

Multimodal Analysis: Exploring Character Portrayal in Harry Potter and the Philosopher's Stonethrough Visual Elements and Subtitle Translation—A Case Study of Harry Potter and Hermione Granger

Yejie Xu

School of Foreign Languages,
East China University of Political
Science and Law, Shanghai, 201620,
China

Corresponding author:
23060601079@ecupl.edu.cn

Abstract:

Against the backdrop of global cross-cultural communication research on film and television, studies on character portrayal in the Harry Potter film series have long been constrained by limitations. To transcend the traditional unimodal research paradigm, this study adopts a qualitative case analysis approach, focusing on two representative scenes in *The Philosopher's Stone*—the protagonists' encounter with the mountain troll in the bathroom and their perilous entanglement with the Devil's Snare—to systematically investigate the synergistic mechanisms between the visual and subtitle translation modalities. The findings reveal: visual elements dominate the externalization of character traits; subtitle translation reinforces the expression of personality. Ultimately, the intertextuality of the two modalities significantly enhances character distinctiveness. This study confirms that multimodal synergy is fundamental to cross-cultural character portrayal in film and television. The established "picture-subtitle" dynamic adaptation model not only provides a new paradigm for subtitle translation practice but also, through the mirroring contrast between Harry and Hermione, reveals a universal pathway for differentiated character construction. Future research could extend to other dual-protagonist pairings and explore the potential of artificial intelligence in optimizing multimodal translation.

Keywords: Multimodal Analysis, Subtitle Translation, Character Portrayal, Visual Grammar, Harry Potter

1 Introduction

In the context of globalization, film and television works have become core vehicles for cross-cultural communication. The Harry Potter film series, with its global influence, serves as a quintessential sample for studying character portrayal in cross-cultural communication. Fantasy literature adaptations like the Harry Potter series construct character images through multimodal narratives, shaping global audiences' perceptions. However, due to language barriers and cultural differences, most non-English-speaking audiences primarily understand the plot through subtitle translation. The potential loss of cultural metaphors and linguistic styles during translation can lead to the misinterpretation of character personalities.

Film visual elements, as a signifying system independent of text, co-shape characters alongside translated text. Wei, Qiu, and Wang argue that the visual symbolic system in films (including costumes, settings, props, and cinematography) possesses deep cultural narrative functions beyond surface aesthetics [1]. These meticulously designed visual elements, through symbolic representation, not only construct specific spatiotemporal coordinates but also establish channels for cultural immersion at the cognitive level. Filmmakers encode visual vocabulary, enabling audiences to transcend real-world boundaries and experience the unique charm of heterogeneous cultures [1]. You point out that the film's visual styling system and composition jointly shape the audience's primary aesthetic cognition, serving as the core carrier for the cross-cultural dissemination of cinematic art [2]. These visual elements are not only the film's "first language" but also key artistic codes constructing the foundation of audience perception [2].

In the field of subtitle translation, current research focuses on obstacles to overcome, the priorities and purposes of subtitle translation, the guiding role of Skopos theory, and the multimodal theoretical frameworks to follow. For example, Liu claimed that there are numerous challenges in subtitle translation. These include the preservation of language styles, the conversion of cultural images in various contexts, the compression of information under time and space limits, and the disparities in comprehension brought about by the audience's diverse cultural backgrounds [3]. Hou emphasized the importance of timely, clear, and succinct movie subtitles. The goal is to allow viewers to rapidly understand the main points in the little amount of time allotted for viewing. [4]

Ke thinks that the three teleological concepts can effectively direct the subtitle translation process. Particularly when addressing the alteration of cultural representations [5]. Making ensuring the translated subtitles allow the intended audience to comprehend the movie is the main goal

[5]. Su emphasizes that film subtitle translation should adhere to a multimodal theoretical framework, precisely conveying the multidimensional information of the source language—including semantic connotation, cultural load, and emotional tension—under spatiotemporal constraints [6]. Through the dynamic coordination of subtitle text with visual and auditory elements, it optimizes the reception effect for the target audience and achieves cross-cultural resonance in cinematic artistic expression [6].

