

Negotiating Meta-functions in Multimodal Translation: A Systemic Functional Analysis of the Chinese–English Subtitling in Chang'an

Yuqing Zhou

Faculty of Humanity and Social Sciences, Beijing Normal-Hong Kong Baptist University, Zhuhai, China

s230025117@mail.bubu.edu.cn

Abstract. One of the most important studies in audiovisual translation (AVT) is the 2023 Chinese animated film *Chang'an*, which receptively recreates the golden age of the Tang Dynasty and intertwines the stories of Li Bai and Gao Shi with classical poetry and historical images. It presents a special challenge, a chance of delivering the Tang-era poetical, historical and cultural overtones of the film to the foreign viewer through English subtitles. This study examines its English subtitles through an integrated framework of Systemic Functional Linguistics (SFL) and intermodal relations analysis. Findings indicate that translators strategically simplify cultural-specific terms and interpersonal cues to accommodate spatial constraints, while the film's visual and auditory channels compensate through reinforcement and contextualization. This compensatory process indicates that successful subtitling is not executed in an ambiguous language accuracy, but rather in a multimodal negotiation strategy. The study shows the role of the cultural-dense cinematography based on the interaction of both verbal and non-verbal in maintaining the central meaning of the movie, which is significant knowledge about the cross-cultural media localization and passing of the poetic tradition.

Keywords: Multimodality, Audiovisual translation, Subtitling, Chinese animation

1. Introduction

Audiovisual translation has emerged as an essential mode of multimodal meaning-making in the contemporary globalized media era. The 2023 Chinese animated epic *Chang'an*, directed by Xie Junwei and Zou Jing, vividly reconstructs the golden age of the Tang Dynasty, weaving the narratives of poets Li Bai and Gao Shi into the threads of rich historical and cultural motifs. Rooted in classical Chinese poetry and aesthetics, the film presents significant challenges as a cross-cultural film-English subtitles need to be as accurate and nuanced as possible, portraying the linguistic meaning of the speech alongside its cultural, historical and emotional tones of the speech and their significance to the audience.

This complexity demands a theoretical framework that can encapsulate the interaction between language and other semiotic resources. Halliday's Systemic Functional Linguistics (SFL)-with its

ideational, interpersonal, and textual meta-functions-provides a strong analytical tool for meaning-making [1]. Its strength in subtitle analysis lies in the disclosure of the ideational content (event, characters, and cultural concepts), interpersonal dynamics (relationships, attitudes, and evaluations) and textual structures which means flow of information and cohesion of themes as linguistically constructed in the original text. This framework, together with the taxonomy of intermodal relations presented by Diaz-Cintas [2], redundancy, complementation, contradiction, and anchorage, helps to examine how the subtitles will interact with visual and auditory channels in detail. SFL finds meaning in dialogue; the intermodal taxonomy presents the way subtitles support (redundancy), expand (complementation), subvert (contradiction), or clarify (anchorage) simultaneous imagery and sound. Taylor endorses this integrated multimodal approach [3], arguing that meaning is distributed and reconstructed across different modes, thus a comprehensive analysis must therefore consider their functional interrelations [4]. O'Halloran further highlights the value of exploring "inter-semiotic texture" between language and other semiotic resources in meaning construction. Even though SFL has been effectively implemented in the context of monolingual multimodality research, its methodical application to intermodal relations is still underdeveloped in the field of subtitle translation, especially in culturally rich films such as the movie *Chang'an*, where language, culture, and sense interact with each other on several levels of semiotic representation.

This study addresses this gap through a contrastive multimodal analysis of the film's original Chinese dialogue and its English subtitles. The following research questions guide the investigation:

1. How are the ideational, interpersonal, and textual metafunctions in the original Chinese dialogue of *Chang'an* reconfigured in the English subtitles, and what translation strategies are employed to navigate these shifts?
2. What are the interactions of the English subtitles with the visual and auditory modes of the film (e.g. by redundancy, complement, contradiction, or anchorage), and how do these intermodal relationships influence the construction of poetic, historical and cultural meanings?

The research is both theoretically and practically relevant. Theoretically, it will contribute to the existing prosperous body of research on the studies of multimodal translation as it will be thoroughly tested and refined to demonstrate the effectiveness of the integrated model of SFL-intermodal relations and prove its efficiency during the delicate negotiations of meaning in a restricted translation environment. Practically, the results shed light on the targeted strategies and challenges of mediating a deep cultural heritage to a global audience. This study may be necessary for subtitlers, cultural intermediaries, and policy makers to improve the efficacy of cross-cultural cinema and create a more discriminating international recognition of Chinese history and classical aesthetics.

