
| RESEARCH ARTICLE

On Text-Image Relations and Information Conveyance Optimization in Subtitle Translation: A Case Study of *Chang An*

Zhao Ruixue

School of Foreign Languages, Chongqing Normal University, Chongqing, China

Corresponding Author: Zhao Ruixue, **E-mail:** wsruixue@126.com

| ABSTRACT

Subtitle translation, in comparison with other forms of translation, is more closely related to visual symbols. This article examines the correlation between images and translated subtitles in movies viewed from the information focus and the information content. Through a detailed exploration of translation cases from the animated film *Chang An*, this study discusses in which way the interplay of images and translated subtitles optimize the expressiveness of the movie, based on which, it is suggested that by applying specific translation techniques, the translation of subtitle text should be tailored to harmonize with visual elements, so as to achieve better information dissemination in visual-text mediums.

| KEYWORDS

Text-Image Relations; Subtitle Translation; Optimization of Translations; Information Focus; Information Content.

| ARTICLE INFORMATION

ACCEPTED: 15 March 2024

PUBLISHED: 04 April 2024

DOI: 10.32996/ijllt.2024.7.4.5

1. Introduction

Compared to traditional texts, audiovisual productions enjoy greater popularity among audiences due to their figurative and dynamic nature. Within the digital era, there exists a burgeoning trend in the market for audiovisual productions, characterized by an inclination towards expanding access overseas. Given that films have long been ingrained within culture, serving as reflections of a nation's collective personality, the dissemination of films from various countries abroad, to some extent, contributes to shaping the collective personality image of each respective country.

For films to be accurately comprehended by foreign audiences, especially those whose native language differs from the film's source language, the quality of subtitles assumes paramount importance. In practice, for target audiences proficient in the source language, subtitles in foreign language films merely serve as textual aids, often swiftly scanned to decipher dialogue content. However, the viewership of foreign films extends beyond those adept in foreign languages; many foreign viewers, lacking proficiency in the source language, heavily rely on subtitles provided in their native languages. From this perspective, subtitle translation emerges as an indispensable tool for enabling audience comprehension of foreign language films.

According to the definition by American scholar Mark Shuttleworth (2014), subtitling refers to one of the two main methods of language transfer used in translating types of mass audio-visual communication such as film and television (see also dubbing). It involves the provision of synchronized captions for film and television dialogue. Differing from other translation modalities, subtitling operates within constraints of time and space, necessitating simultaneous consideration of textual and audiovisual symbols. As Ketola (2016) contends, verbal text paired with images may undergo distinct translation compared to the same text without accompanying images. Notably, in the realm of subtitle translation, a multitude of factors warrant consideration. Among the challenges shared by all forms of audiovisual translation is the limited control exerted by translators over elements of multimodal discourse. Visual images are typically static and immutable, constituting a fixed component that cannot be altered

during the translation process (Braun, 2016). Consequently, translators tasked with subtitle translation are compelled to judiciously adjust translations within the constraints of time and space, thereby optimizing text-image expressiveness.