Domestic analyses often isolate discussions on subtitle translation strategies or cinematographic language usage [7]. Although foreign theories emphasize multimodal synergy, they lack empirical validation specifically for Harry Potter and the Philosopher's Stone, resulting in three major issues: the disconnection between visual elements and subtitle texts, the absence of comparative analysis between Harry and Hermione's personalities, and the gap in adaptation mechanisms for Chinese subtitles. Harry and Hermione, as core characters, embody the archetypal paradigms of the "trauma hero" and the "rational mastermind" respectively. Their contrasting personalities are not only crucial for constructing narrative tension but also a key window for cross-cultural audiences to understand Western adolescent coming-of-age narratives. Yet, few studies combine the two. This issue highlights the necessity of multimodal analysis in film translation and character interpretation. Without a systematic examination of their synergistic shaping mechanisms, it is difficult to explain the deep-seated reasons for the characters' global resonance or provide reusable character modeling pathways for cross-cultural film production.

Therefore, this study innovatively examines visual symbols and subtitle translation as a synergistic system, focusing on their personality differences. Through key scenes, it contrasts the differentiated portrayal of Harry and Hermione and reveals how intensifiers in subtitle translation intertextually interact with the presentation of visual crisis. The research objective of this study is to decode the operational logic through which the "picture-subtitle" combination achieves character psychology externalization in specific clips. It aims to establish a reusable analytical model for film and television characters and provide a new pathway for multimodal research. Based on the above objectives, this paper proposes two core research questions: In typical scenes of Harry Potter and the Philosopher's Stone, how do visual elements and subtitle translation dynamically and intertextually synergize to externalize the core personality traits of Harry Potter and Hermione Granger? Does the cognitive reinforcement effect of bimodal synergy on the differentiated character portrayal of the two protagonists show significant differences?

2 Research methodology

This paper adopts the case analysis method within qualitative research. Case studies effectively enhance the clarity of theoretical guidance, improve the fit and relevance between presupposed theories and research phenomena, and allow theoretical constructs to be more fully reflected in data analysis. Simultaneously, this method significantly increases the rigor of data analysis and optimizes the match between research questions and case study methodology [8].

Key scenes were selected through a three-stage process: Firstly, an initial screening of narrative nodes. Based on the *Philosopher's Stone* novel and screenplay, extract all high-conflict scenes involving the externalization of Harry and Hermione's personalities. Secondly, evaluating modal richness. Exclude scenes dominated by dialogue or with singular visuals, retaining fragments simultaneously featuring significant visual symbols (e.g., lighting, composition, body language) and contentious points in subtitle translation. Thirdly, validation of character contrast validity. Ultimately selected two fragments: the protagonists' encounter with the mountain troll in the bathroom and their entanglement with the Devil's Snare. Both scenes feature intensive "visual-subtitle" modal interaction, maximizing the need for analyzing the personality contrast between the two protagonists.

In summary, this study explores two scenes from the film *Harry Potter and the Philosopher's Stone* from the perspectives of subtitle translation and visual elements: the protagonists' encounter with the mountain troll in the bathroom, and the protagonists' accidental peril in the Devil's Snare, investigating the role of subtitle translation and visual elements in character portrayal.

The analytical framework integrates film semiotics and multimodal discourse analysis theory: For the visual modality, Kress & van Leeuwen's Visual Grammar theory is applied, focusing on two dimensions: Interactive Meaning (composition, perspective) and Compositional Meaning (color, lighting, action sequences), as they systematically deconstruct the externalization mechanisms of character psychology within the imagery [9]. For the subtitle translation modality, the focus is on semantic compression strategies, intensifiers (e.g., modal particles, emphatic adverbs), and the treatment of cultural items, as these directly impact the cross-cultural transmission of character linguistic style. These dimensions are already widely used in film character studies. This study accordingly constructs a "visual-subtitle" synergy model to avoid unfocused conclusions due to overly broad analytical dimensions.

