Multimodal Discourse Analysis of Coca-Cola Video Commercials in Kazakhstan

Zhang,Ling   Dong,Mei

Xi 'an Shiyou University, Xi'an, Shaanxi, 710300, China

Abstract: A video commercial is a multimodal discourse that integrates multiple elements such as text, images, sounds, colors, and spatial layout. In recent years, China and Kazakhstan have become increasingly frequent in exchanges, and this study is of great significance to promoting the good-neighborly and friendly development of the two countries. Based on the theoretical framework of visual grammar of Kress and Van Leeuwen, this study adopts the method of qualitative and quantitative analysis , and explores how the multiple modalities in the video commercial integrate national culture into video from three aspects: representational meaning, interactive meaning and compositional meaning, so as to provide reference for international commercial advertising communication.

Keywords: Multimodal discourse analysis; Coca-Cola video commercial; Visual grammer; Qualitative and quantitative analysis

1. Theoretical Framework

Multimodal discourse analysis began with Kress and Van Leeuwen's book *Reading Images: The Grammar of Visual Design*, whose theory of visual grammar encompasses three aspects: representational meaning, interactive meaning, and compositional meaning. The representational meaning is defined as the ability of semiotic systems to represent objects and interrelationships in a world outside the represented world or in the symbolic system of a culture(Kress & Van Leeuwen, 1996:45), which is divided into two aspects: narrative representation and conceptual representation. The interactive meaning mainly explores the interactive relationship between the image and the image viewer, and expresses the attitude of the image viewer towards the world presented by the image. The interactive meaning is mainly manifested in four aspects:contact, social distance, perspective and modality. The compositional meaning is discussed from three aspects: information value, framing and salience(Kress & Van Leeuwen, 2021).

Based on the theoretical framework of visual grammar proposed by Kress and Van Leeuwen, this text selects Coca-Cola as the research object, which is a commercial advertisement in Kazakhstan, and analyzes the construction of the multimodal of this video from representational meaning, interactive meaning, and compositional meaning, which can enlighten the production of commercial advertisements in China and promote the better development of commercial advertisements.

2. Research Objects and Methods

In recent years, China has become increasingly close to the five countries in central Asia. Kazakhstan as a member of the five countries in central Asia, it is of great strategic importance for China to learn more about the culture of Kazakhstan. As a good-neighborly and friendly partner of China, Kazakhstan is a permanent comprehensive strategic partnership with China, and China attaches importance to its long-term development. This study took a 1 minute

and 32 second Coca-Cola advertisement in Kazakhstan as the research object, which showed the unique national culture of Kazakhstan while promoting Coca-Cola video commercials.

This paper adopts the method of qualitative and quantitative analysis , using ELAN6.8 video analysis software to conduct multi-modal discourse analysis on the video Coca-Cola of Kazakhstan according to the statistical results through multi-level synchronous annotation of video files, and calculate the proportion of time of different categories in the promotional video.

3. The Multimodal Discourse Analysis

(1) Representational meaning

The representational meaning mainly reveals the intrinsic connection between the elements of the image and their behaviors, which can be mainly divided into two aspects: narrative representation and conceptual representation.

By using the multimodal video analysis software Elan6.8, the factors in the promo are labeled and the data are counted. Table 3.1 shows the proportion of the images of narrative representation and conceptual representation in the Coca-Cola advertisement. Obviously, narrative representation appears more often in the video, and the time share is much more than conceptual representation. Shot 1 belongs to the analyzing process in conceptual representation, there is snow everywhere at the beginning of the video, which gives people a cold feeling. But when we look closely, there is a person relying on a tree, which contrasts with the whole atmosphere in the image and draw the audience's attention. Shot 2 belongs to the action process in narrative representation, the heroine in the video stumbles into the city where people sing and dance to make the whole atmosphere become cozy, thus contrasting with the atmosphere at the beginning of the video.

