

Research on the Evaluation Mechanisms of the Design Form and Connotation of Digital Games from a Multimodal Interaction Analytical Perspective

Tianshu Ji^{1,a,*}

¹University of Adelaide, 115 Grenfell Street, Adelaide, South Australia, 5000, Australia
a. Tessa20220828@outlook.com

*corresponding author

Abstract: This research endeavor primarily focuses on computer, or electronic transport games. Electronic games are based on a computer's ability to utilize a proportional calculation pattern that simulates imagined human behavior. While the game industry has grown substantially in recent years, methods for evaluating such games remain underdeveloped. Existing evaluation systems also tend to overlook gender bias, overt violence, and a lack of sexual imagination. To construct a more appropriate evaluation system in terms of both form and function, and operating from the perspective of linguistic multimodal interaction analysis, this study analyses the form design of video games with regard to game action levels and modal configuration, as well as the multiple functional meanings in game design at the ideational, interpersonal, and textual, and expressive levels. In addition, the principles of multimodal discourse analysis are applied to popular games, such as *Harry Potter: Magic Awakened* and *It Takes Two*, to enrich discussions of this theoretical framework and provide systematic suggestions for game design's diversified and innovative development. The goal of this article is for scholars in various fields to use the analytical paradigm offered herein to examine multimodal expressions in more evaluation mechanisms of game design in the future.

Keywords: evaluation mechanisms, game design, multimodal interaction analysis, research paradigm

1. Introduction

Much like the news, movies, social media, and other contemporary communication mediums, digital games have gradually been granted a certain status in people's lives. As a form of interactive content expression based on new electronic technology, digital games have become a new platform for cultural transmission and spiritual expression. Books are adapted into movies, and movies are re-created into games, such as *Harry Potter: Magic Awakening*, one of the most popular games in China as of 2021, along with other movie-turned-games, such as *Tomb Raider* and *Warcraft*. The blurred lines between film and game can be further noted in television programming; as the medium became more and more interactive, it absorbed certain elements that originated from games. For example, *Black Mirror: Bandersnatch* represents the interactive drama, which mainly borrows from the game form related to the player's choice. Furthermore, content sharing, in recent years, from a game to the movie *Squid Game*, is an example. The vivid style and flexible structure of computer games were

likewise absorbed into the work of musical composers and visual artists [1]. From an economic standpoint, the game industry has become one of the fastest-growing sectors in today's financial market. According to an annual report published by Super Data in 2020, gross digital game revenue worldwide reached a staggering \$126.6 billion. In terms of scale and monetary value, the potential research value of digital games cannot be understated; however, when game production becomes increasingly scaled up, "mom and pop" or other independent game studios, though they may be highly innovative, are doomed to fail in that they cannot compete with those larger game companies. A vicious circle of this stereotypical trend will inevitably continue if a comprehensive method for evaluating the game is not developed. At present, there are two commonly used evaluation systems, both of which have significant limitations. One is built into mainstream game platforms, such as the steam platform, which provides players with dichotomous choices. The player can "recommend" or "not recommend" the game as a means of expressing their overall attitude towards it. Another evaluation system of digital games mainly focuses on the game's ability to satisfy certain educational objectives, and thus particularly focuses on cognitive learning. In other words, games that appear to develop users' skills along the lines of problem-solving and critical thinking, for instance, tend to be rate higher. But many popular games lack a formal, objective, and comprehensive evaluation system that translates their value in terms other than just profit.

Overall, the potential for further research and the existing problems inherent in the growth of small-scale studios all point to the lack of rational and systematic evaluation mechanisms for digital games. And though YouTube creators offer relevant guidance for developing such evaluative criteria, more research on formal evaluation systems within the game industry is still required. In light of the well-established, but general, evaluation standards for digital games that currently exist, this research aims to explore the following three questions: How does one go about analyzing video games in general? Moreover, how does one evaluate a game that combines various modalities? Finally, how do different forms of expression of the game allow for an exploration of the connection and meaning between them?

