Visual Narrative: Multimodal Discourse Construction on Sustainable Development Video for Beijing Winter Olympics

Ting Ding[1], Wenzhong Zhu[2*]

[1]School of Foreign Studies, Nanjing Institute of Technology, Jiangsu, China

[2]School of Business, Guangdong University of Foreign Studies, Guangzhou, and Business School，City University of Macau, China

[*]Corresponding author, E-mail:wenzhong6409@163.com

Abstract

Sustainable development (SD) has become a major concern in recent decades and sports mega-events (SMEs) such as the Olympic Games are considered an important way for host countries to convey their concepts, policies, and measures of SD. This paper takes the official video of the 2022 Beijing Winter Olympics (BWO) named *Sustainability for the Future* as the corpus and intends to analyze the multimodal constructions of the SD concept of the BWO from a Meta-functional approach, including interpersonal meaning, ideational meaning, and layout, based on the framework of the Visual Narrative Theory (VNT). The results show that four types of discourse are applied to this video, succeeding in constructing a green, sharing, open, and honest SD concept of the BWO and shaping the image of China as an environmental-friendly and responsible host country. Furthermore, characteristics of the construction of the SD discourse of SMEs are fully discussed in this paper. This research is a profound application of the multimodal discourse(MD) analytic approach to the SD discourse from the perspective of the VNT and it also provides valuable reference and inspiration for host countries of SMEs in the future to improve their official communication level and convey the SD concepts to the outside world in the cyber world.

Keywords: visual narrative theory; sustainable development; multimodal discourse; Beijing Winter Olympics

1 Introduction

Sustainable development (SD) is one of the three basic principles of the Olympic Agenda 2020 [1], and it is also one of the main concepts of Beijing's bid to host the 2022 Winter Olympics. The definition of SD was first proposed in the report *Our Common Future* in 1987 by the World Commission on Environment and Development. It is defined as "development that meets the needs of the present without compromising the ability of future generations to meet their needs" [2]. In 1992, the United Nations (UN) issued the Rio de Janeiro Declaration on Environment and Development and Agenda 21 [3]. Since then, the SD concept has gradually been accepted by most countries in the world. In the year 2000, the UN first proposed the Millennium Development Goals (MDGs) [4] and released the more comprehensive Sustainable Development Goals 2015-2030 (SDG) in 2015 [5]. From then on, the SD concept has become a global consensus.

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

However, due to the huge differences in development levels among countries, there have been diverse ways of understanding and practicing the concept of the SD despite the global consensus on the SD. In terms of the sustainable discourse construction, previous researchers tend to hold a variety of interpretations of "sustainability", and they have not made any convincing and innovative explanations about the relationship among the three dimensions (ecology, economy, and society) of the SD [6]. As for the practice aspect, current research on sustainable discourse construction mostly focuses on the environmental and economic field, but far too little attention has been paid to the studies in sports, especially, the mega ones. In addition, although the transmitting channels of sports event publicity have changed from traditional papers or radios to currently multimodal networks with the wide application of cyber media, few studies have been conducted on the multi-modal discourse constructions under the background of the internet [7]. To fulfill the research gap, this paper takes the official video of the BWO named *Sustainability for the Future* as the corpus and intends to achieve the following primary goals: (1)analyze the multimodal construction methods of the SD concept of the BWO from three aspects, namely interpersonal meaning, conceptual meaning, and version layout respectively; (2) summarize the characteristics of SD discourse construction in mega sports events; (3)explore how to exert the characteristics of multimodality of network media discourse to convey the concept of sustainable development more precisely and vividly so as to create a positive image of the host country.

2 Literature Review

2.1. Multimodal discourse analysis (MDA)

MD refers to a compound discourse that includes gestures, expressions, images, sounds, etc. which incorporates traditional linguistic signs and non-linguistic signs into a unified category. MDA theory was developed based on Halliday's (1978) assertion that "Language as social semiotic" in *Language as a Social Semiotic: Towards a Comprehensive Sociolinguistic Theory* [8]. Since the 1990s, MDA has gradually become one of the most significant discussions in linguistics and communication studies as the multimodal trends of functional linguistics and discourse analysis. The representatives of early research on MDA include O'Toole (1994), Kress & van Leeuwen (1996/2006), O'Halloran (2004), and so forth [9-12]. Among them, Kress & van Leeuwen argue that the three meta-functions of language proposed by Halliday (1985) [13] exist in all human communication systems, and MD should not be the exception. Therefore, they studied the realization process of language meta-function in MD and divided the meaning of MD into three intertwined types: representation, interaction, and composition. In their foundational work *Reading Images: The Grammar of Visual Design*, Kress & van Leeuwen put forward the significant theory of Visual Grammar (VG) [10-11], which exerts strongly positive impacts on MD studies, making MDA attract highly continuous attention of scholars and become an indispensable part of discourse analysis research.