It can be seen that the video producer prefers to use narrative representation of the whole storyline, which enhances the interaction between the video and the audience and helps to attract the audience's attention, and at the same time, the clithes and food shown in the video plays an important role in the spread of the culture of Kazakhstan.

Table 3.1 The distribution statistics of representational meaning

	Hit count (unit：frequency）	Average duration (unit：seconds)	Time ratio(unit：%）
Narrative representation	34	1.91	70
Narrative representation	12	2.25	30

Shot 1

Shot 2

(2) Interactive meaning

Interactive meaning mainly describes the interplay between images and audience in real life, which explore in four dimension: contact, social distance, perspective, and attitude.

1) Contact

Kress and Van Leeuwen attach great importance to the eye contact between the participants and the viewers of an image, which is the establishment of a 'vector'. Eye contact is the most direct way to establish a connection between the image and the viewers, when "contact" appears between the images and the viewers. Contact can be divided into two categories: demand and offer act.

When there is eye contact between the character in the image and the viewer, and try to establish emotional connection, the image is called demand image, while the participant in the image does not make eye contact with the viewer, the main function of the image is to provide some basic information to the viewer, which belongs to the category of 'offer'.

As shown in the table 2, there are 11 contact factors in the video, which occupy 89% of the time in the whole promo. Specifically, the offering images appeared 7 times, occupying 84% of the time of the video, while the requesting images appeared only 4 times, occupying 5% of the video, and the average duration was smaller than that of the offering images. From the data in table 2, it can be seen that the promo tends to use the offering images to present information objectively. For example, shot 3 is the offering image, which shows a large gathering of friends and family to celebrate, everyone raises a glass to celebrate and drink, and at this time Coca-Cola becomes a popular drink on the table, which attracts the audience's attention, the advertisement always unintentionally highlights the importance of Coca-Cola, which makes the audience feel that it is a natural propaganda and not intentional.

In addition, the 'demand act' of the video is mainly realized through gazing and smiling (shot 4), and more happened at the end of the video, thus achieving the promotional effect of the advertisement and guiding people to buy Coca-Cola.

Table 2   The distribution statistics of contact

	Hit count (unit：frequency)	Average duration (unit：seconds)	Time ratio (unit：%)
offer	7	11.03	84
demand	4	1.26	5

Shot 3

Shot 4

2) Social distance

Social distance can usually be categorized into close-up, medium shot, and long shot, and the connection with the represented participants and the viewers is established through this degree of proximity.

According to the statistical data in Table 3, the near scene appeared 13 times, accounting for 55% of the time in the video, followed by the middle shot, and the distant shot appeared least. Close-up is mainly used to show the activities of the characters, as shown in shot 5. Close-up highlights the positive and hospitable mental outlook and state of the characters, thus narrowing the psychological distance with viewers. As shown in Figure 6, Coca-Cola uses close-up for many times, highlighting the theme of Coca-Cola and achieving a good publicity effect for Coca-Cola. The video is interspersed with a large number of medium-distance shots, accounting for 33% of the time, which are mainly used to show the urban atmosphere and the interaction between people, and play a good role in promoting the regional culture of Kazakhstan.

The flexible use of close-up shot, medium shot and long shot can not only enhance the viewing experience of the viewer, attractting the attention of the viewer and causing psychological resonance, but also highlight the theme and achieve effective publicity effect.

Table 3   The distribution statistics of social distance

	Hit count (unit：frequency)	Average duration (unit：seconds)	Time ratio(unit：%)
Close-up shot	13	3.88	55
Medium shot	10	3.01	33
Long shot	4	2.61	12

Shot 5

Shot 6

3) Perspective

Perspective refers to the image viewer's attitude towards the image world , which includes horizontal perspective and vertical perspective, and differences in perspective can reveal differences in viewers' attitudes toward the image participants. Horizontal perspective is categorized as front angle, side angle, and back angle, while vertical perspective includes high angle, eye angle and low angle. The high angle emphasizes the high status of the image viewer, the eye angle reflects the equal status between the image viewer and the image participant, and the log eye angle emphasizes the high status of the image participant.