2. Research Background

A multimodal perspective is constructive for the purposes of the present research in that it introduces the analysis of non-verbal expressions, such as images, to the study of the game [2]. Throughout the comparative research that has been conducted over the last few decades, the quality of theoretical studies operating from a multimodal perspective has been progressively improved. Indeed, multimodal discourses have increasingly been used to understand a variety of phenomena, such as film, television and social media. However, there remains little systematic elaboration on video games' meaning-making abilities or their communicative functions [3]. Moreover, game analysis and evaluation from the perspective of multimodal research, specifically, are rarely studied. Therefore, this research intends to use a multimodal perspective to explore different forms of expression for digital games, as well as the functions embedded in the different modalities, to study how the text and narrative of the game convey the ideological connotations it seeks to express.

"Multimodal" refers to the expression of phenomena via two or more modalities [4], and so encompasses modalities accompanied by expressions, actions, and other non-verbal forms, in addition to speech. Together, these symbols become social and cultural resources to represent meaning, abstractly constructing meaningful actions in various situational contexts through various means and symbols. In the last 10 years, multimodal analysis has transcended the boundaries of linguistics to extend into the fields of semiotics, philosophy, sociology, anthropology, and countless other disciplines [5].

Currently, there are three main branches of research that are subsumed under the broader heading of "multimodal analysis." One is a conceptual metaphor research perspective based on the principles

of cognitive linguistics [6] [7], and the other is a socio-semiotic perspective rooted in systemic functional linguistics [8], which proposes that language is a semiotic system for expressing meaning [9] [10]. Along with language, there are various non-linguistic symbols, such as symbols derived from images, sounds, emojis, actions, color, and physical space [11] [12].

But, the latest trends in multimodal research come from multimodal interaction analysis [12]. The objectives of multimodal interaction analysis and multimodal discourse analysis based on social semiotics are largely the same, with the exception that interaction analysis explores how different modalities work together to achieve a particular communicative goal [13] [14]. Combining the analytical modes of systemic functional linguistics and multimodal interaction theory, this research proposes a new comprehensive framework for multi-layered multimodal interaction analysis by deconstructing the game's various components. From this multimodal perspective, this research will systematically identify the core spiritual symbols expressed by each game, offer a diversified evaluation of the game under review, and provide corresponding suggestions for future game production in terms of both textual design and narrative content.

3. Analysis of Form

Working from the theoretical premises of multimodal interaction analysis as proposed by Scollon [15] [16] [17], the comprehensive analytical framework for the formal analysis of evaluation mechanisms for digital games is comprised of the following 6 research objects:

3.1. Actions - Hierarchies of Game Design

Within the context of multimodal analysis, actions may be classified as “high-level” or “low-level.” In general, the high-level actions [12] of a game refer to the goals and tasks in specific scenes within the game design, such as a party or a dinner in the case of *Harry Potter Magic Awakening*. Low-level actions are those modes of communication that collectively make up high-level actions [6], such as a gesture or an expression from the game characters. Low-level actions include the most diminutive elements of the game characters’ communication, such as spell content in *Harry Potter Magic Awakening*.

Actions within the game can be further refined depending on whether or not they are considered to be “frozen.” Frozen things will not change, such as the passerby walking around in the neighbourhood scene in *Detroit: Become Human*. Although his behaviour is not solidified, “functionally, he is solidified and will not change, so it is classified as a solidification activity” [18]. All actions are mediated and are usually multiply mediated. Actions come in frozen form, performed on a higher and lower level [18].

3.2. Level of Attention and Awareness - The Focus of the Game Design

In the process of multimodal interaction, high-level actions can also be categorized according to foreground, mid-ground and background high-level actions depending on the degree of attention given by the participants [19].

The primary focus of the player's attention is the foreground of the high-level activity, and in granting such attention, the attention allocated to the mid-ground and background actions decreases. The foreground in the game is the first shot the players pay attention to; the mid-ground high-level actions, by contrast, comprise the game scene we see, upon which the entire game’s content is built.

