A Multimodal Analysis of Traditional Chinese Paintings: Based on the Grammar of Visual Design

HOU Haotian[a]; WANG Yue[b],*

[a] Undergraduate, Department of Foreign Studies, North China Electric Power University, Baoding, Hebei, China
[b] Lecturer, Department of Foreign Studies, North China Electric Power University, Baoding, Hebei, China
*Corresponding author.

Abstract

With the development of multimedia technology, the modalities in culture interpreting are diversified, and it is essential to investigate the function of multimodal semiotics in intercultural communication. Based on Kress and Van Leeuwen's grammar of visual design, we discussed its adaptation in the traditional Chinese count painting and the meaning of depicted semiotics in cultural innovation in the Application of the Chinese paintings the *Twelve Beauties of Prince Yong* published by the Palace Museum. As one of the representatives of Chinese count lady painting in the Qing Dynasty, it is found that the *Twelve Beauties of Prince Yong* could conform to the fundamental principle of the grammar of visual design. Although two distinct modalities—image and text—are related tightly, they express meaning and load information through different visual positions. Interpreting the traditional culture with multimodal perspectives can highlight the core based on containing originality and close to contemporary aesthetics. It is hoped that this study can promote the conversion of some ancient authentic arts or relics from single mode to multimode, enhancing the efficiency of intercultural communication.

Key words: Grammar of visual design; Multimodal analysis; Traditional Chinese paintings

1. INTRODUCTION

Translation is a vital process in cultural communication. According to Roman Jakobson (1959), the question of translation becomes much more entangled and controversial in jest, dreams, and poetry above all. Traditional Chinese paintings, calligraphy, and other symbols of China are reflections of Chinese cultures, if relying on a lexical monomodal approach to translate them is very hard and will probably lose their original meaning. Hence, Chinese traditional culture translation is a significant concern of Chinese linguists and translators in the New Era. With the development of technology, modern media has diversified, providing a multimodal view to deal with Chinese culture translation. From the "single-modality" to "multimodality," visual and auditive modality expands and multiplies the narratives of China's story, enhancing the cognitive of China's value worldwide (Wu & Mou, 2022). This paper will investigate precisely the intersemiotic translation in the Chinese painting *Twelve Beauties of Prince Yong* by a method of visual grammar. The aim is to explore the feasibility of multimodal translation in Chinese cultural heritages and provide references for cultural publicity.

2. LITERATURE REVIEW

Jakobson (1959) first labeled the three ways to interpret a verbal sign: intralingual, interlingual, and intersemiotic translation. The practices of interpreting and translation have been centered on lexical language studies and aimed at exploring the phenomena, principles, and methods of *translation proper*, i.e., interlingual translation. Therefore, the other semiotic modes like images, sounds, and colors were ignored until the 1990s; multimodal discourse

analysis rose in the field of linguistics from the western, helping scholars overcome the limitations of monomodal discourse analysis or lexical-oriented translations (Zhu, 2007). After Roman Jakobson, western studies have been increasingly focusing on interpreting verbal signs utilizing signs of nonverbal sign systems or between two nonverbal model systems (Wu, 2021). Kress & Van Leeuwen (2006) applied Halliday's metafunctional theory from verbal modality to the field of other nonverbal semiotics and put forward the Grammar of Visual Theory, which derived and expanded systemic functional linguistics.

Researchers have investigated the functions of multimodal symbols in cultural communication in various aspects of society and with a diversified approach in different disciplines. Intersemiotic multimodal translation can break through the traditional relation between the source language and the target language, for instance, in cartoons (Mo, 2020), poems (Hu, 2008), picture books (Teng & Miao, 2018), posters (Li, 2021), news (Li, 2021), advertisements (Shi, 2021), and traditional Chinese literature (Wang & Zhang, 2022). At the logical and methodological level, the grammar of visuals has been applied to Forceville's multimodal metaphor (Mo, 2020), Martinec & Salway's graphic relations theory (Mo & Li, 2021), and critical discourse analysis (Puspitasari et al., 2021). It is also widely developed in computer science, with convolutional neural networks (CNN) (Bäuerle et al., 2021), a subcategory of machine learning. From the point of view of cultural communication, multimodal interpreting of work, compared with single-mode translation, could remove language barriers to communication activities. With the support of high-tech, translation based on viewers' cognition and experience can adapt to the contemporary reading style by linking history with reality and eventually satisfy the viewers' needs. From the perspective of disciplinary development, the grammar of visual design based on systemic functional linguistics has shifted the focus of translation research from text to multimodality, proving that translation is not a pure lexical conversion but a rewriting and recreation in a foreign culture. Scholars have turned their research objective in this field from summarization to empirical in specific practical applications in recent years. However, most of their research objects are contemporary works or technologies, and there are still gaps in cultural relics. Therefore, the authors attempt to explore the analysis of ancient works under the grammar of visual design and provide more reference value for the development of this theory.