2. Literature review

Audiovisual translation (AVT) research has evolved from a peripheral subfield of translation and linguistics to a core area of multimodal discourse investigation. Such metamorphosis was traced by bibliometric work by Liu et al. [5]. They plotted an exponential increase in AVT-related publications since 2012 and found the core authors in the research-these include Díaz-Cintas, Chaume, and Pérez-Gonzalez, as well as corpus-based AVT, cultural Otherness, audience reception, and computational approaches, which substantiate the move toward an empirical and audience-oriented scholarship [5,6].

The social semiotic paradigm serves as a dominant theoretical foundation for AVT. Based on the Halliday's SFL [7], Kress and van Leeuwen developed a "grammar" of non-linguistic modes, differentiating between ideational, interpersonal and textual metafunctions [8]. This model enables

analysts to break down the contribution of visual and auditory resources in making meaning to filmic discourse. However, as Forceville notes, this approach often overlooks how inference, attention modulation, and schema activation mediate multimodal meaning construction [9].

Zhang proposed a Sinocentric adaptation of SFL in Chinese translation studies, emphasizing the synergy and interplay among image, sound, and verbal modes within Chinese cultural semiotic systems [10]. While theoretically rich, this framework has seen limited empirical application in AVT research—most studies remain expository or methodological, lacking rigorous multimodal analysis.

Social semiotic scholars have developed more functional or relational taxonomies to explain the interactions of modes. A foundational example is Perego and Díaz-Cintas' typology of intermodal relations (redundancy, complementation, contradiction, anchorage), which has become a staple to characterise the interaction between subtitle and image and audio [11]. Ramos Pinto and Mubarak's Multimodal Corpus Analysis of Subtitling implements this taxonomy to examine how subtitles mediate nonstandard speech varieties in film, showing how subtitling can preserve, cancel, or reshape intermodal relations [12]. Their study represents a significant step toward corpus-based multimodal analysis.

Nevertheless, two persistent limitations shadow this strand of work. First, most case studies and corpora are drawn from Western or Anglophone audiovisual products. As a result, their findings may reflect Eurocentric conventions in visual-verbal interaction, which may not generalise to East Asian or Chinese cinematic practices. Second, these relational taxonomies often foreground how modes relate (i.e. the pattern of interaction) but do not systematically theorise what meanings are semantically produced, negotiated, or lost in these interactions—that is, they seldom re-invoke the metafunctional lens (ideational, interpersonal, textual) that gave impetus to their design.

In response to these methodological gaps, AVT research has expanded into reception and cognitive approaches. Forceville initiated one such turn by arguing for cognitive models of viewer inference, mental visualisation, and integrative comprehension [9]. Subsequent empirical studies—especially via eye-tracking, reaction times, and cognitive load metrics—have enriched our knowledge of subtitle readability, attention split, and information processing in AVT. However, this reception-oriented strand sometimes downplays the socio-cultural conventions embedded in semiotic resources—that is, visual or auditory symbols that carry culturally specific affordances may be treated as “neutral stimuli” rather than as semiotically loaded sign systems.

In parallel, Chinese scholarship on Chang'an and related historically and culturally dense texts has produced analyses of Tang poetry translation, cultural allusion, and lexical choices in subtitles [13,14]. These works help preserve poetic resonance and cultural connotations in translation, but they often treat the visual-aural dimension as static context rather than as co-constitutive in meaning-making. They seldom adopt a fully multimodal analytical posture.

A critical gap thus emerges—Western AVT scholarship offers advanced methods and taxonomies, it has rarely been applied to Sinocentric audiovisual texts in a culturally attuned way. Meanwhile, Chinese translation studies rightly emphasise cultural depth but typically lack an analytic apparatus for exploring cross-modal interaction dynamics. To date, no study has systematically merged the metafunctional perspective of SFL (to ask what meanings emerge) with the intermodal relations taxonomy (to ask how modes interact) in the holistic scrutiny of a film like Chang'an.