Research on MDA in China began after 2000. In 2003, *Social Semiotic Analysis of MD* published by Zhanzi Li introduced the main content of *Reading Images: The Grammar of Visual Design* (Kress & van Leeuwen, 1996) to Chinese readers for the first time, and established the social semiotics framework of image analysis, which is regarded as the foundation for MDA in China [14]. Yongsheng Zhu (2007) introduced the generation, definition, nature, and theoretical basis of MD in The Theoretical Basis and Research Methods of MD Analysis, enriching the theoretical framework of MD [15]. Delu Zhang's (2009) Exploration of a Comprehensive Theoretical Framework for MDA discussed it from different points of culture, context, meaning, form and media while also exploring the relationship between the various modalities [16]. In addition to theoretical research, a number of domestic scholars have also engaged in the introduction of MDA tools and the construction of multimodal corpora. For example, Lifei Wang and Yan Wen (2008) introduced the use of the multilayer annotation software ELAN [17]; Xuening Li and Lixia Ma (2010) explained the application of the MAC multimedia information retrieval tool [18]; Lihe Huang (2015) dis-

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

cussed the construction of a multimodal corpus of impromptu discourse [19]. Furthermore, there are also quantities of extensive and in-depth discussions on MD forms such as advertising discourse, news reports, movies, classroom teaching, and natural conversation from the perspectives of systemic functional VG, multimodal metaphor, and multiple reading and writing [20-22].

While the VG theory established by Kress and van Leeuwen (2006) lays a solid foundation for systematically explaining the meaning construction of graphic texts, Painter et al. (2013) pointed out that the current tools of VG are "insufficiently developed for ad-dressing key aspects of picture books", especially in terms of the relations among images in a sequence [23-24].

2.2. Visual narrative theory(VNT)

VNT is considered the most significant development in visual image analysis theory since the foundational work of Kress & van Leeuwen (1996/2006). It was proposed by Clare Painter, Jim Martin, and Len Unsworth in their masterpiece *Reading Visual Narratives: Image Analysis of Children's Picture Books* [23,25]. The contribution of VNT to VG can be summarized in two aspects: firstly, VNT expands research objects from a single image, such as advertising images, to more complex visual narratives composed of multiple images. Secondly, VNT improves the theoretical framework of VG in two ways: on the one hand, it abandons or re-interprets the controversial views, such as "the left side as the known information" and "the eye expression", etc.; on the other hand, it adds important frameworks of emotional representation, characterization, event relationships and so forth [25]. There have been a number of works applying the VNT to the analysis of cartoon books, TV, movies, etc. since the VNT was proposed in 2013. Recently, some studies focusing on the BWO based on the VNT approach have also begun to emerge. Unfortunately, these studies are merely restricted to the implications of the symbols and slogans of the BWO and there are few studies concerning how multimodal images and language work together to construct and convey the SD concept of the BWO. Therefore, analysis in detail of the official sustainability promotion video of the BWO will be fully examined by a meta-functional approach [13] from three aspects of interpersonal meaning, ideational meaning, and lay-out in Section 3.

3 Analysis of the multimodal constructions of the SD concept of the BWO

It is noticeable that the BWO is the first Olympic Games that practice the SD concept throughout the whole process. In the past six years, the Chinese government has fully fulfilled the sustainable commitments made during the bid and achieved a series of achievements that have attracted worldwide attention in the aspects of the environment, society, and people's livelihood. Accordingly, the corpus of this study totally comes from *Sustainability for the Future* released on the official website of the BWO on January 14, 2022. The video lasts only 6 minutes and18 seconds. It narrated a short story of a little boy called XIA Dongdong aged 6 years old about how he completed the homework of "building the ideal Winter Olympic Paradise in my mind" in a way that would not harm the earth. This study attempts to interpret how the BWO conveys the concept of SD to the audience through multimodal expressions based on the VNT and MDA.

3.1 The interpersonal meaning perspective

In the 'Enacting Social Relations', Painter et al. examine the visual interpersonal meaning, introducing the systems of FOCALIZATION, PATHOS, and AFFECT, which are presented as complementary to Kress and van Leeuwen's (2006) SOCIAL DISTANCE, INVOLVEMENT, ORIENTATION, and POWER [23].

3.1.1 FOCALIZATION

FOCALIZATION can be seen as an extension of Kress and van Leeuwen's CONTACT system [10-11], with more focus on the "affectual dimension of meaning". "gaze" is seen as an important parameter to measure the relationship between the viewer and the depicted characters, and they suggest that facial and bodily

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

postures and stances can be seen as attitudinal resources to realize interpersonal meanings. According to Figure 1, FOCALIZATION consists of two aspects: one is whether the characters in the image participate in the interaction with the reader; the other is what kind of reading perspective the image provides the reader. Painter et al. (2013) distinguish two types of images of Contact and Observe according to gaze or not [23]. As far as perspective is concerned, Painter distinguishes between mediated and unmediated images. The former refers to the reader reading through the perspective of the characters in the figure, and the latter refers to the reader's perspective that is inconsistent with the perspective of the characters in the picture. Since the meaning of the word "mediated" is confusing to some extent, we borrow the terms of subjective shot and objective shot in film studies in order to facilitate understanding: "unmediated" as the objective perspective, and "mediated" as the subjective perspective [25]. Generally, the subjective perspective can integrate the reader into the story, making it empathetic, and thus more likely to arouse the reader's emotional resonance.

