, it is the center of the form of realistic communication, not just as a supplement to language research (Kress,2006). The hypothetical basis of multimodal research is that almost all practices can be regarded as multimodality, if the original "multimodality" proposal was only out of interest in the study of other symbolic modes, so now it has become a paradigm that can provide interpretation for most texts.

The main theoretical basis of multimodal discourse is system-functional linguistics founded by Halliday, a British linguist. Specifically, on the basis of systemic functional linguistics, O'Toole, Kress and Van Leeuwen, regard the meaning potential of text, sound and picture as a series of interrelated systems and structures, and analyze them, thus laying the foundation of multimodal research. O'Toole proposed a "bottom-up" grammatical approach. He applied the system function mode of Halliday to the symbolic analysis of display art, sculpture and painting; Kress and Van Leeuwen focus on image and visual design, and summarize the general principles of visual design through a bottom-up contextual analysis method.

In the field of discourse analysis, multimodal discourse analysis provides a new vision and has been active in western academic circles for more than ten years. At present, the more famous scholars in the world are divided into social semiotics schools, interactive sociology schools and cognitive schools. The representative figures of the school of social symbolism are Kress and Leeuwen, who first put forward and perfect the theory of social semiotics analysis of multimodal discourse. This theory is based on Halliday's functional linguistics theory. The functional grammar put forward by Halliday summarizes the function of language into three metafunctions. The first is conceptual function, which is used to convey conceptual meaning; The second is the interpersonal function, which is used to reflect the relationship between the communicator and the receiver, and the attitude of the communicator to the content of the communication; the third is the textual function, which is used to express the meaning of the text (Halliday,1994). Meaning exists in all kinds of social symbol resource systems, such as vision, hearing, action, expression and so on. This kind of existence is both extensive and profound. Therefore, the understanding of social meaning can not only focus on language and character symbols, but must be combined with the examination of all symbol resources. No single symbol can be completely understood in isolation. Therefore, Kress and Leeuwen applied the three metafunctions of language in functional grammar to the understanding of non-linguistic symbols, they see language and non-verbal symbols as separate symbol resources, but play a role in each other, the role of linguistic symbols and other different modes of symbols is comprehensively investigated in discourse analysis.

II. VISUAL GRAMMAR

As mentioned earlier, Kress and Leeuwen believe that image is also a kind of social symbol, which can be analyzed and described by Halliday's social semiotics theory. According to the three metafunctions of the language, the two men

create the visual grammar of the analytical image with the representational meaning, the interactive meaning and the composition meaning as the core content, which provides the theoretical basis and analysis method for the multimodal discourse analysis.

A. Representational Meaning

The meaning of representation can be understood in two categories: narrative representation and concept representation. Vector is the symbol of narrative representation image, and there is no vector in the conceptual reproduced image. Vector is a quantity of direction and size, which is usually represented as a strong diagonal line in the image. Narrative representation can be divided into three categories: action narrative, reactive narrative, speech and psychological narration. In the narrative of action, the representation participants send out vectors, and they are often the most prominent participants. When the vector in the image is only the eyes of the participants, the image belongs to the reactive narrative image, and discourse analysis focuses on the "responder" and "phenomenon". Speech and psychological narrative images usually use a circle to frame what the participants think or say. The image of concept representation generally represents the relationship between participants, including their categories and characteristics. Concept representation and narrative representation are mainly different from time factors in the representation of image meaning. Narrative images focus on reflecting the process of change and development of an event or thing, while concept images focus on representation the category, structure, meaning and so on of participants.

Concept representation image includes three different contents: concept classification, concept analysis and concept symbol. There is always one part of the participant in the conceptual classification image that is subordinate to the other part of the participant, which can be either dominant or implicit. The dominant classification will appear at the same time as the upper level of the category and the subordinate category. However, there are only subordinate categories in the implicit classification, and the symmetric composition is generally formed between the participants, and the distance and size between them are almost the same. The relationships between participants in the conceptual analysis are generally partial and overall relationship. The whole is the bearing, and the part is the characteristic of any number of people.