This study addresses this gap by operating at two levels. Firstly, deploying intermodal relational analysis to reveal how subtitle, image, and sound interact (redundantly, complementarily, contradictorily, anchoring), and then layering on top of that a metafunctional reading (ideational, interpersonal, textual) that shows what semantic, attitudinal, and coherence-level meanings are

(re)constructed or transformed in translation. Only such an integrative multimodal lens can do justice to the complexity of Chang'an's semiotic fabric.

3. Research design

This chapter outlines the study's theoretical framework, research data, and analytical procedures.

3.1. Theoretical framework

This study is anchored in a synthesis of Halliday's SFL and the model of intermodal relations advanced by Perego and Díaz-Cintas [7,11]. This combined methodology enables a two-level analysis-the linguistic meaning-making via subtitles, and the interplay of these selections with other semiotic means of the film.

The ideational meta-function provides the framework of how Tang history, social structures, and embedded poetry are encoded in the Chinese and re-encoded in the English subtitles based on the three metafunctions of SFL, emphasizing the translation of historical and metaphorical content.

The interpersonal meta-function, which deals with the question of interpersonal relationships, discusses the expression of linguistic markers of power, solidarity, and emotion in the process of translation, particularly in conjunction with paralinguistic aspects of tone and facial expression, in the preservation of meaning and interpersonal subtlety across languages and media

The textual meta-function, which concerns structuring the discourse into a unified whole, enlightens the evaluation of the information flow control and rhythmic construction of poetic sequences in the constrained space of English subtitles.

Four intermodal relations explain these meta-functional analyses. Redundancy happens when subtitles reinforce information presented previously by other means (visuals or audio). Complementation is observable when subtitles complete the cultural or narrative information lacking in different modes. There is yet another contradiction of the subtitles with the audiovisual context and they can be culturally misconstrued. Anchorage assists viewers in being appreciative of ambiguous or symbolic imagery. Such a multidimensional construct enables the comprehensive study of the performance of the linguistic translation, along with the workings of the subtitles and the multimodal assembly of the film in the establishment of the interpretative experiences of transnational moviegoers, how language, image, and sound interrelate in the meaning-making processes across the cultures.

3.2. Research data and analysis procedures

The main data of this analysis includes a carefully collected and purged bilingual corpus of subtitles of the film, the total of which is 25,869 words. This data is the empirical ground of a multi-stage analytical process that was based on studying the interaction of linguistic translation and multimodal semiosis.

The former is the subtitle corpus metafunctional analysis. The step compares and contrasts the original Chinese talk and the English one in subtitles, which are compared and contrasted with the three SFL metafunctions.

The second step, in turn, extracts the results of the linguistic analysis into the multimodal movie setting. At this stage, the elements of the translated subtitles, which have been identified in the first step, are recontextualised in distinct audio-visual scenes. Such language decisions are to be interpreted in the context of intermodal relations.

The final analytical step is the selection and in-depth exemplification of representative cases. From the broader patterns identified in the previous steps, specific, clear instances that typify redundancy, complementation, contradiction, and anchorage will be selected for a fine-grained discussion in the findings section.

4. Analysis and discussion

4.1. Metafunctional analysis of Chinese-English subtitling

This section presents a qualitative analysis of the corpus, focusing on the three meta-functions. Table 1 presents the frequency distribution of strategy shifts for each meta-function.

Table 1. Statistical distribution of metafunctional shifts in subtitling

Metafunction	Sub-category	Frequency	Percentage
Ideational	Culture Specific Items	52	46.4%
	Historical Term Literal Translation	31	27.6%
	Poetic Image Reconfiguration	29	25.8%
	Appellation System Adjustment	17	13.7%
Interpersonal	Modal Particle Omission	12	9.7%
	Affective Metaphor Explicitation	9	7.3%
Textual	Info-Structure Reorganization	25	20.2%
	Cohesive Device Simplification	20	16.1%
	Thematic Structure Explicitation	14	11.3%

4.1.1. Ideational meta-function

The ideational Meta-function in Chang’an involves reconstructing historical-cultural semiosis. The Tang Dynasty history and poetics contain rich and profound semiotic complexity, forcing the translator to negotiate between cultural authenticity, untranslatability, temporal constraints, and readability in the film, which was quite challenging. The translation frequently employs naturalization, exploitation, and literal translation methods, following Aixelá’s ways of translating culture-specific items (CSI) [15].