Unlike activity that happens in the fore- and mid-ground, background high-level activity includes the stylistic, artistic, and humanistic value attached to the game. It can include technical aspects of processor, graphic, and digital media technology, as well as artistic concerns pertaining to gameplay, map, and level design. The soundtracks for the specific atmosphere and feel of each scene and the

overall plot are also typically designated background high-level activities. Much attention goes into selecting such elements of the game; the individual designing the game's music, for example, needs to pay careful attention to whether the music matches the logic reflected in the character's dialogue, whether it is commensurate with the general plot, and the desired emotional effects of the music. Given this extreme attention to detail, the construction of these activity scenes is often the result of the painstaking efforts of several designers, who draw upon these elements to construct and convey the game's narrative.

Take, for example, the popular game *It Takes Two*, about the antagonistic couple, Cody and May, two humans turned into dolls by a magic spell. Trapped in a fantastical world, they are reluctantly challenged with saving their fractured relationship by the suave love guru Dr. Hakim. Various gleefully disruptive gameplay challenges stand between them and their return to normal life. The space where the game narrative transpires assumes the form of various warehouses, tree holes, gardens, and lofts. The construction of these activity scenes can be used as a sort of shorthand to advance the narrative.

In *It Takes Two*, different game scenarios and the corresponding gameplay in a two-player row failed to illustrate the common in-game background high-level activity organisation and music design. All game soundtracks were in the interest of the game itself. However, because of the ability of the game's music to shoulder the emotional expression of the music, game music, in addition to the emotional effect of the music, should be accounted for, paying particular attention to whether the music and the logic inherent in the character's dialogue and the plot match.

3.3. Means - The Cohesive Organisation of Game Design

Various means are used to organise and connect prospective high-level actions. The means of spanning and connecting between different activity scenes complete the narrative process that conveys the spiritual connotation of the game. In the game, means can be understood to refer to a level, trap, flashback, or animation transition used to complete or configure solid gameplay so that the game's interface can be switched out. The term could also be used to explain how the player switches scenes; a simple game means can be a slingshot in *Angry Birds*, which switches scenes and so allows the content of the game to progress. To a certain extent, the means can indicate the level of detail in the production of the game.

In order to experience a metaphorical merging of gameplay and narrative that pushes the boundaries of interactive storytelling, players can master unique and connected character abilities with every new level, such as bobsledding through a magical snow globe and parachuting a pair of underpants in *It Takes Two*.

All those means—jumping, zapping, swinging, riding, racing, throwing, shrinking, hacking, and slashing—are the cohesive elements that connect different game activities.

3.4. Site of Engagement- Specific Events in the Game Design

The site of engagement in multimodal interaction analysis refers to the specific place on the social timeline where the interaction between social practice and intermediary means enables social activities to be carried out. These are the actual windows of time created through the intersection of social practice and intermediary means, which allow high- or low-level activities to become the focus of participants' attention and further radiate to other participants and activities with a high or low degree of secondary attention [6].

For example, in Alipay's tree planting game, although this game is situated within a specific setting, through the player's daily watering and interaction with other users, the small saplings on their screen grow, with the user ultimately receiving a certificate congratulating them on protecting the

environment. This environmental certificate means that the player's watering contribution has allowed real saplings to grow offline, in the Ala Shan Desert. Although the platform handles the intermediate link facilitating this process, the environmental protection certificate does not have any actual authority, but it serves to make the player feel that they have truly participated in environmental protection. In this way, a simple example of an intervention point setting, just "watering" in this case, cannot attract a large number of players. Excluding the volume of the platform itself, this kind of game behaviour radiates to other users so as to reduce the distance between the virtual world and the real one, and this is the charm of the intervention point setting.

The social interactions offered by the game have allowed many people to survive the social isolation of the past few years. From this a linguistic point of view, when people are looking for online community building, social network games are bound to continue to grow. Online games open the door to the global gaming community, and users can communicate with each other in games or through multi-channel platforms, such as Discord. But, more in-real-life (IRL) gaming communities are likely to be born after the pandemic has eased because people will be more willing to leave the house. There are already many examples of this move to offline networking in San Francisco, where game nights centered around card games (like *Magic*), board games (like *Settlers of Catan*), and role-playing games (like *Mothership*) are prevalent in local game stores and cafes.

