International Conference on Humanities, Education and Economic Development (ICHEED 2026)

Cross-Media Translation of Inner Narrative from the Perspective of Narratology—A Comparative Study of the Original and Movie of the Girl with a Pearl Earring

Jingyi Guo
Faculty of Letters and Journalism, Xiangtan University, Xiangtan, Hunan, China

Abstract: Using narratology as a theoretical framework, this study makes a comparative analysis of the textual differences in inner narrative presentation between the original book "Girl with a Pearl Earring" and its film adaptation in 2003. From the analysis results, the original work is based on the first-person limited-knowledge perspective, and with the help of inner monologue and free indirect quotation, it realizes the multi-layer juxtaposition of the psychological level and the flexible adjustment of the narrative distance; When it comes to movies, video text methods such as behavior design, dialogue blank space, and scene narrative are used to build a psychological expression system, and the expression has changed from "direct narration" to "hint". The essential reason for these differences, in the final analysis, is the difference in media characteristics between language symbols and image symbols; The core of cross-media translation does not lie in anything else, but in reconstructing that psychological expression with the language of the new media itself. This study makes a systematic comparison from the five dimensions of narrative perspective, expression form, narrative level, narrative distance, and interpretation space, and provides a pure text research path for the cross-media adaptation of psychological narrative literature works.

Keywords: Narratology; Inner Narrative; Translation Across Media; Text Contrast; "Girl with A Pearl Earring"

1. Introduction

Inner narration is an important component of literary narration, which undertakes the core function of shaping characters and conveying themes. However, when literary works are adapted into movies, how to transform those delicate and subtle inner worlds from words to images has always been a difficult problem that has not been systematically resolved in theoretical discussions. George Bruxton once keenly pointed out that novels and films are completely different in dealing with internal life and external behavior, but since then, the focus of cross-media adaptation research has mostly focused on "visible" elements such as plot deletion, character image changes, and "fidelity" [7]. However, there is obviously insufficient attention to the hidden core mechanism of inner narrative. The Girl with a Pearl Earring provides a typical case of great value for discussing this problem: the original work is narrated from the first-person limited knowledge perspective of the maid Griet, and the skillful use of inner monologue and free indirect quotation makes it a literary text highly dependent on psychological depth; While the film adaptation in 2003 retains the core psychological context, it must not be unsuitable for the expressive logic of the image medium itself. Although the academic circles have discussed this work from the perspectives of feminism and psychoanalysis, they have not been able to systematically answer: What kind of written means did the original work use to build an inner narrative system? What kind of video text design does the film use to complete the translation of psychological expression? What are the core differences between the two in the context of plain text? Based on this, this study takes narratology as the theoretical framework, comprehensively utilizes Genet's narrative perspective theory, Mick Barr's narrative analysis and Shen Dan's identification method of free indirect speech, adopts the text close reading and cross-media comparison method, and conducts a systematic text comparison between the original text and the film image from the five dimensions of narrative perspective, expression form, narrative level, narrative distance and interpretation space,

trying to reveal the internal logic and literary evolution of inner narrative cross-media translation.

2. Review of Research Status at Home and Abroad

Judging from the domestic research situation, the exploration of "The Girl with a Pearl Earring" started relatively late. For example, Fu Shouxiang and Li Xin made an analysis from the perspective of feminism [1]. As for the research of the first person perspective, it basically only stays in the general theory of "can shorten the distance with the reader", and does not go deep into the level from pure text to construction mechanism. Looking at the adaptation of literature, film and television, most of the relevant achievements revolve around issues such as plot deletion and fidelity. For example, Cheng Xiangli used McLuhan's theory to discuss media differences, but in the end she still stayed at the comparison at the macro level. However, the research that combines "inner narrative" and "cross-media translation" is still in a blank state at present.

Let's look at the situation abroad. The study of novels over there is more detailed. Some scholars have specifically analyzed Griet's psychological activities, and some scholars will discuss novels with Vermeer's paintings and the social background of 17th-century Holland. None of the research has established a connection with the film adaptation. Cross-media narrative theory started relatively early in foreign countries, and its theoretical system is relatively mature. The achievements of people such as Stam are very important theoretical resources. Stam once proposed the "loyalty/betrayal" dichotomy replaced by "intertextuality" [2], but even so, those studies did not regard "inner narrative" as the core analysis object. Occasionally, there are some film studies from the perspective of psychoanalysis, but they also do not systematically compare with the inner narrative in the original text. Based on the above, this study is just based on this research gap.