For instance, “湖人” (hú rén) is rendered as “northern tribesmen”, enhancing geographic clarity but erasing the Tang-era ideological nuance of “Hu” as a term laden with frontier tension and civilizational contrast. Likewise, “节度使” (jié dù shǐ) is translated into “Governor,” replacing a familiar Western administrative title for the military-governor role in the Tang Dynasty. It compresses the institutional distinctiveness of Tang governance into a generic frame. The Confucian concept of “知己” (zhī jǐ), representing a genuine soul-mate who truly understands each other, is delivered as “Do not fear loneliness on the road ahead,” universalizing the sentiment while detaching it from its cultural-ethical roots. Such cross-cultural modulation recasts how the ideational meta-function operates under audiovisual translation’s pragmatic pressure.

4.1.2. Interpersonal meta-function in multimodal subtitling

The interpersonal meta-function concerns how linguistic and visual resources construct social relations and emotional stance jointly. In multimodal texts, these interact dynamically with gesture,

gaze, and visual framing, which will be mentioned in more detail in 4.2. The Tang-era honorifics and kinship appellations such as “高兄” (Gāo Xiōng) and “节度使” (jié dù shǐ) are translated as “Brother Gao,” and “Sir.” While lexically faithful, these renderings perform a pragmatic flattening (see Figure 1).

In the scene where Gao Shi is addressed as “剑南西川节度使” with the English “Southwest Sichuan Governor”, it simplifies the Tang administrative hierarchy into an accessible modern political term. While this lexical compression erases the feudal-military nuance of “节度使”, the accompanying visual frame compensates through multimodal anchorage: Gao Shi’s ornate armour, solemn expression, and the flaming fortress in the background visually reinstate the gravity and militarized authority implicit in the original title. Thus, the loss of interpersonal deference in text is balanced by a visual reinstatement of status and ethos.

Similarly, the translation “高兄” retains lexical fidelity but shifts from Confucian seniority to Western camaraderie. Nevertheless, the multimodal ensemble visually maintains the hierarchy, which is the bowing, the low posture, and the respect. Although the English subtitle nullifies it for universal viewers, the picture reinstates the manners of the Tang era, and accessibility and authenticity coexist.



Figure 1. Interpersonal interpretation in the name of “Gao Shi”

4.1.3. Textual meta-function analysis

The textual meta-function organizes information to maintain coherence across modes. In Chang’an, poetic rhythm and discourse flow are restructured for intercultural accessibility, as illustrated by two examples:

Example 1: In Gao Shi’s Farewell to Dong Da (别董大)

Source Text (ST): 莫愁前路无知己/天下谁人不识君

Target Text (TT): Do not fear loneliness on the road ahead, For who in this world does not know you?

The initial parallelism is turned into a cause-and-effect relationship, and formal symmetry is changed into semantic development. The intensity of emotions and English narrative logic dominate, and the unity of the image is formed due to the edits’ rhythm and musical continuity.

Example 2: The Yellow River scene

ST: 黄河之水奔流到海不复回

TT: The Yellow River racing to the sea, never to return

The translation intensifies the inter-semiotic cohesion, depicting the river’s motion, replacing linguistic linkage, embodying temporal flow and existential reflection. Thus, textual meaning emerges from the synergy of language, image, and rhythm, where visual pacing compensates for linguistic compression and sustains the poem’s metaphoric coherence.

4.2. Intermodal relations analysis: negotiating meaning across semiotic systems

Based on meta-functional analysis, this section examine the interaction of the English subtitles and the visual and auditory modes in film. Table 2 provides an overview of the determined intermodal relations based on a systematic analysis of the corpus and indicates the subtitles as key semiotic forces that affect the multimodal experience.

Table 2. Distribution of intermodal relations in Chang’an English subtitling

Intermodal Relation	Frequency	Percentage	Primary Function in Context
Redundancy	78	52.0%	Reinforcing historical facts, character emotions, and poetic imagery
Complementation	48	32.0%	Scaffolding cultural & historical context; explaining visual metaphors
Anchorage	20	13.3%	Disambiguating symbolic imagery and clarifying temporal shifts
Contradiction	4	2.7%	Highlighting character irony and socio-cultural untranslatability
Total Instances	150	100%	