B. Interactive Meaning

Compared with representation meaning, the realization of interactive meaning needs to mobilize more viewers' ideology. The interactive meaning embodies the tripartite character relationship: the producer of the image, the representation participant in the image and the viewer of the image. At the same time, it guides the viewer's attitude towards the representation of participants. The realization of interactive meaning of image mainly depends on four basic elements: contact, distance, attitude and modality. Contact refers to whether there is a docking between the participants in the image and the eyes of the image viewer, and whether it causes some kind of emotional experience of the image viewer. If there is eye docking, the image belongs to "request" interaction, if there is no eye docking, "provide" interaction. Participants in the image can express what they ask through expressions, eyes and postures. For example, if you look down from a high place, you can ask for obedience; from a bottom look up, you can ask for mercy; when you look directly at the viewer with a smile, you can get approval. On the contrary, if the image viewer does not have virtual contact with the participants, he will "stay out of the way" to treat the people and things in the image.

Distance is social distance, which is the distance between social beings in space, time and psychology, which symbolizes the closeness of people's social relations. Distance is social distance, which is the distance between social beings in space, time and psychology, which symbolizes the closeness of people's social relations.

Modality is the truthfulness or credibility of the statements which we make to the world we pay attention to. According to different standards, we can divide modality into high modality, medium modality and low modality. In color saturation, from high sensory modality to low sensory modality, the image will use high saturation color, not too saturated color, only black and white. They also range from eight visual markers to discuss the realistic value of modality in the image , such as color discrimination, color harmony, situational (from no background to extremely meticulous background), reproduction (from maximum abstraction to maximum reproduction of detail), depth (from no depth to maximum perspective), lighting (from the maximum reproduction of light and shadow to the absence of light and shadow) and brightness (that is, from the maximum number of different brightness to only light and dark)

C. The Meaning of Composition

The meaning of composition is realized in three ways: information value, scene selection and significance. The so-called information value is the position where a symbol is placed in the image. The position of the symbol also affects the embodiment of meaning, for example, the position on the left and right forms the structure of the known information-the new information, and the upper and lower positions form the structure of the ideal information-the real information. Ideal information indicates the idealization or generality of information, true information indicates certain specific information, or more "down-to-earth" information. Significance refers to the intensity of one symbol in the image to attract the attention of the viewer relative to other symbols. The symbol is placed in the foreground or background, the size of the symbol, the contrast of colors, the intensity of light, and so on. Whether or not there is a way to pick up a scene, through which some elements can be classified or excluded from others, thus reflecting whether they belong or not in a sense. Kress and Leeuwen believe that the principles of information value, significance and scene

selection are applicable to both image and text, and try to use the same terms in multimodal discourse analysis.

D. The Relationship between Different Discourse Modes

The overall meaning of the communicator's discourse is embodied by the discourse of different modes. When the discourse of one mode can not fully express its meaning, the discourse of the other mode is supplemented and constitutes a complementary relationship. Complementary relationship can be specifically strengthened relationship and non-strengthened relationship. Reinforcement relationship means that there is one mode is the main form of communication, and the other modes are only the enhancement of it. There is no primary and secondary in the unstrengthened relationship, and all kinds of discourse modes can not be missing, which complement each other and embody the complete meaning. There are coordination, joint and cross relations in the unstrengthened relationship. Coordinated relationship refers to the lack of any mode of discourse can not become a complete meaning; joint relationship refers to the combination of different types of media to reflect the meaning; cross is the cross embodiment of meaning. Compared with the complementary relationship, there is no mode between the modes of the non-complementarity relationship. This relationship can be overlapping, that is, although the discourse of all kinds of modes appears at the same time, it does not strengthen each other, or it can be included, that is, all kinds of modes have no new contribution to the expression of the whole meaning, but they provide more specific information, or they may not participate in the context in a positive way.