Based on the above-mentioned situations, although domestic and foreign academic circles have accumulated a certain amount of research on "The Girl with a Pearl Earring", the kind of research that combines "inner narrative" and "cross-media translation" and systematically compares the inner narrative of the original text with the psychological expression of film images is still blank. And this is exactly the core foothold of this research.

3. A Contrastive Analysis of Inner Narration in Original Works and Movies

By juxtaposing the original text narrative and the film video narrative, the author makes a comparative analysis from five dimensions: narrative perspective, means of expression, narrative level, narrative distance, and interpretation space, and reveals the text evolution law of inner narrative in cross-media translation.

3.1 Narrative Perspective: Limited Perspective vs. Quasi-Internal Focus

The Girl with a Pearl Earring uses Griet's first-person "I" as a narrative, which is a typical inner-focused narrative. Genet summarizes the characteristics of inner-focused narration as "the narrator only says what a certain character knows"-the reader can only see through Griet's eyes and feel the story world through her heart [3]. The "limitation" of this perspective just constitutes the advantage of inner narrative: it naturally brings the reader into the subjective world of the characters, and makes the psychological distance compressed to the minimum. Shen Dan further pointed out that there is a tension between "narrative self" and "self" in the first-person narrative, the former being the narrator of retrospect after the fact, and the latter being the character experiencing the event at that time [4]. In the original work, this tension is particularly obvious: Griet is not only the self who experienced the event at that time, but also the self who reflected afterwards. The coexistence of the double self provides the inner narrative with a hierarchy in the time dimension. Mick Barr distinguishes between "focusser" and "narrator", and in the original work, the two highly overlap with Griet alone, which makes the whole narrative show strong subjectivity and psychological depth [5].

Movies cannot replicate the first-person perspective of text, but through narrative design, "quasi-internal focus" at the image level is achieved. The film sets Griet as the absolute core narrative subject-the plot always revolves around her experience, and the camera always follows her action trajectory, so the audience shares roughly the same time and space vision with her. This kind of "follow-up narrative" is

functionally close to the limited-knowledge perspective of the original work. However, the "inner focus" of the image is fundamentally different from the first person of the text: the text can directly present the inner voice through the word "I", while the image can only present the external world through the perspective of "she". Cheng Xiangli sums up this difference brilliantly: "Film is a figurative art, which can only present what is visible, while literature can present the invisible mind [6]." George Bruxton also points out: "The strength of film lies in presenting visible movements, while the unique power of literature lies in revealing the invisible flow of consciousness [7]." The film Girl with a Pearl Earring precisely does not use any narration, which means that all inner information can only be transmitted indirectly through externalizable narrative elements.

3.2 Means of Expression: Text Direct Narrative vs. Video Hint

There are four main ways to present Griet's inner world in the original work. One is the direct inner monologue, that is, to directly express the feelings and desires of the characters with leading words such as "I think", "I feel" and "I am afraid". For example, in Chapter 4, "I want to tell him that I don't care about the onions, what I care about is what he looks at me. But I didn't say anything"-this sentence presents the great contrast between the real thoughts of the heart and the external words directly in front of the reader. Shen Dan calls this technique "direct psychological quotation", and its function is to "enable the reader to directly enter the inner world of the character and experience it immersively [4]". The second is free indirect speech, which is between direct speech and indirect speech. Although the narrator tells the characters' hearts in the third person, he retains the characters' own language habits and emotional colors. For example, in Chapter 5, "She should leave the room. But she didn't move"-from the perspective of grammar, it seems that the narrator is speaking, and from the perspective of semantics and emotional color, it is clear that Griet herself is urging herself. Shen Dan pointed out that this technique "can present the characters' hearts without interrupting the narrative flow, while at the same time having a certain degree of narrative distance" [4]. The third is the psychological projection of sensory details, that is, to reflect the inner state through the description of external sensory details-the changes in light, color, and smell perception are not simple descriptions, but external projections of psychology. The smell description of "sweet and strange, like a honey cake baked in the sun" in the ear piercing scene hints at Griet's faint sense of expectation through the association of "sweetness". The fourth is the echoing interpretation of behavior and psychology. The original works often follow the psychological interpretation after the behavior description, forming a structure of "behavior psychology"-after "her movement of rubbing tiles is getting slower and slower", she must use her inner monologue to explain "Why slow down" to achieve internal and external linkage presentation.