2. Literature Review

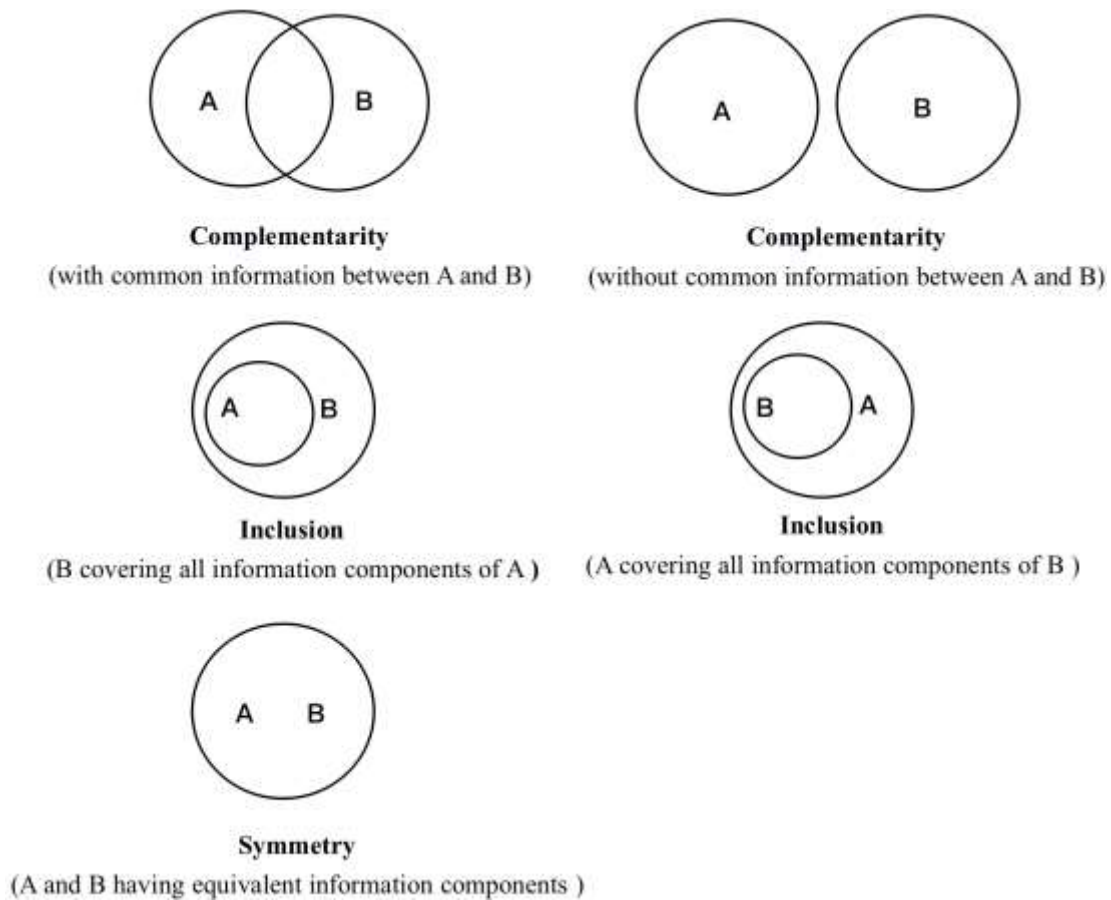
Subtitling entails not the translation of pre-existing source language subtitles but rather the addition of translated subtitles to audiovisual works (regardless of the presence of source language subtitles) (Yang & Xue, 2023). As per the definition provided by Díaz-Cintas and Remael (2014), subtitle translation constitutes a translation practice involving the presentation of written text (typically located at the bottom of the screen) to describe the original spoken dialogue of the speakers, various speech elements appearing in the imagery (such as letters, inserted texts, graffiti, inscriptions, signboards, etc.), and the information contained in the soundtrack (songs, off-screen sounds).

From the aforementioned definition, it is evident that audiovisual subtitle translation inherently embodies multi-modal translation characteristics. A salient feature of multi-modality is the prioritization of language without compromising other semiotic modes (Pérez-González, 2014). Consequently, while subtitle translators concentrate on translating textual content, they must also attend to other modes, such as visual and audio effects, to ensure that the translated text maximizes the effectiveness and richness of information dissemination in audiovisual works. Subtitle translation is circumscribed by temporal and spatial limitations, rendering it impractical to solely discuss the maximum of meaning activated and represented by the image and text (Qiu, 2020). Adherence to the language economy principle in audiovisual translation, i.e., ensuring that subtitles in audiovisual productions are concise and accurate in conveying the information contained in the source language, is imperative (Yao & Cai, 2010). Hence, this paper adopts the term “information conveyance optimization” rather than “richness of information expression” to describe the success of subtitle translation.