3.5. Modal Media Design Level

One or more modalities often embody actions [11], and in in-game design, different modalities focus on different game needs. The differences at the modal design level can be divided into four parts, namely modal types, modal configurations, modal density, and modal fineness.

3.5.1. Modal Type

Take the visual modal, which is the primary mode that most games focus on. The critical use of other modalities, such as tactile and audio modalities, in addition to last year's widespread Elden Ring fever and Nintendo's PS series game console update to the fifth generation, demonstrates the public's love of the multimodal organization of the game. The modalities used in future game design will inevitably be more than the combination of technologies such as XR, VR, and AR—they will ideally break through the constructed reality of the game world to ultimately attract greater numbers of people to adopt pursuits outside of reality.

3.5.2. Modal Configuration

Social action is generally performed via several modes, each of which plays a different role [18]. Therefore, it is necessary to select each modal according to the importance of its tasks. Determining the interrelationship between each modal and the corresponding weight value requires attention to the hierarchy system organizing each modal and its importance in the game.

3.5.3. Modal Density

Modal density affects the design content of the game. For example, in some secret room adventures and haunted house break-in games, the density of the auditory mode is crucial; different sound effects mean different plots. Only by mastering the proper auditory modal density can the designer ensure that the player stays fixated on the screen, or that they feel frightened in a tense atmosphere.

Secondly, understandings of modal complexity can be combined with those of the collective design of the above mentioned modes, with complexity being mainly reflected in the number of modal types. All modes constitute actions, but a change in a modality will not affect the sudden change of

action. The definition of modal density is also reflected in the number of modal configurations. In a game such as *It Takes Two*, the players who play the role of husband and wife can "double screen" on their PCs, thus realizing the plural presentation of the same set of modal configurations, which reduces the cost the player incurs (since players do not need to buy more game consoles or speakers to create their gaming environment) while innovating the expression of the game.

3.5.4. Modal Fineness

This study also introduces the concept of modal precision as relevant to the subsequent analyses through discussions of modal fineness. The fineness of the individual mode is particularly important for sensors. On the one hand, this fineness can be observed with the continuous evolution of game consoles and the perpetual updating of versions of the game's software. On the other, fineness in the game refers to the ways in which the game, as a virtual engine technology, simulates the actual social actions of the art form. For the user, the structural configuration of each mode, modal density, and modal complexity are readily apparent, because in real life, the configuration and complexity of each mode are much greater than that of the game, so the feeling brought by the modal precision in the game becomes more exciting. The more seamless the modal feels, the more it can give people the feeling that the game they are playing is "real," which, in turn, shortens the distance between virtual and reality and enables users to achieve a fully immersive experience.

3.6. Media

Any action featured within the game is an intermediary, and as such, needs to be implemented through various mid-grounds.

Completing all character actions in the game requires the user to perform actions in the real, offline world. For this reason, media can be divided into two categories - body media and non-body media, depending on whether or not it requires physical input from the user.

In the game, the player/user usually moves the character's body mid-ground so that the protagonist it represents is active; examples include the voice chat feature in *Honor of Kings* and the limbs that the player needs to move in *QQ Dance*. Non-body media are usually emitted by the content of the game design, and include elements like worldview, animation style of the surrounding environment, and background music, all of which serve to build a mid-ground for the action to occur.

4. Functional Analysis

Systematic functional linguistics not only proposes the research direction of multimodal discourse analysis, but also provides a more mature theoretical framework for it. In this framework, there is little difference between monomodal verbal and multimodal discourse communication at the contextual, communicative, and discourse levels. According to the framework proposed by Lim in the study of the multimodal discourse formed by the language of the picture domain, language and picture as communicative modes share four levels: ideology, genre, register, and discourse meaning. The use of multimodal game media is also applicable; in the field of games, multimodal function and significance evaluation indicators are mainly reflected in the following four levels:

4.1. Cultural Context Level

This level refers to the cultural context; that is, the ideology composed of people's mindsets, worldviews, habits, social rules, and the communicative procedures or structures that can specifically realize this ideology [2]. Without ideological context, however, there is no situational context. At this level, we can see the "shadow" of ideology that is composed of the cognitive constructs and inertia

of people and society from multiple sources of multimodal media, so that its "shadow" takes the specific "shape" of potential structures within this genre. Common structures of this sort include the stories, reports, explanations, and arguments that occur in daily life.