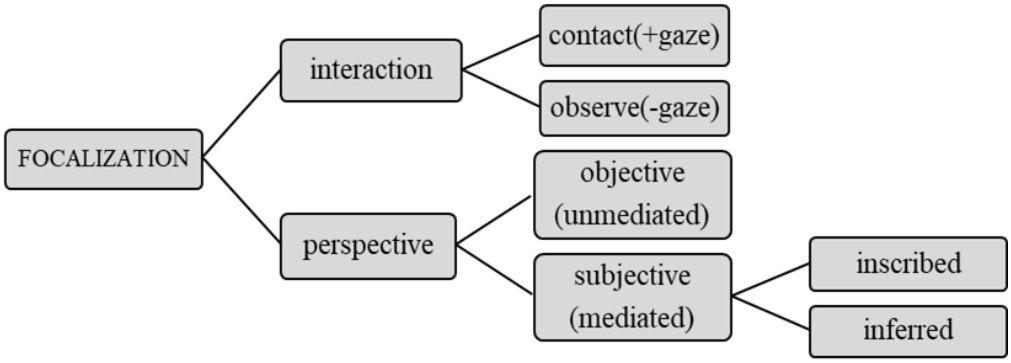


Figure 1. FOCALIZATION in interpersonal meaning

As shown in the video, the interaction is established by eye contact with the audience at the very beginning. *Sustainability for the Future* begins with a scene of a running young boy and his monologue: "My name is XIA Dongdong. I love sports very much. My parents said this may be related to my birthday on which there was a piece of good news spread all over Beijing." At the same time, when making eye contact with the little boy, the audience can easily notice the prints on his T-shirt read "Small Things with Great Love" which implies the SD concept of the BWO starts from small things and finally achieves big goals.

Another example of the interaction can be seen in the scene when IOC President Thomas Bach announced that Beijing won the right to host the 2022 Winter Olympics and Paralympic Winter Games. It is noticeable that there is a close-up on Bach's face when he uttered the decision, and his eyes are full of anticipation and trust, indicating that the IOC believes China has the ability to hold a successful Winter Olympics. It is mentioned that a complementary method is also used to combine the image with the text "China won the right to host the 2022 Winter Olympics and Paralympics on July 31, 2015" at the bottom right, which properly explains the little boy's name XIA Dongdong: his family name "Xia" means summer, while the given name "Dong" means winter in Chinese; he was born on July 31, 2015, the same day as the Beijing won the Winter Olympic hosting right; he is a child growing up together with the preparation for the BWO. According to the above shots, it is easy and natural for the audience to build interaction with the characters of the video, which would quickly get the audience engaged in the story and evoke their interest in the following storyline.

With regards to the perspective, it mainly is used in XIA Dongdong completing the handwork assigned by his teacher. Firstly, there is an example of the objective perspective in the shot. In this picture, the teacher was introducing the basic concepts of SD of the BWO to the children in an easy-to-understand way to make sure that students can understand what she said, "Sustainability means making the environment around us

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

more beautiful, making the city develop faster, and making everyone's life happier". Secondly, the video-maker applied a subjective perspective in order to show the specific SD contents of the BWO. As shown in another picture, a listing drawn by the litter boy is provided: on the left, there are the symbol of the BWO and the homework "building the ideal Winter Olympic Paradise in my mind" in Chinese; on the right are the five main measures taken by the Chinese government to achieve the SD goal in the BWO, which are "protecting animals and plants, making ice with CO2, double Olympic venues, low Carbon transportation, and green energy."

It can be seen from the above analysis that the subjective perspective is more likely to integrate the reader into the story. However, both objective and subjective perspectives are of equal importance in the VN and can convey main ideas effectively. On the one hand, by the application of the objective perspective, the audience could obtain the general contents of the theme, which is crucial to promote the story and the construction of the SD concept. On the other hand, with the help of the subjective perspective, the audiences can quickly enter the story and easily comprehend the SD concept of the BWO.

3.1.2 PATHOS

Another important subsystem of interpersonal meaning in VNT is PATHOS (shown in Figure 2). Painter et al. (2013) argue that the realistic and abstract styles of image representation are not just a matter of modality, but a discourse strategy for authors to set the emotional relationship between readers and image characters [23,25]. They first distinguish two representations of Engaging and Alienating: the former means that the reader has an emotional relationship with the characters in the picture, and the author tries to attract the reader's emotion; the latter means that the reader reads with an objective and critical attitude with no emotional input [25]. According to the degree of readers' emotion-al involvement, the intervention methods are further divided into Appreciate, Empathic, and Personalizing. Appreciate images are in the simplest style and abstract; empathic images are in generalized style, in which we can easily identify the emotions of the characters, but cannot recognize who they are; Personalizing images are in natural style, generally in the form of individual photos, which has authentically facial expressions. Using this framework, we can analyze the differences in the degree of emotional involvement of different types of visual narratives. Due to the short duration of the promotional video and the compact picture, it is usually necessary to increase the degree of Engaging to attract the audience's emotion and interest.