3. GRAMMAR OF VISUAL DESIGN

Kress & Van Leeuwen (2006) regarded images as social semiotics and proposed the grammar of visual design, which is the theory to investigate the profound meaning of integrated texts and the interaction between images and verbal texts with their composition. There is a one-to-one correspondence between visual grammar theory and Halliday's systemic functional linguistics, as shown in Figure 1.
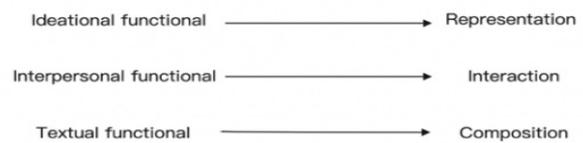
Figure 1

The representations include narrative representation and conceptual representation; the difference is whether there is the participation of vectors. Vectors are usually marked by a diagonal line or formed by the participant's movements and other elements. Narrative representation is a dynamic process consisting of action, reactional, speech, and mental processes. It is frequently used to reveal the implication of vectors between participants, which forms by their gestures or behavior. But the conceptual representation is relatively static and does not form vectors, consisting of the classification, the analytical structure, and the symbolic structure, representing participants' more generalized and stable essences.

Kress & Van Leeuwen (2006) asserts that the interactive meanings of visual modalities mainly express a particular relationship between the producers, the represented participants, and the viewers. The interactive meanings of contact, social distance, and attitude are usually achieved by visual communication, which is ubiquitous in daily life and occurs immediately when we have eye contact with objects.

Ultimately, the third element, the meaning of composition, could make representational and interactive participants a meaningful whole, illustrating the spatial layout of an image, including information value, salience, and framing (Kress & Van Leeuwen, 2006).

Not only linguistic semiotics like the text can express meanings, but visual semiotics can also achieve this function with their meaning potential and becomes an essential event in interpreting the meanings. Grammar of Visual Design explores how visual elements consist of integrated meaning in images or other forms. In this direction, the grammar of visual design is a powerful tool for analyzing visual semiotics.

4. DATA SOURCES

Twelve Beauties of Prince Yong is one of the remarkable paintings in the palace collections of the Qing Dynasty, made of silk, but the painter(s) is indeterminate. It is a series of traditional portraits, a total of twelve, in which the beauties are all dressed in costumes of the Han dynasty, with dignified postures and graceful faces. The background decorations are incredibly luxurious, representing a solid feeling of royalty. (Yang, 2011)

In May 2013, the Palace Museum released its first official application with the same name *Twelve Beauties of Prince Yong*. However, the Palace Museum did not locate its function as a comprehensive guide tool, but as an introducer, by analyzing specific objects or participants as precise incisions for interpreting the original paintings, providing information on collections of genuine cultural relics in the Palace Museum with the same or similar furnishings in the paints. The application has won the favor of the masses and received many comments and became a topical issue and raised much research among scholars. From now, it seems that it is a novel and efficient way and a successful case for the innovative development of splendid traditional culture. Therefore, we have selected this application as the analysis object and investigated its layout and the original painting system from the perspective of the grammar of visual design.

5. ANALYSIS

To uncover the role and visual or textual function of multimodal semiotics in the innovative translation of traditional Chinese meticulous brushwork painting, some of the pictures were taken from the Application as samples to analyze from the point of view of the *representational*, *interactive meaning* and *the meaning of the composition*.