A genre with a highly structured and routine communicative structure can allow specific cultural content to carry a specific "shape" to hint at the specific "shadow" of culture in a purposeful and scripted manner. We collectively refer to ideology and genre as a "worldview" in the game. The choice of ideology affects the choice of genre, and the choice of genre reflects the content of the ideology. The two influence each other and jointly build the worldview for each game. Usually, the choice of ideology, which can be very grand or extremely mundane, will define the game's target players at a deeper level. For example, *QQ Farm* targets Chinese users with "idyllic" hobbies. The form is quite similar to the traditional Chinese smallholder economy. The game also capitalizes on the cultural manner in which most Chinese people comport themselves; they are often introverted and advocate self-sufficiency, independence, and solitude. It is presumably for this reason that planting different fields is the first task in this top-rated game, as this is something that one can do on their own. Social tasks, such as stealing vegetables, are also limited to known online friends. In large part because the game design is compatible with established cultural ideologies, the game itself may effectively be classified as a member of the leisure genre.

4.2. Situational Context Level

Situational context level means that in a specific context, the communication is constrained by various contextual factors, including the scope, tone, and style of discourse. In different situations, context will have an important impact on one's ability to both understand and express particular discourses: to a certain extent, it can improve the clarity of the discourse and thus help people communicate better. Context is the foundation of language in use, through which the processing of the selected register is also realized.

Since the context in the game is all human-created and human-controlled content, it can be regarded as the "register" of the individual creator of this part of the content, within this specific context. For example, in the mobile game *Carpenter's Wood*, all the instructions and hints are in the form of ancient Chinese words so as to implicitly praise the science behind and preservation of China's fifth-greatest invention: knowledge related to mortise and tenon. By using classical writing forms, the game's designers are effectively making people feel more connected to the ancient context of the game and conveying its cultural connotations.

In contrast to *Carpenter's Wood*, in the cute game *Animal Restaurant*, the words spoken by each animal character are more childish and endearing. Since this is a game about make-believe animals for a young audience, the register is influenced so as to include an unsophisticated vocabulary and operates on a low reading level.

4.3. The Level of Discourse Meaning

The level of discourse meaning is constituted by conceptual, interpersonal, and textual meaning [19].

4.3.1. Conceptual Meaning

For example, in comparing the abovementioned *Carpenter's Wood* and *Animal Restaurant*, the former is more cerebral and intellectual, and the latter is more cartoonish and emotional, meaning that they are, conceptually speaking, significantly divergent. Conceptual meaning defines the context to a certain extent and determines the range of language that may be used. It should be noted that conceptual meaning here refers more to the discursive meaning in the language field, the conceptual field and the language field in the game design content. The conceptual field, as mentioned earlier,

and the language field in the conceptual meaning are embodied in the independent language of some games, such as the spells in Harry Potter, the Middle Elf language in The Lord of the Rings, and the Dothraki language in Game of Thrones. The unique meaning contained in the game design can often make the spirit of the game come across as more special and unique, and enrich the overall connotation of the game to attract more players to this niche subculture. Ultimately, if a game is to make its conceptual meaning resonate with a great number of people, it must first attract a particular audience.