3 Research findings

3.1 The protagonists' encounter with the mountain troll in the bathroom

3.1.1 Visual modality

Through the action scenes, lighting symbolism, and environmental design, the visual modality in the bathroom huge scene communicates the characteristics of the actors. Tension in the story is produced by the cramped space's uncomfortable design. With the giant's enormous form taking up over two-thirds of the frame, the small bathroom setting heightens the sense of urgency and danger due to the sharp spatial contrast. Hermione's powerlessness in the face of such a power imbalance is strongly illustrated by the contrast between her horrified countenance as she looks up and the giant's downward glance.

The characters' characteristics are depicted effectively through the series of character movements. Harry's dynamic sequence of activities, including "hugging the giant → leaping on its shoulder → poking his wand into its nose" were primarily taken from his point of view, using tracking shots. This demonstrates his courage and agility. Hermione's body language and dilated pupils reveal her "curling up → crawling → hiding" body motions. This illustrates her panic and suggests that she will eventually guide Ron logically. This creates a character development subplot.

Setting the mood involves the use of metaphorical lighting. The first icy gray tones imply hopelessness. A dim, warm glow appears from Harry's wand as he attacks, signifying the intervention of hope. The warm light emerges when Ron casts a spell, dramatically changing the mood. It draws attention to the characters' moment of triumph.

3.1.2 Subtitle Translation Modality

By modifying imperative sentences and changing modal particles, this subtitle translation method accurately depicts the contrasts between characters. "赫敏, 快跑!" is the translation of Harry's phrase, "Hermione, move!" This structure's strong short phrase emphasizes the sense of urgency by omitting the subject and predicate. In a Chinese emergency, "快跑" (move!) is a better choice than a literal translation of "移动", emphasizing Harry's prompt action in defending his friend.

Character development is also substantially improved by the way interjections are handled. "Help! Help!" Hermione uses repetition to heighten a girl's fear of danger while she is by herself. The slang translation of Ron's "Hey, pea-brain!" was "嘿! 大傻瓜." It is consistent with his nature and maintains his witty snark. "快想办

法!” is the translation of Harry’s “Do something!” The word “quickly” makes him feel even more compelled to save his friend.

The translation of magical terminology is also adept. The incantation “Wingardium Leviosa” retains its phonetic translation “羽加迪姆勒维奥萨” and adds a gesture description—“先一挥再一抖”(swish and flick)—cleverly resolving cultural barriers through form-meaning complementarity while consolidating Hermione’s image as the knowledgeable “know-it-all.”

3.1.3 Character portrayal through the two modalities

The dynamic synergy between the visual and subtitle modalities significantly enhances the characters’ core traits. When the troll grabs Harry’s ankle and dangles him upside down, the rhythmic delivery of the translated subtitles intertextually aligns with the visuals of his dangling struggle, deepening the audience’s impression of Harry’s trait of “relying on partners.” The character differences between Harry and Hermione are reinforced through modal complementarity: Hermione’s shriek “Help!” combined with cowering actions embodies a pure instinctive reaction in crisis; Harry’s action-oriented “快跑!” paired with his attacking movements reflects his instinctive response under pressure. This “visual-subtitle” mirroring in crisis not only effectively illustrates the character development process but also validates the reinforcing effect of bimodal synergy on character portrayal.

Synthesizing the portrayal through both modalities, Hermione Granger’s character is revealed as follows: Visually, the stark contrast between her initial extreme panic and instinctive survival reaction upon encountering the troll and her subsequent frantic gesturing of the spell under the sink foreshadows the core traits ultimately revealed. In subtitles, the repetitive scream “救命救命”(Help! Help!) powerfully amplifies her panic. More crucially, the precise supplementary description of the spell “羽加迪姆勒维奥萨”(Wingardium Leviosa) and its execution point “先一挥再一抖”(swish and flick) accurately depicts Hermione’s “emergency wisdom” and “know-it-all” essence. Even in extreme panic, her profound knowledge base and rational thinking can be rapidly accessed to guide her companions, revealing a resilient rational core beneath the fear.