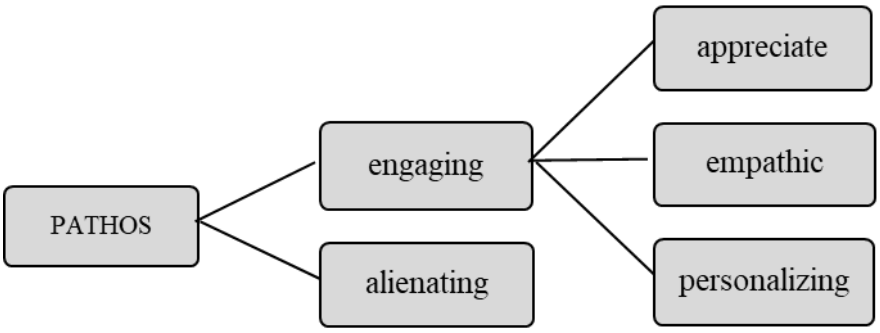


Figure 2. PATHOS in interpersonal meaning

The engaging of PATHOS in *Sustainability for the Future* is mainly reflected in the following two aspects. For one thing, as mentioned in the last part, the video chose a Beijing boy born on July 31, 2015, as the protagonist, which may quickly establish an emotional connection between the audience and the little boy since he is very cute, active, and growing up with the BWO. As a result, it can effectively convey the SD of the BWO to the audience. For another, it is interesting that a vast number of close-up individuals

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

smiling faces appeared in this video, especially when presenting the sustainable goal of fulfilling new urban development and happier people's lives. From a group of smiling shots of people from all walks of life in Beijing and Zhangjiakou in the video, the audience can share the happiness in these common people's faces, which indicated that a great deal of attention has been paid to the BWO in the preparation stage and various SD measures have been taken in the construction of the BWO venues. Therefore, the process of hosting the BWO does not have any negative effects on the local residents, instead, it improves the quality of life of local people, creates more employment opportunities, and promotes urban development.

3.1.3 AMBIENCE

While it is interesting that in the case of AFFECT, no specific system network is given, Painter et al. also acknowledged that color is an important meaning-making resource and one that carries ideational, interpersonal, and textual meaning (Kress and van Leeuwen, 2002; van Leeuwen, 2006) [10,11] in their introduction to AMBIENCE. As shown in Figure 3, Ambience refers to the emotional tone constructed by color in the VNT. The framework distinguishes whether the emotional atmosphere is activated by the use of color or not. Color images include three dimensions: vibrancy, warmth, and familiarity [23]. First, high vibrancy is the brightest color with the highest saturation, and low vibrancy means the darker color with less saturation. As far as atmosphere is concerned, bright colors create an exciting, energetic emotion; while dark colors create a somber, restrained emotion (Painter et al. 2013: 37). Secondly, the warm colors mainly refer to red and yellow, while cool colors refer to blue and green, which construct a warm and comfortable atmosphere or a cold and alienated atmosphere respectively. Third, the familiarity mainly refers to the richness of colors. The richer the color, the more it can reflect its natural color; on the contrary, a single color appears more abstract. In terms of emotional atmosphere, natural colors give people a feeling of familiarity and closeness, which can shorten the distance with readers. According to the above conclusion about colors, authors of promotional films tend to choose various colors to create different emotional atmospheres, so as to arouse the emotional resonance of the audience [23,25].

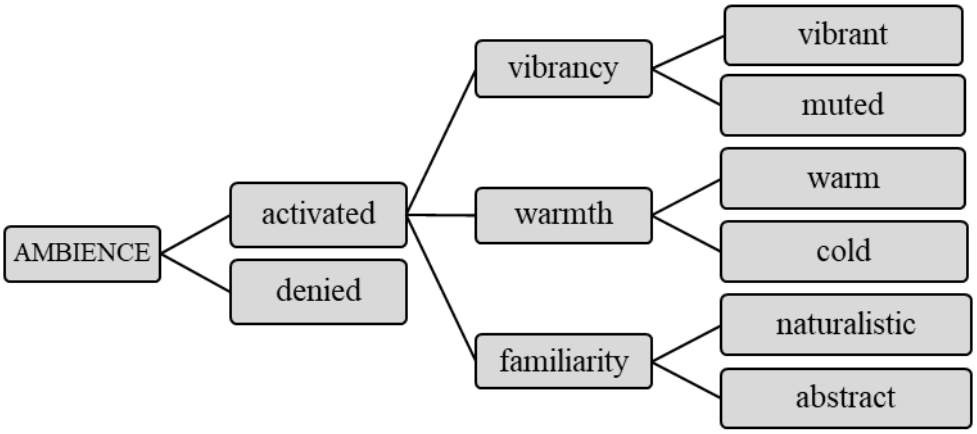


Figure 3 AMBIENCE in interpersonal meaning

As *Sustainability for the Future* is a video in color, its emotional atmosphere is activated. In terms of vibrancy, 95% of the images in the video use bright colors with high saturation, creating a vibrant atmosphere. For example, in the opening screen of the video, the blue sky accounts for nearly 80% of the figure, constructing an atmosphere full of vitality and hope, which implies that the BWO will be an Olympic that truly implements the concept of SD. It is noticeable that the six Chinese characters " 可持续 向未来 " in the center of the figure applies white hollow fonts, and what is more impressive is that the hollows of the characters are embellished with various images of the Winter Olympic events. With regard to warmth, blue, green, and white are classified as Cold by Painter et al. (2013) in the VNT. However, this video aims to pro-

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

mote the SD concept of the BWO, and it is known that blue and green are the typical colors of nature and white is the color of ice and snow. Therefore, the above three colors here did not make the audience feel alienated, instead, successfully conveyed the concept of SD during the preparations for the BWO, and at the same time pointed out the theme of ice and snow sports naturally. As for familiarity, the entire video is rich in color, which easy to create a feeling of familiarity and closeness between the depicted images and the audience and narrow the distance in a short time accordingly.