Movies cannot achieve any of the above-mentioned direct psychological presentation, and can only rely on the narrative design of the image text to complete indirect translation. Its core means include: behavior design instead of inner monologue, such as Vermeer's behavior of piercing Griet's ears in the ear piercing scene itself is a kind of psychological externalization-she delivers physical pain to him, this choice It already contains multiple meanings such as trust, desire and dedication; Dialogue blank space and restraint, the movie dialogue is simplified to the extreme, Griet has almost no lines in a large number of scenes, this kind of "silence" itself becomes a means of psychological expression-when a person has something in his heart but does not say it, Her silence can convey inner depression and restraint better than any words; Plot trade-offs and reconstruction, such as ear piercing and earring wearing in two acts in the original work (Griet pierced her ears alone in the attic, Vermeer assisted in wearing earrings in the studio) were merged into a continuous scene, and all of them were completed by Vermeer. Although it sacrificed Griet's subjectivity to bear the pain alone, it strengthened the intimacy and sense of ritual between the two; The psychological mapping of the scene narration, the transition between the attic (Griet's private space) and the studio (Vermeer's space), constitute the visual metaphor of the psychological state in itself.

3.3 Narrative Layers: Multilayer Juxtaposition vs. Single Line Focus

In terms of narrative levels, the original work

can simultaneously present multiple psychological levels such as physiology, emotion, desire, and cognition, while the film is limited by the media and can only focus on the main levels that can be externalized.

Taking the scenes of piercing ears and wearing earrings as an example, the psychological levels presented at the same time in the original work include at least the physical level ("a fire-like pain pierces my body"-pain is not only a physical reaction, but also an emotional metaphor, implying "piercing" and intrusion, pain and intimacy are inseparable), the emotional level ("My jaw is stiff, but I still try to hold my head up"-shyness and restraint are intertwined; the erasure of tears becomes an opportunity for intimacy) and the desire and cognition layer ("I just keep recalling his fingers in my mind"-although the touch is over, it is played back infinitely in my heart; "I want to know what he is thinking"-pointing to the narrative boundary, and also pointing to Griet's inability to enter These three levels are presented at the same time without any obstacles, forming a complete psychological picture.

The film can present the physical layer (the audience can see painful frowns and tears) and the emotional layer (nervous expression, shy head bow, tearful stare), but the desire and cognition layer-those most secretive, most complex, and most Deep psychological content-can hardly be effectively presented. The audience can feel that Griet "has feelings for that person", but they have no way of knowing whether this feeling is love, gratitude, attachment, reluctance, or dependence after fear-these subtle distinctions are reduced in the film. Simplified to the overall "affection". Genet put it brilliantly: "The greatest advantage of written narrative is that it can penetrate infinitely into the bottom layer of the character's consciousness, which is beyond the reach of any other medium. [3]"

3.4 Narrative Distance: Flexible Adjustable vs. Relatively Fixed

In the original work, the narrative distance can be flexibly slid between "getting into the characters' hearts" and "exiting external observation". "Just recalling his fingers in my mind" is extremely close-the reader "dives" directly into Griet's thoughts, in sync with her inner flow. "He looked at me but didn't draw, and I wondered what he was thinking"-this sentence is both extremely close (she is telling the reader what she thinks) and at the same time opens up a subtle space: she doesn't know what Vermeer is thinking, and the reader doesn't know, and the reader is in the same "don't know" state as Griet. The original book can freely switch between different distances, sometimes allowing readers to integrate with her, and sometimes allowing readers to watch with her.