5.1 Representation

5.1.1 Narrative Representation
In *Narrative Representation*, the *Action & Reactive* process is non-projective, consists of *participant*, *vector*, goal, or *phenomenon*, and is often the most salient part of an image and can be identified by the saturation of color or the sharpness of focus and so on.

Figure 2

Note. Figure 2,3,4,5,6,7,9,10,11, are taken from "Twelve Beauties of Prince Yong" © [2022] The Palace Museum. All rights reserved.
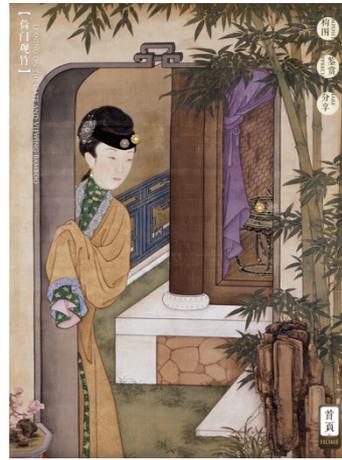
Figure 3

Figure 4

The *Action process* in the above three Figures predominantly occurs around beauties, i.e., the *Actor*, and still can be generally divided into three subcategories: unidirectional action, bidirectional action, and non-transactional action. In Figure 2, for instance, a unidirectional vector is formed from the *Actor*. Still, it does not lead to another depicted element, maybe the bamboo fence, which signifies an un-transactional *action process* that allows us to imagine what or why she is contemplating and makes the *Actor* more mysterious. From a comprehensive view, the *actor* with Wood Ruyi Scepter is arranged as a conspicuous element at the center of the whole image. At the same time, Ruyi, in the form of lingzhi fungus itself, constitutes an exhaustive analytical structure and conveys the metaphorical meaning that everything is as one wishes and wealth. The beauty in Figure 3 leaning on the gate and posting an inclined gesture slightly to the right side (from the Viewer) constitutes a diagonal *vector*, starkly contrasting the gate's vertical and straight frame. The beauty as the *Actor* views the bamboo, hence the bidirectional action formed, may entice the viewer's attention between the two elements. In Figure 4, the *Actor* does not engage in other parts. Still,

she is detached from them, like in Figure 2, constituting an un-transactional *vector* by herself, which will benefit the viewer to concentrate on appreciating the other objects around the *Actor*; thus, they would not be excessively attracted by the center of the image.

The *reactional process* consists of a *Reactor* and a *Phenomenon*. Reactions also can be divided into transactional or non-transactional without the *Phenomenon* same as the action process. The medium between the *Reactor* and *Phenomenon* is formed by the direction of the eye-line of one or more *participants*. According to Kress & Van Leeuwen (2006), only humans and animals with facial expressions or expressions of emotion with their eyes can be called *Reactors*. Therefore, the beauties in the above three images now become the *Reactors*, and their eye-line form a *vector*, respectively pointing to the direction of the fence, the bamboo, and Songhua inkstone on the square table, as we have discussed. Their facial countenance reflects their subjective reactions to the *Phenomenon*, which may be pleasure, solemnity, or thoughtfulness, also echoes from the verbs in the titles of these paintings, for instance, "standing," "viewing," and

"facing." *Verbal* and *mental* processes are represented through dialogue balloons and thought bubbles, which do not occur obviously in this application; thus, they are not a pivotal part of our research.

In this part, we could draw a brief conclusion that, in this series of paints, absent-minded beauties isolated from the environment are represented as the agentive *Actor* or *Reactor* dominantly in the unidirectional or non-transactional process.

5.1.2 Conceptual Representation

The *conceptual representation* in this application usually happens in the form of hypertext that is used to descriptive illustrate the elements and objects in the painting, which is a disruptive innovation in the field of traditional culture interpretation. Hypertext technology plays another critical role in linking the details in the original image tightly with the homologous or similar collections in the Palace Museum, for instance, the Wood Ruyi Scepter and the Wood Ruyi Scepter Carved in the form of a Lingzhi Fungus (as in Figure 5), or Earrings and the Drop Earrings Inlaid with Pearls (as in Figure 6).