As can be seen from Table 4, the side angle in the horizontal perspective appears most frequently compared with the front angle and the back angle, which indicates that the video mainly leads the viewers to watch the video from the third perspective, thus making the viewers further immersed in the Coca-Cola advertisement, achieving the effect of publicizing Coca-Cola effectively. The video is interspersed with some front angle to enhance eye contact with the viewer and getting the viewer's attention.

Eye angle is more often used in the vertical perspective of the video, accounting for 41% of the time of the video, indicating that the image viewer and the image participants are in a relatively equal position. Actually, eye contact will enhance the sense of intimacy between image viewer and the image participants. For example, shot 7 shows a scene of people sitting around a big table and having a meal, and through the eye angle of the vertical

perspective, it seems to pull the viewer into the scene and make video viewer feels like sitting with the people and drinking delicious Coca-Cola at the moment. At the same time, we can clearly see a large bucket of Coca-Cola placed prominently on the table, inadvertently emphasizing the theme of the video. The high angle accounts for 15% of the time, mainly used to show the cityscape and special food, which is a good way to publicize the culture of the country of Kazakhstan. Meanwhile, the low angle is mainly used to show the natural landscape and solemn activities. Finally, the advertisement enhances the viewer's visual experience through the comprehensive use of different perspectives.

Table 4 The distribution statistics of perspective

	Hit count (unit：frequency）	Average duration (unit：seconds)	Time ratio(unit：%）
front angle and	4	1.15	5
side angle	9	0.95	9
back angle	1	2.46	2
high angle	7	2	15
eye angle	13	2.94	41
low angle	4	2.24	9

Shot 7

4) Attitude

Attitude refers to the degree of truthfulness of an image's reflection of reality, which is reflected in eight dimensions: color saturation, color differentiation, color coordination, contextualization, reproduction, depth, illumination, and brightness. Attitude can reflect the realism and credibility of an image to the objective world it depicts, and Kress and Van Leeuwen categorize attitude into three types: high modality, medium modality, and low modality. The level of modality is directly proportional to the colorfulness of the elements in the image. The images of high modality using highly saturated colors that are vibrant and distinguishable from one another. The images of medium modality have low saturated colors, low contrast and brightness, and are predominantly washed out. The images of low modality are relatively monochromatic, with images in black and white tones only.

We can see from Table 5 that the time of high saturation color images of the video accounts for the largest proportion, accounting for 77% of the whole video, and the overall video mainly uses images in red color, which is high modality. By comparision, medium-saturated color images account for less time in the overall video. At the beginning of the video, it is mainly black and white, showing a snowy scene, and the sun on the horizon shines on the afterglow, which seems a little peaceful, which belongs to the medium modality. At this time, the heroine is wearing a red dress and leaning under the tree, as shown in Figure 8, the color is bright, it belongs to the highly saturated color picture, which forms a sharp contrast with the previous picture, and the color discrimination is very large, which brings a visual impact.

At the same time, the difference in color saturation in the video not only brings about the difference in color contrast, but also has another deep meaning. As shown in shot 9, the heroine appears in the center of the picture in a long red dress, which contrasts with the surrounding snow-capped mountains. While the right side of the figure is covered with snow, the left side of the images is full of flowers with the arrival of the figure, which implys the hope of life. Shot 8 shows us the bleakness and silence of the city with the coming of winter. However, at this time, the heroine enters the city in a bright red dress (shot 11), which brings a color impact, meawhile, countless flowers bloom behind the heroine with every step she takes, symbolizing vitality and hope that the heroine brings to the city. It wasn't until the video at 1 minute and 14 seconds (shot 12) shows Coca-Cola on the dinner table that the viewer of the video suddenly realized that the video personifies the Coca-Cola product and gives it life to the video from beginning to end, making the ad full of storytelling.