4.2.1. Redundancy: reinforcing the historical and poetic fabric

Redundancy is a reaffirmation of the meaning by means of semiotics. Redundancy makes information more coherent, as Perego argues that information is spread across the visual, verbal, and auditory channels, so there is consistency in the interpretation when language compression is needed in subtitling [16]. In the frame captioned “轻舟已过万重山”, the translator translated it to mean My little boat passes by ten thousand peaks, as contrasting words were used, little and mountains, to emphasize the broadmindedness of Li Bai. The English subtitles are relatively short and prosaic, and it is impossible to reproduce the rhythmic cadence and historical sound of the Tang poem. However, the synchronized visuals, an expansive shot of misty mountains and the swelling orchestral score restore the grandeur and transcendence embedded in the original line. The convergence of semiotic cues thus compensates for the linguistic reduction. The redundancy between image, sound, and subtitle enables this equivalence beyond the verbal layer, sustaining the poem’s emotive and cultural density (see Table 3).

Table 3. Redundancy interpreted in the frame

Screen Shots	
Time Frame	02:34:22,958 --> 02:34:29,957
Chinese	轻舟已过万重山
English	My little boat passes by ten thousand peaks.

4.2.2. Complementation: scaffolding cultural and historical context

As the second most frequent relation, complementation is the primary mechanism for bridging the cultural knowledge gap. It frequently works with the ideational meta-function, supplying essential context that the visuals and original dialogue assume but the target audience lacks.

A prime example is the recitation of Gao Shi's poem Farewell to Dong Da. It used limited universalization ("twilit sky" for culturally dense imagery) and synonymy ("know you" for the honorific "君"), prioritizing the poem's core interpersonal function for comforting over literal fidelity. On one hand, the related reframes do not show the responding pictures or explanation of 董大, since the poem here is borrowed to indicate friendship enduring personal and political failures between Gao shi and Li Bai. On the other hand, the multimodal elements, including visual mode shifts from close shots showing Gao Shi's facial expression, conveying solemnity and sorrow, to high-angled long shots, depict a desolate, snowy barrack. The auditory mode features Gao's voice, which is heavy with emotion, reciting the lines in Chinese. The subtitle does not merely translate the words but reinforces the sombre mood and the poem's thematic essence (see Table 4).

Table 4. Complementation interpreted in the frames

Screen Shots		
Time Frame	00:09:40,042 --> 00:09:46,457	00:09:57,875 --> 00:10:03,416
Chinese	千里黄云白日曛	天下谁人不识君
English	Sunset's clouds color a twilight sky	For who in this world doesn't know you?

This intermodal synergy serves as the crucial semantic bridge, enabling the international viewer to access the core meaning of these poems, expounding the ideational, interpersonal, and textual meaning in multimodal ways.

4.2.3. Anchorage: disambiguating symbolism and temporal flow

Anchorage in the film *Chang* anchors and fixes the interpretation of cultural and linguistic elements into the mainstream direction of the plot, guiding and controlling the illustrations of other modes. Here, it stabilizes Chang'an as both a geographical and metaphysical space. The line “只要诗在，书在，长安就在” epitomizes the film's anchoring mechanism, converting dispersed memories into a single cultural thesis (see Table 5).

Table 5. Anchorage interpreted in the frames

Screen Shots			
Time Frame	02:38:07,667 --> 02:38:09,041	02:38:09,625 --> 02:38:10,916	02:38:11,792 --> 02:38:15,666
Chinese	只要诗在	书在	长安就在
English	so long as poetry exists	and books exist,	Chang'an will still be with us.

(Notes: Table 5 continue)

The current subtitle is faithfully rendered in a foreignized way, as proposed by Venuti, but it fails to deliver the historical connotation and lyrical tone. Its rhythm is prosaic, and “with us” flattens Chang'an's symbolic dual meanings, simplifying the connotation of civilization into the city. However, the multimodal elements assist understanding. The solemn close-up shots of Gao Shi, pale

dawn lighting, and lingering silence, invite a translation mirroring this transcendence: “As long as verse and ink endure, so shall Chang’an.”

This version compresses lexical redundancy, reinstates poetic symmetry, and activates intermodal resonance. The subtitle limitations exist because of the verbal sparseness and nonverbal and emotional changes. For example, the lighting and facial micro-movement from frowning to smiling makes its semantics heavier and richer. Here, therefore, anchorage is tri-modal: verbal condensation, visual stillness, and aural suspension. These jointly attest to the film’s main conclusion that culture maintains civilization, not time.