The narrative distance of the film is relatively fixed. Regardless of the painful or intimate moments, the audience is always standing "outside" to watch Griet's expression and Vermeer's actions. Close-ups can bring physical distance closer-see the tears in her eyes, the trajectory of his fingers-but it's still "watching" rather than "entering". A closer physical distance is not the same as a closer psychological distance: the audience can see things more clearly, but still cannot enter her inner world. Those tearful eyes can express "emotion", but they cannot tell the audience whether the emotion is pain, happiness, shyness or reluctance. The film's narrative distance is basically a relatively fixed "bystander" position. George Bruxton summed it up brilliantly: "Fiction combines the reader with the character, while film distances the audience from the character. [7]"

3.5 Interpretation Space: Open Blank vs. Closed Guide

The blank space in the original work is "deliberately not telling the truth, but implying the direction". "I've always wanted to wear pearl earrings"-why? Since when? Has anything to do with Vermeer? The original book does not answer directly, but provides ample clues for readers to imagine in a specific direction. "I wonder what he's thinking"-she doesn't know, nor does the reader, the narrative leaves an eternal void here. But this kind of blank space is directional: it leads the reader to think about the unknowability of the relationship between characters, and the inability to really enter the hearts of others. The original clearly distinguishes what to say from what not to say, and the unsaid parts are often more worth thinking about.

The interpretation space of the film presents another form. Due to the absence of narration and inner monologue, a large amount of Griet's psychological content is in a state of "absence"

in the film, rather than "conscious blank space". Viewers can "guess" her psychology from her expressions and movements, but different viewers may draw very different conclusions-is that love or dependence? Happiness or sadness? This polysemy is an advantage in some cases (greater openness), but it becomes a limitation when it is necessary to accurately convey complex psychology. In order to make up for this limitation, the film uses lens language, light and shadow design, music, performance rhythm and other means to "guide" the audience's understanding-it tells the audience that "this moment is important" and "this is a kind of affection", but does not give specific psychological content. This is essentially a "closed guidance": the boundaries of meaning have been set (the viewer knows this is an important emotional moment), but the concrete content of meaning remains unfilled. Dai Tong and Lei aoyan pointed out when analyzing the unconscious psychology of the film that "the process of wearing the pearl earrings is processed as a ritualized intimacy, and the language of the camera continuously strengthens the audience's imagination of the relationship between the two" --This kind of "strengthening" is the embodiment of closed guidance [8]. Stam also pointed out: "The visual presentation of a movie is always more" decisive "than the text. It gives the reader the freedom of imagination, but at the same time provides a stronger emotional impact. [2]"

3.6 Comprehensive Comparison(see Table 1)

Table 1. A Comparison of Core Dimensions in Cross-media Translation of Inner Narrative

Dimension	Original (text)	Movies (video text)	Difference essence
narrative perspective	In the first-person limited-knowledge perspective (inner focus), readers can only perceive the world through Gerrit's eyes and hearts, and the narrative self and self-coexist, forming a hierarchy in the time dimension	Quasi-internal focus (follow-up narrative), the lens always revolves around Gerrit, the audience shares the vision of time and space with her, but always takes "she" as the object of observation	Words can say "my" heart, images can only present "her" eyes; The former directly presents the internal voice, while the latter can only be conveyed indirectly through the external perspective
Expression morphology	Speak directly about inner activities, sensory experiences, memories, questions	Hint: Transmitting psychology through behavior, expression, eyes, touch	Words can directly "speak your heart", images can only "use behavior to let you guess"
Narrative level	Multi-layer juxtaposition: physiological pain + immediate emotions + cravings + doubts about others, presented at the same time	Single-line focus: mainly presents externalizable physiological reactions and emotions, and deep desires are blurred	Text can present multiple layers of psychology at the same time, and images can usually only focus on the first two layers that can be "externalized"
Narrative distance	Flexible and adjustable: can drill into the heart, can withdraw from observation, the last sentence is confused with the reader	Relatively fixed: watching the whole process, the audience can only "speculate" their hearts through their expressions	Words can allow readers to "enter" characters, while images can only allow viewers to "watch" characters
Interpretation space	Open blank space: open with direction-the clues are sufficient, but the answers are not filled	Closed guidance: use lens, atmosphere, and rhythm to guide the audience to pay attention to "important moments", but the inner content is still unknown	The blank space of the text is "intentionally not saying but hinting at the direction", and the blank space of the image is "can't say but it's important to tell you"