4.3.2. Interpersonal Meaning

Interpersonal meaning manifests more in the context of the relationship, and in discursive exchanges, especially in the game [14]. Games provide us with the opportunity to rethink and reshape social relationships. When constructing multimodal interpersonal meaning, text and image modalities complement each other in terms of interaction and attitude transfer. In the textual mode, this function can be displayed as a character, relationship, or behavioural expression; in the image mode, the specific embodiment is in the form of dressing up characters as a means of expression. The concept of the flexible use of interpersonal meaning has the potential to attract more players. *Kim Kardashian: Hollywood* is a very typical example. Although in terms of playability and word-of-mouth reputation, this game is often described as ordinary or even imperfect, players engage with the game to walk in the Kardashians' shoes, so to speak. Therefore, if a popular interpersonal meaning can impart such benefits to the game design, it will bring benefits in the short-term.

Indeed, the interpersonal meaning expressed by image modalities cannot be underestimated. This not only conveys the ingenuity of the game but, like all other artistic creations, highlights the mapping of the world values by the game designers themselves. For example, *Honor of Kings* is a game that has acquired thousands of users in China. The content is based on role-playing combat mobile games. However, in their selection of characters, male users usually choose historical heroes with solid combat power, such as Zhang Fei and Zhao Yun. The female characters available include Xi Shi, Da Ji, and other beauties renowned in ancient China. Although this limited selection could be chalked up to the consistent narrative habits of a patriarchal society, we can see that many game designers do not seem to notice anything wrong with this practice. In such a multidimensional role-playing game, we are treated to only two-dimensional characters; we still rarely see Athena not showing her breasts, for example, nor do we see famous women with advantages other than beauty, such as de Beauvoir.

4.3.3. Textual Meaning

Textual meanings reflect the concept of modal precision proposed earlier in this paper, and have the capacity to connect different concepts embedded in the discourse [20]. Articulation can not only form a discourse of conceptual and interpersonal meaning internally, but also connect discourse with context, so discourse articulation includes the articulation mechanism within the discourse, as well as the mechanism linking the discourse and the context. In this way, We can explore whether the expression design of the content and form of the game is compatible enough to become a complete and engaging system. This coherence concerning the modal precision of the game is crucial to examining whether the game is compelling enough to keep users engaged. Even a small button that one can push to "share" can be interpreted as "showing off."

4.4. Expression Level

Language is a medium composed of multiple modes, so the situational context naturally arises after limiting the game type and cultural level [21]. In the situational context, everything in the game design should be oriented around the culture, situation, and formal relationships between characters to

construct the spiritual content of the game. Only when the game is played by the user is everything then expressed. The final form of the game thus becomes the end expression as the designers intended it.

In game design, the expression level not only involves the meaning of the game character design, but it also affects the interactions between players. In *Glory of the King*, for example, one usually selects their own combat role, with players typically choosing characters of the same gender as themselves. Therefore, when the team has a heroic character like Li Bai or Lu Bu, many players will subconsciously think that the players behind the male characters are necessarily more experienced and will be more confident to start the battle. Nevertheless, if a familiar character by an unknown player happens to be Angela, Daji, or CAI Wenji, the stereotypes and hero discrimination caused by gender differences start to surface, and the attitudes towards the respective characters takes on a negative tone. Although, in a similar game, *Bubble Hall*, this is less often the case, because the player in this game is like the one in *Martin's Morning*, because the player in this game is like the one in *Martin's Morning* in that users are more focused on the tasks they are performing than on the characters they are enacting. In *Martin's Morning*, a story about a little boy who is transformed into fantastic creatures, like a werewolf, super hero, or dragon, gender becomes virtually irrelevant to the game design.

The meaning of different gameplay is also worth noting. For example, in *Double Line*, success only comes with cooperation. The player's participation in this linear adventure is aimed at facilitating the expression of the spirit of the game that the game designers have in mind, meaning that the plot is not the point of the journey. The most important thing is to make the player feel the two protagonists were originally estranged from each other, but by establishing trust through cooperation, the player's feelings in the final finale are characterized by warmth and sincerity.

Therefore, in the future development of the game's level of expression, like other media today, greater attention to the spiritual core of the game can attract more users and promote relationship-building through this medium. Following the lead of games like *Ikenfell* and *Spirit Swap*, which feature queer characters and characters of colour and of different sizes, in the future, with the continuous development of the game industry and the popularity of *Into the Mother Lands*, the inclusiveness of the independent role-playing game (RPG) sector will grow exponentially. At that time, it is likely that game design involving queer, disabled, feminist, and Jewish content will not stop with character design, but also show up in relevant, interactive features.