Table 5 The distribution statistics of attitude

	Hit count (unit：frequency)	Average duration (unit：seconds)	Time ratio(unit：%)
high modality	33	2.16	77
medium modality	13	1.58	22
low modality	0	0	0

Shot 8

Shot 9

Shot 10

Shot 11

Shot 12

(3) Compositional meaning

Compositional meaning is depicted in three aspect: information value, salience, and framing.

The information value is reflected by the position of the elements in the image, and different positions of the elements in the image reflect different information values. As shown in shot 9, the heroine is placed in the center of the images to highlight the character as the main message. Throughout the video, the main character appears in close-up several times and is located in the center of the images, which in turn provides the audience with the main information. At the same time, the video also gives close-ups of other supporting characters, for example, people dancing happily in national costume , as shown in shot 5, which highlights the value of the message of national culture. Most importantly, the product Coca-Cola is placed in the center of the screen and close-ups are given many times especially at the end of the video, such as shot 3、shot 6 and shot 7, which highlights the theme of Coca-Cola and conveys to the audience that Coca-Cola is a drink for both friends' gatherings and large-scale celebrations of newcomers. There are many images in the video that are dominated by the main message, the main purpose is to publicize Coca-Cola and to promote the culture of the nation.

Salience refers to the fact that the image or text conveying information is made more conspicuous and prominent through color differences. Salience reflects the attractiveness of an element in an image to the audience: the larger the area of an element in the image, and the greater the color difference with the surrounding elements, the higher the saliency of the image. As shown in shot 9, the main character's clothes in the video are brightly colored, forming a strong contrast in color with the surrounding image, which brings visual impact to the audience and further highlights the prominence of the character. At the same time, placing the character in the middle of the screen plays an emphatic role for the character. Meanwhile, as shown in shot 13, with the character as the center, the left side of the picture is full of flowers , while the right side of the city is lifeless, and the color difference between the left and right side forms a contrast, which further accentuates the theme that Coca-Cola bringing happiness to people's lives, thus playing a very good publicity role for Coca-Cola as a commodity.

Framing refers to the way an image is framed , through the division of the actual lines or space frame in the picture, to link or distinguish the elements in the picture. As shown in the video at 51 seconds (Figure 13), when the main character walks towards the silent city, the left side and the right side of the screen are divided with the character as the axis of symmetry, which creates a contrast in the screen, and then shows the great changes brought about by the character after he enters the city, thus highlighting the product of Coca-Cola, and playing a better publicity role for it.

Shot 13

4. Summary

Based on the visual grammar theory of Kress and Van Leeuwen, this paper analyzes multimodal discourse of Coca-Cola video commercials in Kazakhstan from the perspective of representational meaning, interactional meaning and compositional meaning, and explores the construction strategies of the Coca-Cola video commercials

in Kazakhstan, which is conducive to promoting the international dissemination of advertising.

References

[1] KRESS G, VAN LEEUWEN T. Reading Images:The Grammar of Visual Design[M].London: Routledge,1996.

[2] Dong Mei, Yuan Xiaolu. Research on the Construction of a Multimodal Aesthetic Criticism Discourse Analysis Framework [J]. Foreign Language Education. 2021, 42(01):77-82.

[3] Zhou Yixing, Dong Mei. A Comparative Analysis of the Interactive Meaning of City Promotional Videos from a Multimodal Perspective: Taking *Eternal Ancient Capital, Always Welcome to Chang'an and The World of Luoyang* as Examples[J]. PUBLIC COMMUNICATION OF SCIENCE & TECHNOLOGY, 2023.

[4] Guan Shouxian. A Multi-modal Discourse Analysis From the Perspective of Visual Grammar Theory: A Case Study on *Foshan City Image and Tourism Promotional video* [J]. Journal of Changjiang Engineering Vocational College, 2024, 41(03):67-72.

[5] Xie Yaping. A Multimodal Discourse Analysis of Creative Short Plays from the Perspective of Visual Grammar: A Case Study of *Little Nian Beast and the Demon Catcher* [J]. Journal of New Chu Culture, 2024, (16):58-6.