Based on the first-person limited knowledge perspective, the original work constructs a multi-layer juxtaposition, flexible distance, and open inner narrative system with the help of direct inner monologue, free indirect quotation and other means. The film completes the translation from "direct narration" to "hint" through behavior design, dialogue blank space,

and scene narrative. The root of the difference between the two lies in the characteristics of the media: the abstraction of the text enables it to directly express the heart, while the concreteness of the image makes it only indirectly imply the psychology through external presentation. This difference is particularly evident in the scenes of piercing ears and wearing earrings-when reading the original book, readers follow Griet through pain, touch and doubt; When watching the movie, the audience can only stare at the silent intimate ceremony, feeling the emotion surging between eyes and touches, but they can never enter her heart at the last moment. The former was a question that had penetrated into her heart, and the latter was a guess that she had made while staring at her expression.

4. Conclusion

Based on the theory of narratology, this study makes a systematic comparative analysis of the differences in inner narrative presentation between the original work and the film from five dimensions: narrative perspective, means of expression, narrative level, narrative distance, and interpretation space, and draws the following main conclusions.

The main findings of this paper are as follows: First, based on the first-person limited knowledge perspective, with the help of direct inner monologue, free indirect quotation, psychological projection of sensory details, and echoing interpretation of behavior and psychology, the original work constructs a set of multi-layer juxtaposition (simultaneous presentation of physiology, emotion, desire, cognition, etc.), flexible and adjustable distance (can slide freely between immersive vipassana and external observation), and the blank space has directional (intentionally not speaking but implying the path) inner narrative system. Second, the film is limited by the image medium, and instead translates the inner narrative into indirect hints through image text means such as behavior design, dialogue blank space, plot reorganization, and psychological mapping of scene narrative, so as to compress it into a single-line focus at the narrative level, solidify it into a relatively fixed onlooker position in the narrative distance, and form a closed guidance in the interpretation space (marking the importance of emotion but suspending the specific connotation). Thirdly, the essence of cross-media translation is not faithful reproduction, but adaptive reconstruction: linguistic symbols can directly express the heart, while image symbols can only indirectly refer to the mind through external representations.

This study provides an analytical path based on pure text narratology for the cross-media adaptation of psychological narrative literature works, and at the same time deepens the understanding of the essential differences between text narrative and image narrative in presenting internality. The five-dimensional comparison framework constructed can be used as a reference tool for subsequent similar studies.

The limitation of this study is that it only focuses on a single text case, and fails to extend the analytical framework to other cross-media adaptations of similar inner narrative works for verification and comparison. Therefore, the generalizability of the conclusions needs to be further tested. Follow-up research can be carried out from the following aspects: First, expand the scope of the case, and apply the five-dimensional analytical framework of this study to the adaptation and analysis of similar inner narrative works such as “The Hours” and “Reminiscence of the Lost Years” to test its applicability; Second, from the perspective of gender, it examines the special laws of female first-person inner narratives in cross-media translation; Third, the introduction of audience acceptance research to explore the cognitive mechanism and empathy differences between readers and audiences on the characters’ psychology in different media.

References

[1] Fu Shouxiang, Li Xin. Artistic Communication and Female Light Spreaded Across Media--A Feminist Analysis of “The Girl with a Pearl Earring” [J]. Women’s Studies Series, 2011 (6): 71-76.

[2] Robert Stam. Literature and Film: A Guide to Film Adaptation Theory and Practice [M]. Beijing: Peking University Press, 2005.

[3] Gerard Genet. Narrative Discourse · New Narrative Discourse [M]. Translated by Wang Wenrong. Beijing: China Social Sciences Press, 1990.

[4] Shen Dan. Research on Narratology and Fiction Stylistics [M]. 4th Edition. Beijing: Peking University Press, 2019.

[5] Mick Barr. Narratology: An Introduction to Narrative Theory [M]. Translated by Tan

Junqiang. Beijing: China Social Sciences Press, 2003.
[6] Cheng Xiangli. A Perspective on the Relationship between Film and Literature in Media Transformation [D]. Urumqi: Xinjiang University, 2004.
[7] George Bruxton. Novels to Films [M]. Translated by Gao Junqian. Beijing: China Film Publishing House, 1981.
[8] Dai Tong, Lei Xiaoyan. The Unconscious Psychology of the Movie "Girl with a Pearl Earring" --Based on the Perspective of Psychoanalysis [J]. Western Journal, 2021 (13): 152-154.