Similarly, for Harry Potter, visually, a series of high-risk, high-difficulty continuous offensive actions, combined with rapid editing and tracking shots, concretize his extraordinary courage and agile instincts—fearless, bold yet meticulous, and decisive. In subtitles, concise, powerful, highly action-oriented commands and urgent calls for help highlight his strong sense of responsibility to protect his companions and his eagerness to find solutions in crisis.

Synergistically, they reinforce Harry’s core personality as an “action-oriented leader” and “fearless protector.” Simultaneously, the visual of his dangling struggle and the subtitle of his urgent plea synchronously reveal his underlying trust in and dependence on his friends beneath his courage.

3.2 The protagonists’ accidental Peril in the Devil’s Snare

3.2.1 Visual modality

From a visual modality perspective, the cinematography establishes a dark, eerie environmental atmosphere, setting the mysterious and dangerous tone for the trio’s predicament. This led to a significant increase in tension. It also brought attention to how dire their circumstances were. The characters were surrounded by the tangled vines of the devil’s net, creating a gloomy visual impact. It effectively symbolized the perils in the environment and communicated the present threat.

Take note of the characters’ facial expressions and body language. Harry’s postures and facial expressions made it clear that he was tense and uncomfortable while he was caught in the devil’s net. Hermione stood in stark contrast to this. “You have to relax. If you don’t, it will only kill you faster,” Hermione remarked at the pivotal moment. Her serious attitude and calm manner clearly showed her calm, logical, and quick-thinking personality features.

One important visual symbol is lighting. It is crucial to expression and storytelling. Light and shade are strongly contrasted by the comparison between localized feeble light and huge darkness. This draws attention to the characters’ hardships and heightens the unpleasant atmosphere. Crucially, Hermione’s observation that “Devil’s Snare hates sunlight” highlights the revealing semantic role of light, whether it be used to symbolically indicate something or to imply an actual source of light. It foreshadows the plot twist that follows.

3.2.2 Subtitle translation modality

Furthermore, the subtitle translation modality in this segment not only accurately conveys the original meaning but also plays a significant role in character portrayal.

“Stop moving, both of you. This is Devil’s Snare.” is translated as “别乱动，都别动，这是魔鬼网。”The translation is concise, forceful, and imperative, effectively conveying Hermione’s decisiveness and authority in the crisis.

“You have to relax. If you don’t, it will only kill you faster.” is translated as “你们必须放松，不放松，你们就会死得更快。”“必须”(have to) strengthens the imperative nature. The blunt phrasing of “不放松”(if you don’t)

and “死得更快” (it will only kill you faster) accurately reproduces the extreme peril and Hermione’s grave warning while amplifying the tense atmosphere.

Regarding terminology, “Devil’s Snare” is translated as “魔鬼网”. While there is no direct equivalent in Chinese, this translation is readily understandable, ensuring immediate audience comprehension of the plot while retaining the mystique appropriate for a magical creature.

The subtitle translation successfully conveys character tone: Hermione’s translated lines reflect her calmness, authority, and intellectual confidence; Harry’s translated lines correspondingly exhibit tension and dependence.

3.2.3 Character portrayal through the two modalities

Focusing on Hermione Granger, visually, her composed and serious expression, focused gaze, and clear-voiced commands during the crisis visually present her remarkable calmness, composure, and level-headed resourcefulness beyond her years. In subtitles, the translation reinforces and amplifies her leadership within the team, her extensive knowledge base, and her rationally-based decisiveness through firm, imperative diction like “必须” (have to), “别动” (stop moving), and clear, knowledge-based instructions. Her rapid recall of Herbology knowledge to direct Harry is served precisely by the translation, shaping her “mastermind” role.

Shifting to the portrayal of Harry Potter in the same scenario, visually, his frightened expression and struggling movements while entangled by the vines directly externalize his nervousness, fear, and sense of helplessness facing an unknown magical threat. In subtitles, the translation of his lines, such as possible pleas or flustered responses, effectively conveys his anxiety and immediate reliance on Hermione’s guidance, highlighting his relatively passive and support-needing state in this specific situation.