Figure 5

Figure 6

This display approach not only increases the authenticity of the images but also enhances users or viewers' understanding of both paintings and existing collections, adding more cultural details to them from a very close distance. The elements in Figure 5 constitute a *Covert Taxonomy*, a set of participants is distributed symmetrically across the visual space – at equal distance from each other, equal in size, and oriented toward the vertical and horizontal axes in the same way, according to Kress & Van Leeuwen (2006, p.82), representing a relationship that they are all subcategories of the Ruyi or the products made of Lingzhi. In China, Lingzhi has been widely spread by the alias "holy grass" and the symbol of immorality since ancient. On the contrary, Figure 7 shows a kind of *Overt Taxonomy*. The earring is a general scope, and the Drop Earrings Inlaid with Pearls is contained; thus, we can infer that earrings are a very common

ornament to compliment women's beauty and elegance from past to contemporary combined with the textual illustration at the right of the image.

Under the grammar of visual design, elements in the images or on the screen are usually involved in multiple structures, such as Ruyi, which can be depicted as a *participant* in the *action process* with the woman (*Actor*), can also be used as an *attribute*, and the woman also acts as a *carrier*. Thirdly, it also has a *symbolic* meaning; as we mentioned earlier, it is a kind of carrier of good wishes. Therefore, hypertext technology plays a prominent function in spreading more information about objects, and it is worthy of being applied and promoted in other fields.

5.2 Interactive Meanings

Before discussing the *meanings of interactive*, to better understand the relationship between producers, represented participants, and the viewers, and how they

"communicate" with each other, it is essential to define who the producer is. Based on the literature review and picture comparison, Yongzheng Emperor (Aisin Gioro Yinzhen) is also one of the creators of this image, and these twelve paintings are not portraits of twelve people but were all painted according to Yinzhen's wife (nee Nala) as a model (Yang, 2011). The original silk painting comes from the screen in the room of the Yongzheng Emperor. The calligraphy on the paintings, inscribed by Aisin Gioro Yinzhen, was deliberately covered by the glaze or some framing lines, which no one can do except himself; hence he is not only the producer but also the viewer (Zhang, 2014).

The *interactive meaning* was condensed into three subcategories: *Contact*, *Social distance*, and *Attitude*, which we will illustrate respectively from different subjects.

5.2.1 Contact

The *Contact* between the represented *participant* and the viewer behind the screen must be either *Demand* (in which the participant gazes at the viewer) or *Offer* (in which absence of gaze at the viewer). Observing the images from a very close distance, the humans in these paintings did not gaze at the viewer, avoided them intentionally or unintentionally, and contemplated the surrounding. Therefore, they are all *Offers*, giving the viewer information about details that infer the inner world of the character's mind or represent the image's cultural background.


Figure 7

For instance, in Figure 7, a poem with the inscription "Qi Chang" hangs on the wall behind the beauty. This quatrain poem illustrates the desolation of a woman in a boudoir. Among them, nouns of plant names like "Chrysanthemum," "Huang Hua," "Orchid," and "Thin Bamboo" are common metaphors in ancient Chinese that can create a cold, lifeless isolated environment for contemplative women and form the *Symbolic* of the *conceptual representation* as we mentioned above. It is worth noting that these symbolic nouns are also the

natural plants in the painting, conspicuously representing the communication between images and texts. Through these *Offers*, it is convenient for viewers to trace back the condition when the producers drew or made the paintings. The objects in the images are all Yinzhen's favorite collections, which convinced Yinzhen "feminize" and put himself in such an ideal environment to some degree. Ostensibly, they are a series of paints. Still, we can regard them as a spontaneous textual manuscript of Yinzhen, expressing his monologue on the political intentions or ambitions behind the famine character. What confused and puzzled Yinzhen was how to gain the appreciation of the Kangxi Emperor and the other ministers who held great power so that he could outstand other Princes and succession to the throne. Nevertheless, due to the political and challenging situation in the palace, no one could express his thought directly but used metaphors to describe through painting, poems, and other forms to "encrypt" the meaning.