The interaction between text and image stands out as one of the most prominent characteristics distinguishing audiovisual works from typical textual works. Consequently, an emphasis on text-image relationships represents a crucial approach to researching subtitle translation. Audiovisual texts encompass various symbols owing to their multi-modal nature. Scholars such as Delabastita (1989), Zabalbeascoa (2008), and Sokoli (2009) have classified audiovisual text symbols into two types: linguistic and non-linguistic symbols, which are further combined with auditory and visual communication channels to yield four different types of symbols: auditory linguistic, auditory non-linguistic, visual linguistic, and visual non-linguistic. Text-image relationships in audiovisual works pertain to the relationship between visual linguistic and visual non-linguistic elements. While understanding text-image relationships in audiovisual productions necessitates dynamic interpretation within the context of moving images (here, “dynamic images” refer to film shots that are edited and presented in a dynamic form in front of the audience while following a certain logical relationship), research on static text-image relationships (such as picture books, illustrated stories, etc.) holds reference value for analyzing text-image relationships in audiovisual productions. To provide a more concrete elucidation of text-image relationships, this paper analyzes text-image information content from the perspective of text-image information.

Barthes (1999) elucidated the mechanism of text-image interaction and meaning production by interpreting an image-text advertisement. He delineated two types of relationships between text and image: “Anchorage” and “Relay” — where language information can serve as an “anchor” to “fix the floating chain of signifieds.” Linguistic information can identify elements in the image and the image itself. Royce (1998) initially proposed the concept of multi-modal intertextuality, positing that visual and textual modes complement each other in discourse application, forming a complementary relationship. Nikolajeva and Scott (2000), through analysis of picture book examples, categorized text-image relationships into four categories: enhancing (symmetry), complementary interaction, counterpointing dynamics, and contradictory interaction. Building on Barthes’ research on text-image relationships and Halliday’s research on logical-semantic units in language systems, Martinec and Salway (2005) further examined new and old media discourses, developing the system of text-image relationships. Tian and Liu (2017) classified the “status” relationship and the “logical-semantic” relationship between textual and visual modes, exploring the configuration relationship between text and image in works from three aspects: schema preservation, schema enhancement, and schema updating. Chen and Zhang (2017) applied the logical-semantic relationships of linguistic units—elaboration, extension, and enhancement—to analyze text-image logical relationships.

Comparing the richness of meaning between text and image information in terms of sensory reception forms proves challenging due to images constituting visual non-linguistic information while texts represent visual, linguistic information. However, from the perspective of text-image information content, the text-image relationships summarized in the above studies can be roughly categorized as follows: symmetry — where the information of text is equivalent to that of the image; complementarity—where information of text and image is either partly the same or totally different, inclusion—all information of one mode (text/image) is included by the other. The Venn diagrams depicting text-image relationships are as follows (A and B represent text and image, respectively, and the size of the circles does not reflect the actual amount of text-image information but only highlights the relationship between text and image information content):



Viewed in terms of function, text-image relationships can also be comprehended through the diagram above. For instance, the phenomenon of “anchorage” of text to images typically occurs when “the textual information contains all the visual information,” as “anchorage” denotes language assisting readers or viewers in locating elements in the image, and this positioning relies on the premise that textual information is contained in and not identical to visual information (Barthes, 1999).

The “anchorage” text-image relationship offers a significant perspective for analyzing text-image information content in the preceding paragraph, indicating that the coordination between the focal information in the visual and the subtitles—rather than all elements of the visual scene—is the most pertinent point of discussion in audiovisual subtitle translation. For instance, the visual elements accompanying character dialogue invariably encompass more information than the spoken language information (i.e., subtitle information). In such instances, the total visual information content almost invariably surpasses the subtitle information content. Consequently, the textual representation of character dialogue only necessitates consideration of the focal visual information (character expressions, intonation) and does not require attention to irrelevant decorative background information (unrelated scenery and characters, etc.). Therefore, when analyzing the translation of subtitles solely conveying dialogue meaning, the examination of text-image relationships is confined to the level of focal visual elements and translated subtitles.

Given the preliminary examination of existing literature on the topic, the present study endeavors to address the following research questions:

- RQ 1:** What is the frequency of the respective relations?
- RQ 2:** Which text-image relations are more acceptable in subtitle translation?
- RQ 3:** What translation techniques are applied, and do they facilitate better information dissemination in visual-text mediums?