Cao emphasizes that subtitle translation often works in concert with visual elements, complementing each other, thus making the original meaning come across more vividly and clearly [10]. Through multimodal analysis of visual elements and subtitle translation, the intricate design of character portrayal in *Harry Potter and the Philosopher’s Stone* becomes clear. The visual modality intuitively constructs the crisis scenario and externalizes characters’ internal states through environment, lighting, expressions, and actions. The subtitle translation modality deepens and complements the linguistic expression of character traits through precise conversion of tone, semantics, and terminology. The two modalities work synergistically within the narrative, successfully shaping Hermione’s image as the rational, calm, knowledgeable “rescuer” in this scene, and Harry’s image as the tense, dependent, instinct-driven “victim,” vividly showcasing the protagonists’ contrast-

ing personality facets in crisis response and laying the groundwork for subsequent character development and interaction.

4 Conclusion

This study, by focusing on classic scenes featuring Harry and Hermione in *The Philosopher’s Stone*, innovatively integrates the visual symbolic system with the subtitle translation modality, validating the effectiveness of multimodal synergistic analysis for character interpretation. The research demonstrates that visual elements are the core vehicle for personality externalization, while subtitle translation plays a crucial supplementary role in reinforcing character shaping. The intertextuality formed by the visual and subtitle bimodal system significantly enhances character portrayal. These findings directly serve the core theme of this paper—the impact of multimodal synergy on cross-cultural character portrayal—and hold practical value for film translators, cross-cultural communication researchers, and film educators. In the context of global film and television dissemination, this study holds dual significance: theoretically, it breaks through the limitations of unimodal research, confirming multimodal synergy as the cornerstone of cross-cultural character cognition; practically, it reveals the need for dynamic adaptation between Chinese subtitles and visual symbols, offering innovative value for enhancing audience character empathy and reducing cultural misinterpretation across cultures. The analysis of isolated scenes in a single film within this study fails to reveal the evolutionary patterns of bimodal synergy in the long-term character development across a film series. Future research could extend to other dual-protagonist comparisons (e.g., Ron and Malfoy) and explore the application potential of artificial intelligence in optimizing multimodal translation.

References

1. X. Wei, L. Qiu & Z. Wang. The Presentation of Visual Elements in Film within the Context of Cultural Appreciation. *JYL*, (15), 85-87 (2023)
2. L. You, Aesthetic Reflections on Film Visual Styling Elements and Picture Composition. *Movie Literature*, (14), 84-88 (2024)
3. X. Liu, Chinese-English Subtitle Translation Techniques from the Perspective of Relevance Translation Theory—Taking Chang’an 30,000 Li as an Example. *Sinogram Culture*, (12), 158-160 (2025)
4. Y. Hou, Research on the English Subtitle Translation of “Bring in the Wine” in Chang’an 30,000 Li Based on Fu Lei’s “Spiritual Resemblance” Theory. *Overseas English*, (10), 31-33 (2025)
5. M. Ke, The Embodiment of Skopos Theory in the Subtitle

Translation of English Idioms—Taking Season 1 of Gravity Falls as an Example. *JinGu Creative Literature*, (18), 102-104 (2025)

6. Z. Su, Subtitle Translation of the Film Big Hero 6 from the Perspective of Multimodal Discourse Analysis. *Delta*, (12), 70-72 (2025)

7. L. Wei, A Study on the Subtitle Translation of Kung Fu Panda 4 from the Perspective of Multimodal Discourse Analysis. *ECM*, (08), 3-6 (2025)

8. J. Mao, & F. Su, The Theoretical Contribution of Case

Study Research—A Review of the Forum on Case Study and Qualitative Research of Chinese Enterprise Management (2015). *JMW*, (02), 128-132 (2016)

9. S. Guan, Multimodal Discourse Analysis from the Perspective of Visual Grammar Theory—Taking the Foshan City Tourism Image Promotional Video as an Example. *JCIT*, 41(03), 67-72 (2024)

10. Z. Cao, Research on Subtitle Translation and National Image Construction from a Multimodal Perspective, Master's thesis, Beijing Foreign Studies University (2024)