5.2.2 Social Distance & Attitude

From the view of the function and meaning of *Social Distance*, the producers use a *medium-long* shot to depict the character, which can show the total figure of the character. This kind of shot corresponds to a social distance between social and impersonal; it expresses the meaning of keeping a distance from viewers, which makes the relationship just like a stranger. Then, viewers only superficially observe the images and are attracted by royal and incredible decorations or furniture. However, why women still seem worried in such an ideal environment is incomprehensible to viewers until viewers could be placed in the position of Yinzhen to consider and eventually better understand the images from a profound scale.

Geometrically, two conceptions of perspective were taken from Kress & Van Leeuwen (2006), *frontal-isometric* and *angular-isometric*, as the distinctive style the authors analyze. The front of the *frontal-isometric* depicted objects has not been distorted and is a rectangle in naturalism, yet we still can see the top and side of them, which is more objective without *involvement* and *detachment*; On the contrary, the front view of the *angular-isometric* is no longer a square but gives people a feeling of looking down, which keeps a distance between the depicted participant and the viewer like detachment.
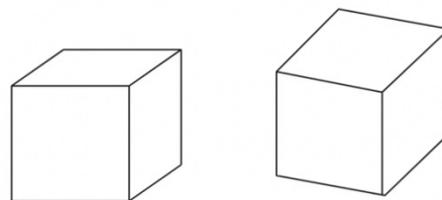
Note. This image is taken from "Reading Images: The Grammar of Visual Design" by Kress & Van Leeuwen, 2021, p.142.
Figure 8
Frontal-isometric and angular-isometric

The *frontal-isometric* perspective is ubiquitous in these paintings; for instance, the different kinds of Square Table shown in Figure 9 are detached and displayed in a separated subpage from the original images, which enhances the objectivity and salience of inanimate elements by depicting them comprehensively. Another homologous example is the treasures on the treasure shelf behind the woman in Figure 5. They are presented in an abnormal top-down perspective (like the *angular-isometric* perspective); hence if the creator's (and viewer's) sight line is parallel to the character, it is impossible to look down on items above the character. According to the painting found in Yinzhen's study room, it is reasonable for viewers to infer that this perspective may be convenient for Yinzhen to appreciate his collections from an objective point of view at any time.


Figure 9

Figure 10

Compared with classic paintings, characters are often highlighted, and other objects are distorted or depicted from a single angle; the practice of perspective in *Twelve Beauties of Prince Yong* reflects the society that the Chinese lady paintings no longer only focus on characters under the influence of Western ideology and methodology was introduced into China in the late Ming Dynasty, and gradually pay attention to the objective and meticulous drawing of the naturalism environment.

In further discussion of the *interactive meaning* of texts and images, it is worth mentioning that although most elements are depicted as detachment by oblique and high angles, there are also some representations of involvement that could close the distance from the viewers, for instance, the calligraphy of poems at the right side of the beauty (Figure 7, Figure 10) that represented by a frontal angle with eye level, as if the viewer is more likely the creator or even the position of Yinzhen rather than the visitor. As mentioned above, these are essential tools to reveal the profound spiritual world of producers.

5.3 The Meaning of Composition

In the above images, the positions of the characters, the table, or the wall are exchanged, however, the characters are still non-transactional in the *medium-long* distance, and their *representational* and *interactive meanings* will not alter relatively. However, *the compositional meaning*, as the mediator between *representational* and *interactive meaning*, will change to a degree. The information value, framing, and salience of the visual texts will be illustrated as follows to clarify their function in interpreting the visual design of culture.