3.2 The ideational meaning perspective

Painter et al. (2013) extended the conceptual meaning framework of Kress & van Leeuwen (1996) to the following three aspects (demonstrated in Figure 4): firstly, they proposed a framework for character representation and the relationship between characters; secondly, they proposed an event relationship in VN; third, it proposes the contextual change relationship of different images [23]. In this way, the conceptual meaning framework of the VN covers the three elements of participants, process, and environmental components in systemic functional linguistics (SFL).

First of all, character representations can be divided into two categories: complete and metonymic. The former uses the facial features of the characters to clearly construct the character's identity; while the latter only uses the non-facial features such as clothing to represent the character's identity, forming a metonymic relationship [23,25]. In *Sustainability for the Future*, a large proportion of close-ups of the characters' faces are made. This kind of close-up eliminates unimportant elements such as lower limbs and highlights the emotional response of the characters in the picture. In this video, all the individuals are smiling dedicatedly, which indicates that they are satisfied with the work conducted by the government and benefit a lot from the process to prepare for the Winter Olympics. For example, the development of urban infrastructure and transportation provides convenience for the local residents; the increasing number of new job occupations are also improved their living quality, etc. which fully demonstrates the SD goal of the BWO of "Better people's life".

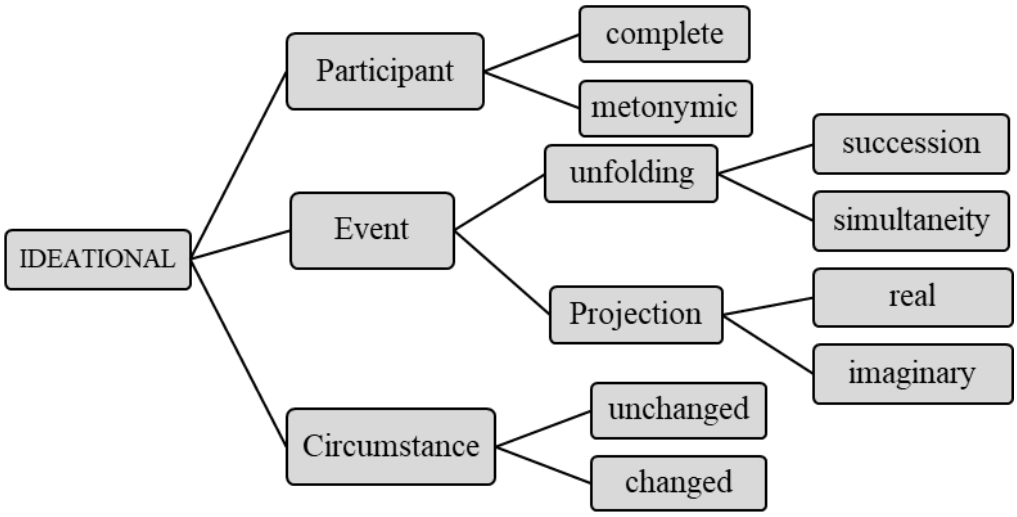


Figure 4 IDEATIONAL system in the VNT

In addition, the event relation system refers to the connections between different events, which are unfolding and projection. Unfolding means that the event development mainly focuses on the timing of two events: simultaneous or sequential. The former, as the name suggests, refers to the simultaneous process of multiple characters as the main body, while the latter obviously occurs sequentially [23,25]. In general, the

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

sequence is the more common way for events to unfold. Then, the event projection is similar to the process of projection in language. For example, in the real event projection, the event in one picture is the object seen by the characters in the other picture (visual projection constructed by behaviors); in the imaginary event projection, the event in one picture is what the character in the other picture had in mind (visual projection constructed by thinking). In the promotional video, the teacher taught the little protagonist several ways to "build the ideal Winter Olympic Paradise in your mind" in ways that would not harm the earth in class before he started. These sustainable methods are Animals and plant protection, Carbon dioxide ice-making, Double Olympic venues, Low-carbon transportation, and Green energy, which are exactly the main sustainable measures adopted in the construction and preparation of the BWO. Particularly, in terms of image representation, the video maker skillfully projected the process of building the Winter Olympics paradise of XIA Dongdong to the corresponding real venues of the Winter Olympics, resulting in a repeated switch between the children's manual work and the authentic scenes of the Winter Olympics. Consequently, it is easier for the audience to understand the originally highly complex and profound concepts of the SD of the BWO in the way of the narrative from the perspective of a six-year-old child, which is conducive to conveying the SD connotation of this Winter Olympics to the public.