5. Conclusion

This study proposes a comprehensive evaluation framework for game design and expression from a multimodal perspective and constructs a paradigm through qualitative analysis. It shows that a multimodal framework can be used to evaluate the game's design and level of expression. This framework can also enrich the resources of the multimodal corpus through the analysis of existing games. Reliable and tested scales and established theoretical frameworks can be used to investigate the actual needs of different target audiences to achieve better game form design, media display, and transmission of spiritual connotation.

Through the use of multimodal analysis, it was found that the ideational and interpersonal meaning of the current design used in today's game market are insufficient in that they are characterized by language limitations and gender discrimination. In the future, more systematic linguistic surveys and paradigms can be carried out for individual cases, and evaluation questionnaires on the design and expression of related game forms can be designed through the different indicators mentioned in this article.

References

- [1] Diane • Carr, (2006). *Computer Games: Text, Narrative and Play*: Peking University press.
- [2] Zhang Delu. (2009). *Exploration of the comprehensive theoretical framework of multimodal discourse analysis* [J]. *A Chinese foreign language*, 6(01):24-30. [in Chinese]
- [3] Clara Fernández-Vara, (2019). *Introduction to Game Analysis*: Routledge.
- [4] Forceville, C. (1996). *Pictorial Metaphor in Advertising* [M]. London: Routledge.
- [5] Feng, D., Zhang Delu, Kay O'Halloran. (2014). *Progress in multimodal discourse analysis* [J]. *Contemporary linguistics*, 16 (1): 88-99126. [in Chinese]
- [6] Norris, S. (2011). *Modal Density and Modal Configurations: Multimodal Actions* [A]. In C.
- [7] Martin, J. (1992). *English Text: System and Structure* [M]. Amsterdam and Philadelphia: Benjamins,
- [8] Halliday, M.A.K. & Hasan, R.(1985) *Language, Context and Text* [M]. Victoria: Deakin University Press.
- [9] Halliday, &M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning* [M]. London: Arnold.
- [10] Jewitt (ed). (2009) *Routledge Handbook for Multimodal Discourse Analysis* C. London: Routledge.
- [11] Norris, S. (2007). *The micropolitics of personal national and ethnicity identity* [J]. *Discourse and Society*, (18): 653-674.
- [12] Norris, S.(2004). *Analyzing Multimodal Interaction: A Methodological Framework*[M]. London: Routledge.
- [13] Zhu Yongsheng. (2007). *Theoretical basis and research method of multimodal discourse analysis* [J]. *Journal of Foreign Languages*, (05):82-86. [in Chinese] DOI:10.16263/j.cnki.23-1071/h.2007.05.034.
- [14] Li Zhanzi. (2001). *Extension of the interpersonal meaning framework in functional grammar* [J]. *Foreign language research*, (1):48-54,80. [in Chinese]
- [15] Scollon. R. (1998). *Mediated Discourse as Social Interaction* [M]. London: Longman.
- [16] Scollon, R. (2001). *Mediated Discourse: The Nexus of Practices* [M]. London: Routledge.
- [17] Scollon. R. & R. Scollon. (2004). *Nexus Analysis: Discourse and the Emerging Internet* [M] London: Routledge.
- [18] Halliday, &M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning* [M]. London: Arnold.
- [19] Clara Fernández-Vara, (2019). *Introduction to Game Analysis*: Routledge.
- [20] Zhang Delu, Guo Enhua. (2013). *Dual perspectives of multimodal discourse analysis —— the connection and complementarity of social symbol view and conceptual metaphor* [J]. *Foreign Languages (Journal of Shanghai International Studies University)*, 36(03):20-28. [in Chinese]
- [21] Norris, S. (2002). *The implication of visual research for discourse analysis: Transcription beyond language* [J]. *Visual Communication*, (1):97-121.