III. VISUAL GRAMMATICAL ANALYSIS OF NEWS PICTURES

A. Representation Meaning

The meaning of representation can be understood in two categories: narrative representation and concept representation. Vector is the symbol of narrative representation image, and there is no vector in the conceptual reproduced image. The background information in Fig 1 is that Putin congratulates President Xi Jinping face to face on his 66th birthday. There is an obvious vector in the picture. General Secretary Xi's eyes meet with President Putin's eyes, and there is also an obvious vector in the direction of the hand, which reflects the process of action and response. At this time, the photo shows the reproducing meaning of the narrative: Embodying the friendly relations between China and Russia through speech and action acts. The vector in Fig 3 exists between the participant and the viewer and belongs to the reaction narrative, attention needs to be paid to "responders" and phenomena, the background of the picture is that 70% of child labourers are engaged in agricultural work. The viewer directly creates sympathy for child labour through the picture and deeply thinks about the ill phenomenon of using child labour. There are also obvious vectors in Fig 4, in the middle of the picture, the participants and viewers produce obvious vectors, and the eyes of the two people in the rear and the map of green action produce vectors, so this picture is a narrative with behavior and reaction narrative.

Conceptual symbols are used to express what the participants are or represent. There are two categories of conceptual symbols: attributes and cues. Symbolism cues means that a certain meaning and identity are embodied in the qualities of the participants themselves. The symbolic attribute, on the other hand, is the external addition of this meaning to the participants. For example, to enhance the significance of participants in the image, can be placed in the foreground, using prominent dimensions, light, details, hue, saturation, etc, or point to it in a position, such as arrows, fingers, etc. There is no sense of vector in the eyes and movement between the characters in the Fig 2, their relationship is defined by the classification process: the classification process associates the participants with the classification relationship and at least one group of participants will play the role of the other participant in the "dependent" role (Li, 2003). Compared with the speaker on the podium, the audience is subordinate to the attribute role of the reporter, the meaning of reproduction here is not narrative, but conceptual reproduction: an official press conference is under way.

B. Interactive Meaning

The interactive meaning embodies the tripartite character relationship: the producer of the image, the representation participant in the image and the viewer of the image. The realization of interactive meaning of image mainly depends on four basic elements: contact, distance, attitude and modality. Contact refers to whether there is a docking between the participants in the image and the eyes of the image viewer, and whether it causes some kind of emotional experience of

the image viewer. If there is eye docking, the image belongs to "request" interaction, if there is no eye docking, "provide" interaction. Participants in the image can express what they ask through expressions, eyes and postures. For example, if you look down from a high place, you can ask for obedience; from a bottom look up, you can ask for mercy; when you look directly at the viewer with a smile, you can get approval. On the contrary, if the image viewer does not have virtual contact with the participants, he will "stay out of the way" to treat the people and things in the image. "Social distance" refers to the distance between the participant and the interactive participant, that is, the distance between the participant and the viewer in the graph. If the picture producer wants to present the picture content better, it must grasp the distance between the visual image and the reader. This distance is called "social distance". Its main yardsticks are close-up, medium vision, big vision. In Kress & VanLeeuwen's view, the close-up lens represents a kind of intimate private space, which can make the viewer feel his inner world through the appearance of the characters. "close-up" is a kind of personal friendly close-up. "Middle close-up" belongs to a long distance private space, but the viewer can still integrate himself into the scene in which the characters in the picture live and communicate with each other. "Middle vision" means a familiar social relationship; "Middle-big Vision" is a close social distance, in which there is still a sense of familiarity with each other; "Vision" means that the social distance is becoming more and more distant; in the "Big Vision", publicity and non-individualization are becoming more and more obvious. Kress summed up the relationship between viewing and distance, showing only the head and shoulders to indicate personal close range; The image above the waist symbolizes the distance of the individual; the whole person appears in the image to indicate the close distance of the society; in addition to the whole character, there is a blank space around it that indicates the distance of the society. The distance of at least 4 to 5 people means public distance. The attitude that the image is provided to the viewer can be achieved from a perspective. For example, when the reader looks at the image positively, it will produce the feeling of empathy and integration, while the inclined viewing image will feel indifferent and distance, the things in the overlooking image symbolize the power, and the things in the flat-looking image embody equality.