Furthermore, Circumstance studies the continuity and change of the context of adjacent image events in the VNT. Specifically, if the context did not change over the course of the event, the author could either choose to use the exact same background or change the perspective (e.g. from elevation to depression, from long-range to close-range). If the context changes, the author can choose to apply a new background or remove the previous [23,25]. In *Sustainability for the Future*, the background of the picture has been replaced with the three major SD themes named "Positive environmental impacts", "New urban development" and "Better people's life" in order to highlight the diversity among them. From the point of view of shooting, most of the characters in this video are mainly shot from an elevation angle, showing the enthusiastic expectation of the Chinese for the Winter Olympics.

3.3 The layout perspective

Modifying the compositional meaning of Kress & van Leeuwen (1996) and combining it with the image-text relationship framework of Martinec & Salway (2005), Painter et al. (2013) proposed the layout relationship between images and text in visual narrative discourse (presented in Figure 5) [23].

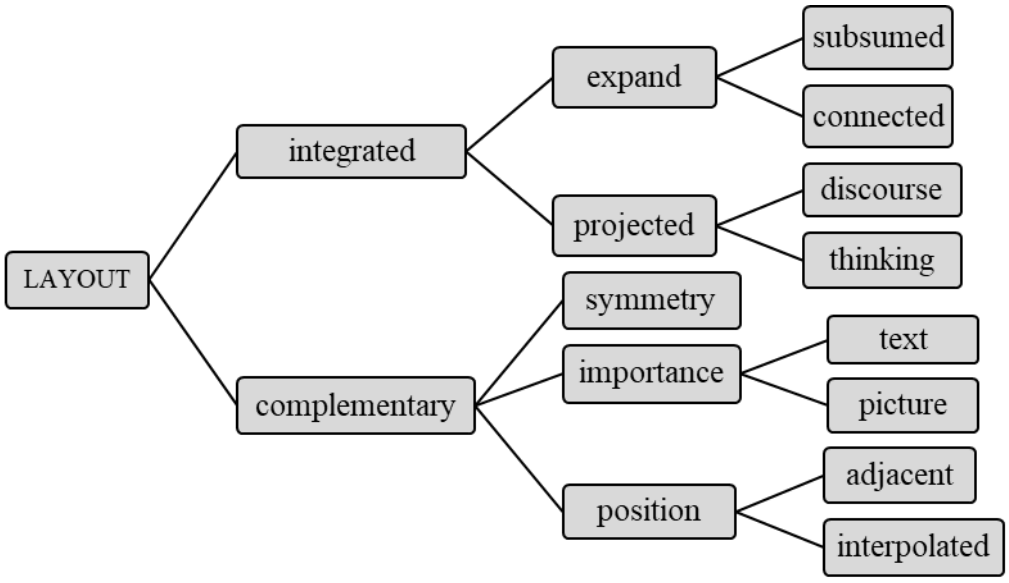


Figure 5 Layout system in the VNT

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

The layout relations can be divided into *integrated and complementary*. Firstly, in the Integrated relation, the language comprises part of the image, which may be further divided into an extension relation and a projective relation [23]. On the one hand, images and words have their own meanings in an extended relationship, but the two are interrelated, such as complementing, reinforcing, or interpreting each other. This relationship can usually be achieved by two layouts, which are Subsumed and Connected. Subsumed refers to the text overlaid on the image, and Connected is mainly achieved through vectors (the direction of the action). On the other hand, discourse projection is the most common picture-text relationship in the video, which is mainly realized through subtitles (discourse) and text description (thinking). The shot that the teacher assigning manual work to students is a typical Integrated example in the layout framework: the text on the blackboard behind the teacher forms part of the figure, and the Chinese characters " 北京 2022 冬奥会 "(Beijing 2022 Winter Olympics) are written on the blackboard, indicating the theme of the video. It is also noticeable that there are some other characters on the blackboard, for example, "保护动植物"(Animals and Plants Protection), "CO2 制冰 "(Carbon Dioxide Ice-making), " 双奥场馆 "(Double Olympic Venues), which are the specific measures of 2022 Winter Olympics. By conducting the measures above mentioned, the SD objects of the BWO, such as *Positive Environmental Impacts, New Urban Development, and Better People's Life* will be achieved eventually. In addition to the text, cartoon images of the mascots "Bing Dwen Dwen" (a panda in an icy coat suggesting purity, health, and ingenuity) and "Shuey Rhon Rhon" (a cartoon lantern representing inclusion and integration) of the BWO and Paralympics were also drawn on the blackboard, as well as a child skating. The above-mentioned texts and images are interconnected and complement each other, building an extended semantic relationship. Specifically, the teacher's hand in the figure can be regarded as a vector, therefore, this figure mainly expresses the Connection relationship. Through the layout shown in this picture, the integration of Chinese characters, words, and images has been created naturally, paving the way for the subsequent introduction of specific measures in detail for sustainable development.