5.3.1 Information Value

Different spatial orientations of elements in an image load specific information values. From a horizontal juxtaposition, the aspect on the left is often the *Given*, the right is the *New*, and the principal elements are represented as the *Ideal* (presented as the "generalized or idealized essence of the message") from vertical conjunction. The bottom is the *Real*, complementing with "factual and practical detail" (Kress & Van Leeuwen, 2006, p.217). The left edge of the wall (in Figure 2) behind the character forms a framing line that separates the image into two parts. The left part comprises the vermilion fence, green bamboo, and rockwork artificial hill, typical representatives of classical ancient Chinese royal landscape art. Thus, symbols of *Given* information would arouse viewers' inherent impressions of China or the Orient. On the opposite, the participants on the right, primarily the dressing box in the window, the character, and the bamboo fence constitute the *New* that is "unknown or unfamiliar" to viewers, representing the beauty's attractiveness and favorable living environment. At the same time, Figure 2 is simultaneously divided into upper and lower parts by the framing line of the bamboo fence covered with flowers. Above the framing line, a spiritual world was illustrated, for instance, by the integrated semiotic Ruyi with a virtuous meaning (as we discussed its *symbolism* before), a beauty's Guanyin standing posture, and it may also represent Yinzhen's wish of the ideal life. As we scroll down the screen, the bottom part dominantly is bamboo fences covered with a compact cluster of flowers that takes up half of the screen, expressing that beauty lives in wealth but still is "trapped" in the complicated fence, with loneliness and a strong desire for freedom. To apply various screen size, the designer of the App become scrolling the screen up and down from *Ideal* to *Real* to see the whole image a reality, which also sets a more conspicuous contrast of composition and enhance the meaning expression of the orientation, and

echo with the previous analysis of the *representational* and *interactive meaning*.

5.3.2 Framing

The second part of the composition is *framing*, which includes two situations of elements in images, *disconnection* and *connection*. The stronger the frame, the more isolated the element is from its environment. The dressing box in Figure 2 (separated by lines), the different bonsai outside the door in Figure 3 (separated by blanks), and the objects on the shelves in Figure 4 (separated by a grid), are typical phenomena of usage of *disconnection* that "highlight" the individuality of each element. From another perspective, these "separated" things are also interrelated, that is, *connected* within each other, in which characters play a more prominent role by overlapping or integrating other elements, as a mediator exists between the *given* and *new*, or the *ideal* and *real*. A structure called visual *rhyme*, in which the meaning of connection is condensed, the common feature in two or more elements, is arranged as visually or typographically harmony. For example, in Figure 10, the curtains, beauty's wind hat and dress, and bookcases form a visual color *rhyme* with blue tones in the vertical direction, explaining the whole image with spacial logic.

5.3.3 Salience

Salience is a polar range, not the neither-nor conception, that creates a hierarchy of importance that guides the viewers' reading order, and it is also a multimodal principle. The listening or speaking process often outlines *salience* by stressed and unstressed syllables. It forms a contrast through their strength or length, and in visually integrated texts, such as films, photographs, or paintings, it is mainly quantized in terms of size, sharpness of focus, tonal contrast, and so on.


Figure 11
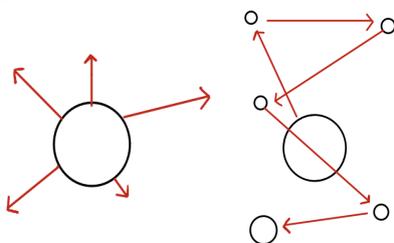
Figure 12

The introductory information about the cultural relics in the painting is compressed into cloud-shaped or lotus-shaped buttons, which are also presented in the form of hypertext. These designed hypertext links have a distinctive shape then as the high-salience point. In Figure 12, circles metaphysically represent the salient point, and the larger the circle, the higher the salience value; the red arrow is the simplified track of the viewer's attention that is "moving" on the image (i.e., the line segment formed by the sum of the infinite intersection of the eye-line and the plane where the image is). In the centralized images, beauty as the central element has the highest salience value, represented by the largest circle, then the viewer will probably gaze at the woman primarily. If viewers appreciate the original painting ostensibly, the track of attention shows a disordered radial shape around the character (as shown in the left part of Figure 12) and may be attracted by any less salient element after looking at the center, for instance, the bamboo or other collections, and thus may pass by the prior elements which are more worthy to investigate or acknowledge in the image.

Hence, it is necessary to improve the image, hint to viewers what elements are important, and guide them comprehensively to browse the whole image without making them bored. After multimodal reproduction, the image becomes multiple salient, then a hierarchy of elements was built that could navigate the viewer's gaze to move between salient points, from high values point to low values point; when salience value is the same or approximate, may from top to bottom or from left to right, as shown in the right part in Figure 12. Therefore, more details and information are conveyed under the intentional guide by the producer while the original image is not destroyed by these distinctive hypertexts because of their antique color and shape.