3. Methodology

The translation challenges encountered in audiovisual productions predominantly revolve around the representation of cultural symbols and the expression of emotional significance. This paper is firmly grounded in addressing these two significant challenges, opting to scrutinize the audiovisual work of Chang An, renowned for its rich cultural and emotional expressions (Chen, 2023), through a specific analysis of text-image cases. The study aims to delve into the three relationships of symmetry, complementarity,

and inclusion while also examining how translators adjust subtitle translations to ensure the optimal transmission of text-image information.

3.1 Research Methods

The research methods employed in this paper are primarily descriptive, with a focal point on delineating the text-image relationships in specific cases. While qualitative in nature, the study also integrates a quantitative analysis of the text-image relations between culturally loaded source subtitles and visuals.

3.2 Research Procedures

To fulfill the objectives of this research, the study was structured as follows:

- Selection of visuals and subtitles from Chang An
- Extraction of Chinese and English subtitles from the film
- Analysis of the relationship between visuals and subtitles
- Evaluation of subtitling techniques to ascertain whether they facilitate audiences' efficient understanding

The analysis involves a comparison between the visuals and the subtitles to ascertain whether appropriate translation techniques were employed to harmonize with the visuals, thereby optimizing the conveyance of meaning from the source subtitles.

4. Results and Discussion

The results derived from the content analysis are delineated in the subsequent subsections. The text-image relations are categorized into three overarching classifications: symmetry, complementarity, and inclusion. Table 1 depicts the frequency of occurrence for each of the aforementioned relations in both the culturally loaded source subtitles and visuals.

Table 1: Frequency of Different Relations

Image-Text Relations	Symmetry		9
	Complementarity	With Common Information	36
		Without Common Information	23
	Inclusion	Image > Text	11
		Text > Image	3

The relation of symmetry exhibits the fewest occurrences in this film, with cases of symmetry primarily associated with the notation of characters or places. Conversely, the relation of complementarity is the most prevalent, indicating a tendency for subtitles to adhere to the language economy principle. Concerning inclusion, the informational components of images surpass those of the corresponding translated texts, necessitating simplification of the subtitles to comply with time and space constraints.

4.1 Symmetry between Visuals and Subtitles

The concept of "symmetry" is drawn from Nikolajeva and Scott (2000), who derived this type of text-image relationship through an analysis of picture books. Originally, it refers to a correspondence between the content conveyed by images and words, where the image and text are entirely correlated. The symmetric relationship also entails an equivalence in the informational content between the image and text.

Example1: [Footage 01:07:10-01:07:45]



[Shot 01:07:17]



[Shot 01:07:44]

Source texts: 床前明月光 疑是地上霜 举头望明月 低头思故乡

Subtitles: Moonlight at the foot of my bed/ Appears to me like frost./ Lifting my head, I behold the moon/ Lowering it, I think of home.

In this scene, Li Bai (One of the most prominent poets in the Tang Dynasty of China) is depicted holding a wine glass and reciting a poem he previously composed with a sorrowful expression. Combined with the earlier scene depicting "Li Bai recalling the pain of losing his father and being ostracized by his brothers," we can distinctly perceive the character's longing for his relatives and hometown at this moment. The English subtitles generally achieve informational equivalence with the Chinese subtitles. However, the alignment of information between the English subtitles and the visual scene is only partially symmetrical. When the subtitle displays the first line of the poem, "Moonlight at the foot of my bed," the visual depicts Li Bai standing by the window without including "my bed", as indicated in the subtitle. Similarly, when the subtitle presents the second line, "Lifting my head, I behold the moon," although Li Bai does tilt his head upward in the scene, the word "moon" is not included in the visual depiction. The true realization of text-image symmetry in this example occurs only with the fourth line of the poem, "Lowering it, I think of home," and its corresponding visual representation: when this subtitle appears, Li Bai lowers his head, aligning with "Lowering it," and his sobbing action clearly conveys the meaning of "think of home."