Secondly, in the Complementary relationship, if the image and text occupy a part of the space respectively, we can examine the complementary relationship from the perspectives of Symmetry, Importance, and Position. Symmetry means that images and text are evenly distributed on both sides of the central axis; Importance refers to whether a page is dominated by images or text [26]; Position includes two cases: adjacent and separated. Adjacent means that although there is no vector connection between words and images, they establish a semantic relationship through close connection, while Separated means there is a border or other contents between the image and the text. As in this case, when introducing the specific sustainable practices of the BWO, the Complementary is widely used in this video, with more images than texts, to vividly highlight the SD objectives proposed by the Chinese government at the time of biding have been perfectly fulfilled. As examples demonstrated at 4 minute 15 second, there are three shots demonstrating the sustainable measures in natural environment, building, and human well-beings. First shot: "3,500 square meters of subalpine meadow back paved and restored" intuitively shows that the original vegetation was carefully stripped by the workers to restore the green land during the construction of the alpine ski resort, posing the leas at negative influence to the environment. Second shot: "Make full use of existing venues" uses animation to demonstrate the entire process of transforming the Water Cube of the 2008 Beijing Olympics into the Ice Cube of the Winter Olympics, which was told to audience through the mouth of the little protagonist: "after the Winter Olympics, this venue can play in both summer and winter, which will become a veritable Double Olympics venue." The renovation of the Double Olympics venues not only maximizes the use of existing resources but also brings benefits to the local people, perfectly illustrating the concept of sustainable development. Third shot: "To engage 300 million Chinese people in winter sports, and 1/5 of the residents of Chongli are engaged in winter industries" indicates the sports event have offered a large number of employment opportunities to local citizens, greatly stimulating local economic development, and increasing their income and living standards.

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

4 Discussion

After a detailed analysis of the sustainable discourse construction of the promotional video of the Beijing 2022 Winter Olympics based on the theoretical framework of visual narrative, the characteristics of the sustainable discourse of mega sports can be summarized as the follows:

One characteristic of sustainable discourse is that discourse narrative fosters emotional identity. The identification under the influence of media discourse is not mandatory, instead, it is realized in the form of acculturation through the interactive communication and guidance of emotion [27], especially the constructional role is more prominent in the network media era. For example, in the case of *Sustainability for the Future*, narratives of a boy aged 6 are used to arouse the audience's empathy so that the audience can quickly enter the story of the little protagonist, gradually build a common SD discourse, and reach a consensus on sustainable concepts.

Another characteristic is that MD strengthens emotional mobilization [27,28]. Different from the traditional propaganda reports dominated by text, the multiplicity of social media platforms provides an ideal medium for the integration of multimodal resources for mega events, and the application of video realizes the adjustment of communicative purposes of different discourse types. On the one hand, MD improves the intuitiveness, transparency, and authenticity of the in-formation. On the other hand, in the process of dissemination of SMEs, online media have a positive impact on the construction of SD discourse from the perspective of emotional mobilization via the interaction of text, pictures, videos, etc. In particular, the interaction between text and video is one of the important forms of network media communication, which is often manifested as using text to explain the connotation, to pave the way for the video, and then to further strengthen the significance of the SD concept in the construction of sustainable discourse in SMEs through video spreading on the Internet.

In addition, the discourse style is usually close to popular culture. It is acknowledge-able that SMEs can attract high public attention and wide participation, which reflects in the network is the public is playing a much more important role in the communication process and the spreading of information become wider and faster. Therefore, in the narrative of mainstream media, the discourse style should pay more attention to popularization and diversification [29]. Although SD has become a familiar concept to the majority of netizens, the concept of SD in sports events seems to be relatively abstract to many people [30]. For example, as introduced at the beginning of this paper, since the BWO is the first Olympic Games that implement the concept of SD throughout the whole process, the Chinese government has adhered to the SD vision of green, shared, open and clean Olympics from start to finish and has achieved a positive impact on the environment, regional development and life quality. However, how to effectively convey the information mentioned to the public? As a promotional video released on the official website of the BWO, aiming to convey the concept of SD, *Sustainability for the Future* selected a boy born on July 31, 2015 (the day Beijing won the hosting) as the protagonist, who demonstrates the concept of SD of the BWO in an easy-to-understand way. Besides, with the help of projection techniques, it is much easier for the public to perceive the measures carried out by the Chinese government. Finally, the more people know, the easier it is for them to engage themselves in sports events, which would eliminate misunderstanding due to cultural differences and help gain more support from the public all over the world.

5 Conclusion

This paper used the official video of the BWO *Sustainability for the Future* as the corpus and analysed the multimodal constructions of the SD concept of the BWO from a Meta-functional approach, including interpersonal meaning, ideational meaning, and layout, based on the framework of the VNT put forward by Painter et al. in 2013. The research shows that the case *Sustainability for the Future* is mainly composed of four types of dis-course, namely, pictures, texts, videos, and audio. In addition, the video succeeds in

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

constructing the green, sharing, open, and honest SD concept of the BWO and shaping the image of China as an environmental-friendly and responsible host country. Furthermore, characteristics of the construction of the SD discourse of SMEs are fully discussed in section 4. Firstly, discourse narrative fosters emotional identity. Secondly, multimodal dis-course strengthens emotional mobilization. In addition, the discourse style is usually close to popular culture. This research provides a systematic framework for the multi-modal analysis of sustainable development discourse and can provide a reference for how to use the network platform to improve the effectiveness of communication, construction and transmit SD discourse in the new era.