CONCLUSION

Through the multimodal analysis of *Twelve Beauties of Prince Yong* and its official Application, the results could be summarized as follows:

1. *Twelve Beauties of Prince Yong*, as one of the representatives of Chinese count lady painting in the Qing Dynasty, conforms to the fundamental principle of the grammar of visual design.

2. Although two distinct modalities, image, and text, are related tightly, they express meaning and load information through different visual positions.

3. Multimodally interpreting the traditional culture can highlight the core based on containing originality and close to contemporary aesthetics.

When creating an intercultural product, it's primarily that we not only should adapt to foreign cultures like reading habits but also pay attention to specific hardware or software conditions, thus could improve the viewers' experience when they appreciate the ancient works.

Multimodality is an irresistible trend in the "digital universe" we look forward to our study being conducive to multimodal semiotic research and the interpretation of the aboriginal culture.

REFERENCES

Hu, D. (2007). Multimodal discourse analysis of Emmett Williams' poem "She loves me not". *Foreign Language and Foreign Language Teaching*, (11), 16-19+40

Jakobson, R. (1959). On linguistic aspects of translation. In L. Venuti (Ed.), *Translation studies reader* (2004, 2nd ed.), Routledge.

Li, H. Y. (2021). Multimodal discourse analysis of propaganda posters from the perspective of visual grammar: Taking the movie "Youth as an Example" as an example. *Modern Communication*, (13), 87-89.

Li, Y. (2021). Analysis of the Representational Significance of Interactive News from the Perspective of Visual Grammar: Taking the winning works of the China News Award Media Convergence Award as an example. *Jin Media,* (08), 54-56.

Mo, H. F. (2020). Multimodal metaphor construction of political cartoons from the perspective of visual grammar. *Journal of Hunan University of Humanities and Science,* (01), 109-113+118.

Shi, Q. (2021). Multimodal research on visual grammar in advertising discourse. *Overseas English*, (07), 245-247.

Wu, Y. (2021). Multimodal translation research under media turning. *Foreign Languages* (Journal of Shanghai International Studies University), (01), 115-123.

Yang, X. (2011). Exploring the secrets of Yinzhen's picture of a beautiful woman surrounding the screen. *Journal of the Palace Museum,* (02), 6-23+158. doi: 10.16319/j.cnki.0452-7402.2011.02.004.

Zhang, Y. (2014). Brows and Eyes" beauty map - Interpretation of Qing Yongzheng's "Twelve Beauties" from the perspective of "eyebrows and eyes. *Rong Bao Zhai,* (11), 176-189. doi: 10.14131 /j.cnki.rbzqk.2014.0040.

Zhu, Y. S. (2007). Theoretical basis and research methods of multimodal discourse analysis. *Journal of Foreign Languages,* (05), 82-86. doi:10.16263/j.cnki.23-1071/h.2007.05.034 .

Bäuerle, A., van Onzenoodt, C., & Ropinski, T. (2021). Net2vis– a visual grammar for automatically generating publication-tailored cnn architecture visualizations. *IEEE Transactions on Visualization and Computer Graphics, 27*(6), 2980-2991.

Kress, G. R., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Psychology Press.

Mo, A. P., & Li, M. (2021). Multimodal pragmatic strategies for Chinese culture translation in the digital background. *Foreign Language Teaching*, (06), 68-74+11.

Puspitasari, D., Widodo, H. P., Widyaningrum, L., Allamnakhrah, A., & Lestariyana, R. P. D. (2021). How do primary school English textbooks teach moral values? A critical discourse analysis. *Studies in Educational Evaluation,* 70, 101044.

Teng, D., & Miao, X. W. (2018). The meaning construction of multimodal metaphors in picture books from the perspective of visual grammar. *Journal of Foreign Languages,* (05), 53-59. doi:10.16263/j.cnki.23-1071/h.2018.05.010.

Wu, Y., & Mu, Y. W. (2022). Research on multimodal national translation strategies of Chinese stories. *Foreign Language Teaching,* (01), 76-82. doi: 10.16362/j.cnki.cn61-1023/h.2022.01.012.

Wang, Q. Y., & Zhang, B. (2022). Multimodal intersemiotic translation from the functional perspective: A case study of part of the corpus of materia medica. *Cultural Industry,* (09), 89-93.