It can be observed that the text-image relationship in the above examples does not exhibit complete symmetry. However, this does not impede foreign audiences' understanding of textual and visual information. The absence of the words "bed" and "moon" in the visual does not affect foreign viewers' comprehension of the poetic imagery. The moon reflected in the wine glass in the previous shot compensates for the absence of the "moon" in this shot, accurately conveying the meaning of "Moonlight at the foot of my bed" in Chinese culture. Additionally, "bed" is not a focal point in the original poem; hence, the omission of "bed" in the visual representation does not affect the visual expression of the poem's emotion.

Nevertheless, for some film scenes, symmetry in subtitle information is crucial, as illustrated in the following segment.

Example2: [Footage 23:32-23:35]



[Shot 23:33]

In this scene, there are no subtitles, only the visual imagery: Li Bai stands before the grave of his departed friend, Wu Zhihan, with a bundle of upright bound grass and a stream of water pouring down. The visual representation here depicts an important ritual in Chinese sacrificial ceremonies known as “libation.” Specifically, this involves erecting a bundle of thatch grass and pouring wine over it, with the dregs of the wine being retained within the thatch while the wine gradually seeps down, offered for the gods to partake. Both Eastern and Western cultures have libation customs; hence, foreign viewers can comprehend that this scene portrays the act of “libation,” but the cultural significance of libation differs significantly between Eastern and Western cultures.

In Chinese sacrificial culture, wine serves merely as a tool for offering to deities and honoring deceased loved ones, remaining largely static throughout the ritual, with participants primarily engaging in the act of “offering wine.” Offering wine to deities is done to seek blessings while offering it to departed loved ones is an expression of remembrance and the hope for the deceased to rest in peace.

In Western cultures’ wine deity rituals, people form grand processions, and choruses perform jubilantly in forests or wilderness, using music and wine not only to express reverence for the wine deity but also to pursue the exhilarating and joyful intoxication brought by wine (Jia & Li, 2010). Thus, participants in the ritual not only engage in the act of “offering wine” but also actively participate in the enjoyment of “drinking wine.” In fact, in Western culture, people commune with the wine deity by drinking together during the ritual, experiencing the freedom and joy brought by alcohol’s marvelous effects on the body, pursuing individual pleasure, which is a significant characteristic of Western wine culture.

According to Lemke (2009), when a visual image carries culture-specific information or is linked with the film dialogue in a specific way, the subtitles may have to be phrased so as to convey visual implicatures that the target recipients would not be able to derive. It can be said that the translated text’s potential for triggering a train of connotations in the target-viewers’ minds is the most important aspect of text design. Images are polysemous, but it is not clear whether all connotations are accessible to viewers in different cultures (Smith, 2008). In this scene, the poet Gao Shi’s libation is for the “underground peace” of Li Bai’s deceased friend. Therefore, adding translated subtitles such as “Libation for friend’s underground peace” in this shot, through text-image symmetry, would enable Western readers to truly understand the genuine meaning of this “libation” ritual in China.

Thus, when visual imagery may lead to misunderstandings (whether of information or emotional nuances) or pose difficulties in understanding, it is essential to add explanatory text in the translation to clarify the information, assisting target readers in achieving a correct understanding of the visuals.

4.2 Complementarity between Visuals and Subtitles

The term “text-image complementarity” has long been in existence, yet most descriptions of it tend to offer relatively abstract conceptual introductions, and its general meaning is often very broad. The text-image relationship of complementarity in this paper is more specific, or rather, its significance is more narrowly defined, primarily referring to the simultaneous inclusion of information in text and image that the other does not contain (they may share information or have none at all), thereby achieving an effect where the conveyed meaning exceeds the sum of its parts, akin to the effect of “1+1>2”.

Example1: [Footage 01:09:44-01:09:53]



[Shot 01:09:45]

Source text: 夜来风雨声 花落知多少
Subtitles: A loud stormy night/ countless petals fall.

In this scene, Li Bai (positioned on the right in the image) and Gao Shi are depicted riding slowly along a rural path leading to the home of another poet, Meng Haoran. The scene is bathed in perfect sunlight, with flowers, grass, and trees creating a tranquil yet vibrant atmosphere. It is evident that the information conveyed in the subtitles, namely “A loud stormy night” and “countless petals fall,” starkly contrasts with the depiction of “bright daylight” and “flowers in quiet bloom” in the image.