Similar to many studies, our study is not exempt from limitations. One limitation is the small sample size. Although the BWO is the first Olympic Games that strictly implement the concept of SD throughout the whole process, the corpus of this study is only the SD report of the BWO, which may cause the limitation of this research to some extent. An-other problem is that the understanding of the SD discourse of SMEs has not been unified worldwide, which means the construction method of the sustainable discourse of the BWO may not be applicable to other countries.

Further study can be carried out from the following two aspects: Collect more Olympic sustainable development corpus and analyse the construction methods of their sustainable development discourse; Attempt to discuss the criteria for the sustainable development of SMEs to improve the universality of our research.

Reference

Olympic Agenda 2020 - Strategic Roadmap for the Olympic Movement (www.olympics.com)

D. Hinrichsen. Report of the World Commission on Environment and Development: Our Common Future. 1987.

Declaration R. Rio declaration on environment and development. 1992.

Millennium goals. United Nations. http://www.un.org/millenniumgoals/.2020

Sustainable development goals (2015-2030). https://www.undp.org/sustainable-development-goals.2015

Huan Qingzhi. Research on green discourse from the perspective of environmental politics. Journal of Jiangxi Normal University: Philosophy and Social Sciences, 2016, 49(4):3.

Li Chaoyuan, Wu Dongying. Analysis of Relational Speech Acts in Social Media Discourse of Multinational Corpora-tions. China ESP Research, 2016(2): 30-40.

Halliday, M. A. K. Language as a Social Semiotic. London: Edward Arnold, 1978.

O' Toole, M. The Language of Displayed Art. London: Leicester University Press, 1994.

Kress, G. & T. van Leeuwen. Reading Images: The Grammar of Visual Design. London: Routledge, 1996.

Kress, G. & T. van Leeuwen. Colour as a semiotic mode: Notes for grammar of colour. Visual Communication, 2012(1): 343-368.

O' Halloran, K. MD Analysis: Systemic Functional Perspective. London: Continuum, 2004.

Halliday, M. A. K. An Introduction to Functional Grammar (1st ed.). London: Arnold, 1985.

Li Zhanzi. A sociosemiotic analysis of MD. Foreign Language Studies, 2003(5): 1-8, 80.

Zhu Yongsheng. Theoretical basis and research methods of MD analysis. Journal of Foreign Languages, 2007 (5): 82-86.

Zhang Delu. Exploration of the Comprehensive Theoretical Framework of MD Analysis. China For-eign Language, 2009(1):7.

Wang Lifei, Wen Yan. Multimodal Analysis Method of Applied Linguistics Research. Foreign Language Teaching, 2008(3):5.

Li Xuening, Ma Lixia. MCA retrieval method and its application in foreign language teaching. Modern Educational Tech-nology, 2010, 20(11): 83-87.

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)

Huang Lihe. Corpus 4.0: Multimodal Corpus Construction and Its Application. Journal of the PLA Foreign Languages In-stitute, 2015, 38(3):8.

Feng Dezheng, Xing Chunyan. Spatial Metaphor and Multimodal Meaning Construction. Foreign Language, 2011(3): 58-64.

Chen Yumin. Analysis on the application of modal analysis in the research of multimodal foreign language teaching materials. Foreign language Teaching, 2012(1): 69-72.

Zhang Delu, Zhang Ke. Exploration of the Comprehensive Framework of Multimodal Criticism (Positive) Discourse Analysis. Foreign Language Teaching, 2022.43(1):1-8.

Painter, C., J. R. Martin & L. Unsworth. Reading Visual Narratives: Image Analysis of Children's Picture Books. London: Equi-nox, 2013.23

Bateman, J. A. Multimodality and Genre. London: Routledge, 2008.

Feng Dezheng. The New Development of Visual Grammar: A Visual Narrative Analysis Framework Based on Picture Books. Foreign Language Teaching, 2015, 36(3):5.

Martinec. R,A.Salway.A system for image-text relations in new (and old) media.Visual Communication,2005(4) : 337-371 .

Feng, D. & K. O'Halloran. Representing emotion in visual images: A social semiotic approach. Journal of Pragmatics, 2012 (44): 2067-2084.

Feng Dezheng. Discussion on Basic Issues of MD Analysis. Journal of Beijing International Studies University, 2017, 39(3):11.

Wang Zhenzhen, Wang Xiangfei, Yan Yiran. New media discourse strategy and national identity construction in mega sports events. Journal of Chengdu Institute of Physical Education, 2021,47(1):5.

Ding Yunliang. The Rhetorical Construction of National Identity in a Mediated Society. Academics, 2018,33(4):110-120.Olympic Agenda 2020 - Strategic Roadmap for the Olympic Movement (www.olympics.com)

© 2026 The Author(s). Published by Mason Publish Group. This open access work is licensed under a Creative Commons Attribution 4.0 International License. (http://creativecommons.org/licenses/by/4.0/)