Fig 1 is just an objective display of some kind of information to the viewer, the line of sight between the image participant and the viewer does not intersect and does not interact, at this point, it's just "providing" information about participating in an event. The image is taken as the middle scene and does not interact with the viewer in the eyes or limbs, there is a more estranged relationship between the image viewer and the participants, that is, the perspective and distance of the "other". The participant in Fig 3 is connected to the viewer's line of sight, resulting in a "request", the position of child labor in the picture is an overlooking perspective for viewers, thus arousing sympathy and pity from viewers. Fig 4 publicity map for green action, the line of sight of the boy riding the bike in the middle docked with the viewer, resulting in "request", the shooting angle of the picture is horizontal, so the picture is "request" for approval, and eager for readers to join the green action, which gives the reader a sense of presence. The distance can be connected to the relationship with the viewer by the proportion of the image participants. If the proportion of the participants' face and head is large in the image, a closer relationship can be constructed; if you want to build a public relationship, you have to show the whole body of the participants, and you should be surrounded by space until you can see the bodies of four or five people. Fig 1 and Fig 2 are the distance of public relations, which reflects the psychological distance between the viewer and the photo.

C. Composition Meaning

The meaning of composition corresponds to the textual function of systemic functional grammar. It contains three elements: information value, significance and scene frame. The information value refers to the order in which each component of the composition is arranged. In photo discourse, each component in the overall composition will vary according to its position, and it reflects the different information value. The composition of the content from left to right can establish the structure of the known information --- the new information. The composition of the content from top to bottom can convey the structure of ideal --- reality, while from the center to the four sides, it can express the important --- secondary structure. The center of Fig 1 is the glass of the two leaders. General Secretary Xi Jinping's glass is slightly lower than that of Putin, which show the quality of General Secretary Xi Jinping's respect for other countries and modesty and courtesy. The center position of the two glasses can also imply the fine quality of humility and courtesy of the Chinese nation. The composition of the photo is more coordinated and does not magnify or shrink a participant in proportion. This makes all the image participants more harmonious integration. In figure 4, the center is a boy wearing green short sleeves and riding a bike. Green is an environmentally friendly color that gives readers a clear background meaning. The information at the top of the graph is "亲近自然，玩转绿色行动". The sentence can direct and clear express the subject of this activity, and the composition of this sentence is an idealized message. On the left and right sides of the picture are two young runners and two young people looking at the map. On the left is known information, and people know how to express green action. On the right is a new message. It tells viewers that there is a prescribed route for this activity. The use of young people as participants in the picture gives the viewer a feeling of spiritual fullness and strengthens the desire to join the event. The background of the picture is urban architecture which suggests that the environmental protection activities are carried out in the city.

IV. CONCLUSION

News photos are a common expression in new media discourses, and it is of great significance to analyze and study them. Visual grammar analyzes the three functional meanings of photo discourse from the perspective of social semiotics, pays attention to the internal representation factors and its structure of discourse, and has a good explanatory power to news photo discourse. News discourse has been paid more and more attention because of its wide audience, great influence and strong effectiveness. This paper expounds how other symbolic resources interact with each other, so as to construct a complete text with linguistic symbols, and then convey more social interactive meaning. The results show that visual grammar is feasible and operational in the analysis of multimodal news texts. The background and text of news discourse can be effectively supplemented and explained, and it is of great significance to improve readers' pictures' reading ability.

REFERENCES

[1] Halliday, M. A. K. (1978). Language as Social Semiotic: The Social Interpretation of Language and Meaning. London: Edward Arnold.
[2] Halliday, M. A. K. (1994). An Introduction to Functional Grammar (2nd edition). London: Edward Arnold.
[3] Hu Zhuanglin, Zhu Yongsheng, Zhang Delu, Li Zhanzi. (2017). Systemic Functional Linguistics. Beijing: Peking University Press.
[4] Kress, G. & Leeuwen, Theo Van. (2006). Reading Images: The Grammar of Visual Design. London: Routledge.
[5] Li Zhanzi. (2003). Social semiotics analysis of multimode discourse. *Foreign language Studies, 5,* 1-8.

Mengyuan Bi was born in Linfen, China in 1995. She is a sophomore of graduate student in Shanxi normal University, majoring in foreign linguistics and applied linguistics.

She majors in foreign linguistics and applied linguistics. Her research interests include multimodal discourse analysis and pragmatics.

Ms. Bi has published two papers about linguistics which can be found in the CNKI.