However, this phenomenon of complete mismatch between text and image information can elicit a sense of novelty in the audience, thereby stimulating their curiosity to understand the reason behind such disparity. As Li Bai recites these two lines from Meng Haoran's poem *春晓/Spring Morning* in this scene, both he and Gao Shi are depicted smiling and appearing relaxed. From this, one can infer that the stormy weather described in the poem has passed, and a sense of renewal is implied by the falling petals. The harmonious natural environment portrayed in the entire scene also reflects Meng's affection for the countryside and his longing for seclusion. Therefore, the phenomenon of disparity between text and image in this context actually complements each other, enabling foreign viewers to gain a deeper appreciation for the poet's yearning for tranquility and harmony. Tranquility and harmony fundamentally epitomize the intrinsic attributes of the collective personality of Chinese culture, which emphasizes non-contention.

Example2: [Footage 01:53:40-01:53:56]



[Shot 01:53:43]

Source text: 烹羊宰牛且为乐 会须一饮三百杯
Subtitles: Stew the lamb, prepare the ox, let us feast/ Tonight we'll drink three hundred cups at least

The selected subtitle and visual segment featured here originate from one of the film's most exceptional poetic design sequences, corresponding to Li Bai's poem *将进酒/Invitation to Wine*. The information conveyed in the subtitles, such as “stew the lamb, prepare the ox” and “three hundred cups,” lacks depiction in the visual imagery. Conversely, imagery such as “white crane” and “flying with the crane” is absent in the subtitles, thus exemplifying a case of textual-visual disparity, consistent with the concept of “complementary textual-visual information” as delineated in this paper. In this scenario, the translated subtitles steadfastly adhere

to the source text without introducing, omitting, or modifying any informational elements, even rendering the indefinite quantity of “三百杯” from the source text as a literal translation of “three hundred cups.” The meticulous attention to lexical detail underscores the translator's unwavering fidelity to the source text regarding informational elements.

Indeed, upon scrutinizing the complete poetic and visual segment of *Invitation to Wine* from the film footage [01:52:40-01:56:08], it becomes evident that the subtitles and visuals exhibit significant asymmetry attributable to the nature of the poetry itself. This poem is not a narrative descriptive piece but rather an extravagant and unrestrained lyric composition characterized by an expansive and uninhibited structure. Given that the objective is “to paint the poem,” the visuals must exhibit a high degree of creativity and vividness to aptly correspond to the poet's exuberant and carefree demeanor. Consequently, this segment astutely leverages the complementary interplay between textual and visual information, markedly enhancing the vivacity of the poetic expression.

In literary translation, the faithful reproduction of aesthetics serves as a pivotal criterion for evaluating the translator's fidelity (Lin & Mao, 2023). When the imagery and the original text's informational components diverge or contradict each other, the translation should strive to maintain utmost fidelity to the original text, conveying all informational elements, ensuring coherence, maximizing communicative intent, and achieving the most favorable aesthetic outcome in narrative presentation.

4.3 Inclusion between Visuals and Subtitles

The relationship of inclusion, as delineated within this text, pertains to a scenario in which the informational elements conveyed by the picture (text) comprehensively envelop and surpass those of the other. Such a text-image relationship is frequently observed in the descriptive imagery prevalent in films, as evidenced by scenes depicting characters' reminiscences or instances of textual visualization.

Example1: [Footage 30:53-30:54]



[Shot 30:54]

Source text: 看这四面的诗板便知

Subtitles: Just look around, sir.

In this particular scene, the corresponding visual portrays a waiter at the Yellow Crane Tower (黄鹤楼) guiding Li Bai and Gao Shi to view the poetry boards hanging from the ceiling inside the tower. Interestingly, the term “诗板” is not translated in the English subtitles as it had been previously (as observed in earlier subtitle translations where “诗板” was translated as “poetry board”). This departure from past practice can be attributed to two factors: firstly, the term “poetry board” has been consistently used in preceding contextual subtitles, and secondly, the visual depiction in this scene vividly presents the physical manifestation of the “诗板,” encapsulating all relevant information conveyed in the subtitles. “诗板” refers to wooden boards traditionally employed in ancient China for inscribing poetry. Despite carrying evident cultural connotations associated with China, the tangible nature of the “诗板” facilitates a more seamless integration of visual and textual elements within this cinematic portrayal.

Example2: [Footage 32:25-32:41]



[Shot 32:35]

Source text: 日暮乡关何处是 烟波江上使人愁

Subtitles: Dusk sets in, where is my home?/Mist and waves, a melancholy scene.

The scene unfolds as a dynamic black-and-white landscape painting: a round sun appears to set at the end of the river while a poet stands with hands behind his back in a small boat slowly drifting on the river. Clear mist hovers over the river, with only two birds flying overhead, and dense mountains and chaotic vegetation line the riverbank.

The visual elements in this scene notably surpass the textual elements in the English subtitles. Details such as the “two birds,” “the poet standing alone,” and “the small boat” are omitted from the subtitles, yet they vividly and aptly describe and enrich the meaning of “a melancholy scene.” Consequently, the inclusion relationship between the visuals and the subtitles significantly enhances the audience’s visual experience and aids in understanding the sense of splendid solitude evoked in classical Chinese poetry—a sentiment unique to Chinese culture.

It is noteworthy that the English subtitles in this instance, relative to the Chinese subtitles, are not entirely faithful. The translator renders the poet’s “愁” (gloomy mood) as “a melancholy scene.” Since the poem lacks a specific description of “what is the poet worrying about,” the translator’s flexible approach, while diminishing the negative emotion conveyed in the original text, avoids potential confusion for foreign audiences and accurately conveys the poet’s profound sense of solitude resulting from the juxtaposition of melancholy and beautiful scenery.

Thus, the inclusion relationship between text and image is applicable when the original subtitles contain excessive cultural information or when the visual cues are explicit. Typically, the inclusion of text by images is a common scenario. In practical subtitle translation, when there is an excess of textual information, translators may omit certain elements of the original subtitles to balance the quality and quantity of information conveyed, ensuring the optimal understanding of the plot by foreign audiences. Conversely, when visual elements include certain textual elements from the original subtitles, translators may omit parts of the original subtitles to adhere to the language economy principle of film subtitle translation and prevent additional burdens on foreign audiences in processing extra information.

5. Conclusion

The quality of subtitle translation is a critical determinant of the success of films in international dissemination, with the relationship between text and image offering a valuable research perspective for subtitle translation. This paper further refines the categorization of the relationship between text and image from the standpoint of information content, integrating specific cases from Chang An to dissect the relationship between visuals and subtitles in cinematic works. The evidence suggests that in cinematic works, the relationship between translated subtitles and images is multifaceted, presenting as symmetrical, complementary, or inclusive relationships based on the logical progression of the narrative.

When translating subtitles, adherence to the language economy principle of subtitle translation is paramount, allowing the relationship between text and image to manifest as either inclusion or complementarity: when original textual elements are excessive or when cinematic images encapsulate all information from the original subtitles, subtitle translation should be streamlined. Conversely, when visual information contradicts or diverges from the original subtitle content, fidelity to the original subtitles should be maintained to ensure the juxtaposition of each information element.

Only in cases where visual information is relatively ambiguous compared to the original subtitles or when misunderstandings arise due to missing information in the original subtitles should the translation align symmetrically with the image. This may involve adding explanatory subtitles to ensure the audience's accurate understanding of cultural nuances or plot elements. Through such meticulous approaches, subtitle translation can achieve optimal expressive efficacy with the minimum amount of translated subtitles.

However, it should be recognized that subtitle translation inherently falls within the realm of multimodal translation, requiring a comprehensive consideration of various elements such as text, image, and sound to ensure the quality and expressiveness of the ultimate translation. Given the constraints of time, this study exclusively examines the interplay between text and image in subtitle translation, with no attention paid to the auditory mode. Moreover, the cases scrutinized in this research are confined to a singular film, thus unavoidably limiting the range of case typologies. It is recommended that future studies on subtitle translation accord equal attention to the textual, visual, and auditory modes while also endeavoring to diversify case selections across a spectrum of audiovisual works. This approach will facilitate a more exhaustive and nuanced exploration of subtitle translation.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Barthes, R. (1999). Rhetoric of the Image. *Visual Culture: The Reader*, 33-40.
- [2] Braun, S. (2016). The Importance of Being Relevant? A Cognitive-Pragmatic Framework for Conceptualising Audiovisual Translation. *Target*, 28(2), 302-313.
- [3] Chen, N. (2023) A New Perspective on the Tang Dynasty and A New Method to Interpret Tang Poetry: A Review of *Chang An. Sichuan Drama*, (12): 93-98.
- [4] Chen, Y. P., & Zhang, C. H. (2017) English to Chinese Translation of Movie Subtitles: A Multimodal Analysis Focused on Text-Image Relations. *Chinese Translators Journal*, (05), 105-110.
- [5] Cintas, J. D., & Remael, A. (2014). *Audiovisual Translation: Subtitling*. Routledge.
- [6] Delabastita, D. (1989). Translation and Mass-Communication: Film and TV Translation as Evidence of Cultural Dynamics. *Babel*, 35(4), 193-218.
- [7] Jia, Q., & Li, X. Y. (2010). Romantic Poetry and Ritualistic Libations: An Analysis of Functional Differences in Sino-Western Liquor Culture. *Masterpieces Review*, (09),135-137.
- [8] Ketola, A. (2016). Towards a multimodally oriented theory of translation: A cognitive framework for the translation of illustrated technical texts. *Translation Studies*, 9(1), 67-81.
- [9] Lemke, J. (2009). Multimodality, Identity, and Time. *The Routledge Handbook of Multimodal Analysis*, 140-150.
- [10] Lin, Z., & Mao, Y. F. (2023). Loyal Resistance: A Study of Lv Shuxiang's Poetics on Translation. *Sichuan University of Arts and Science Journal*, (06),110-115.
- [11] Martinec, R., & Salway, A. (2005). A System for Image-Text Relations in New (and Old) Media. *Visual Communication*, 4(3), 337-371.
- [12] Nikolajeva, M., & Scott, C. (2000). The Dynamics of Picturebook Communication. *Children's Literature in Education*, 31, 225-239.
- [13] Pérez-González, L. (2014). *Audiovisual Translation: Theories, Methods and Issues*. Routledge.
- [14] Qiu, Q. (2020). Complementation of Pictures and Texts and Meaning Construction—A Study of Multimodal Discourse of Children's Picture Books. *Journal of Jiangxi Normal University(Philosophy and Social Sciences Edition)*, (03), 129-136.
- [15] Royce, T. (1998). Synergy on the Page: Exploring Intersemiotic Complementarity in Page-based Multimodal Text. *JASFL Occasional Papers*, 1(1), 25-49.
- [16] Smith, V. (2008). Visual persuasion: Issues in the translation of the visual in advertising. *Meta*, 53(1), 44-61.
- [17] Sokoli, S. (2009). Subtitling Norms in Greece and Spain. In *Audiovisual Translation: Language Transfer on Screen* (pp. 36-48). London: Palgrave Macmillan UK.
- [18] Shuttleworth, M. (2014). *Dictionary of translation studies*. Routledge.
- [19] Tian, Y., & Liu, S. S. (2017). On Image-Text Relations in Multimodal Novels and Meaning Construction. *Foreign Language Education*, (02), 39-43+98.
- [20] Yang, S. Z., & Xue, X. W. (2023). Integrating Core Modes in Film and TV Texts for Subtitling. *Technology Enhanced Foreign Languages*, (03), 58-66+121.
- [21] Yao, X. Y., & Cai, J. M. (2010). On Principles of Subtitle Translation. *Movie Literature*, (18), 140-141.
- [22] Zabalbeascoa, P. (2008). The Nature of the Audiovisual Text and Its Parameters. *The Didactics of Audiovisual Translation*, 7, 21-37.