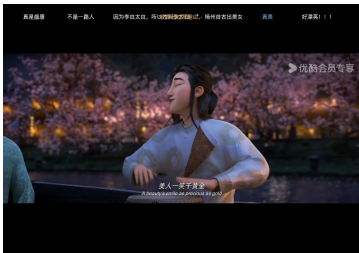
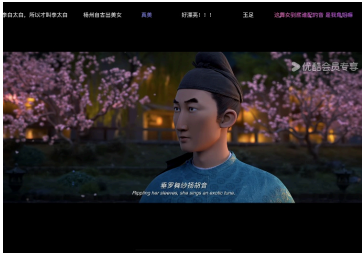
4.2.4. Contradiction: revealing irony and untranslatability

Contradiction is rare in the film Chang’an but creates powerful semantic tension, highlighting character flaws or cultural gaps. These instances frequently involve a clash between the interpersonal dynamics of the dialogue and the visuals.

In a scene depicting Gao Shi and Li Bai in a state of drunken reverie, the poetic couplet “美人一笑千黄金，垂罗舞纱扬胡音” is recited against visuals of dishevelled robes, flickering lamplight, and swaying movements. The English subtitles kept rhythmic elegance and prosodic symmetry, combining literal translation and naturalization to preserve otherness without alienating non-Chinese audiences.

The visual and audible translation turns dissonance into semiotic irony, an artificially musical reflection of cultural realization. Through this, contradiction works as a translation strategy. Rather than curing the difference between modes, it redefines it as an expressive resource, allowing a foreign audience to experience both the sensual grace and existential fatigue of Tang poetics (see Table 6).

Table 6. Contradictions interpreted in frames

Screen Shots		
Time Frame	00:52:20,625 --> 00:52:23,749	00:52:23,750 --> 00:52:26,916
Chinese	美人一笑千黄金	垂罗舞纱扬胡音
English	A beauty’s smile as precious as gold	Rippling her sleeves, she sings an exotic tune.

(Notes: Table 6 continue)

5. Conclusion

This paper reaffirms the fact that the translation of a culturally clogged movie such as Chang’an involves some serious reengineering of the linguistic roles of the original. English subtitles cannot

but make historical ideas and concepts simple, interpersonal understanding thin, and reorganize the poetic flow to fit the spatial restrictions and understanding demands of the audience.

Most importantly, though, these verbal concessions are bargained out in the multimodal environment of the film. The visuals and sound do not serve as the background but as partners in the meaning-making. Redundancy strengthens important emotions and themes, and complementation gives the cultural context that the text has to leave out. The interpretation of symbolic imagery is directed by the concept of anchorage, and even the rare contradictions can draw the attention to cultural untranslatability.

Thus, successful subtitling in this case is not concerned with the independent linguistic truth, as such, but with the strategic decisions in a system of semiotic integrations. The art of the subtitler consists in being aware of what can be orally summarized since it is reinforced by other means and what has to be bridged to a worldwide audience. Such a combined strategy is essential in mediation of deep cultural heritage, so that the spirit of Chang'an its historical magnificence and poetic soul, can be continued through the combination of word, image, and sound. To practice in the future, it highlights the importance of translators being multimodal analysts, where their linguistic decisions are attractive to the entire sensory/narrative setting of the movie

References

- [1] Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014) *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge.
- [2] Díaz Cintas, J., & Remael, A. (2021) *Audiovisual Translation: Subtitling* (2nd ed.). Routledge.
- [3] Taylor, C. (2016) The Multimodal Approach in Audiovisual Translation. *Target: International Journal of Translation Studies*, 28(2), 222–236.
- [4] O'Halloran, K. L. (2008) Systemic Functional-Multimodal Discourse Analysis (SF-MDA): Constructing Ideational Meaning Using Language and Visual Imagery. *Visual Communication*, 7(4), 443–475.
- [5] Liu, J., Wang, Y. and Chen, H. (2023) Mapping the Knowledge Domain of Multimodal Translation: A Bibliometric Analysis. *Humanities and Social Sciences Communications*, 10(1), 1–15.
- [6] Pérez-González, L. (2014) *Audiovisual Translation: Theories, Methods and Issues*. Routledge.
- [7] Halliday, M. A. K. (1978) *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. Edward Arnold.
- [8] Kress, G. and van Leeuwen, T. (2001) *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Edward Arnold.
- [9] Forceville, C. (2006) Non-verbal and Multimodal Metaphor in a Cognitivist Framework: Agendas for Research. In G. Kristiansen, M. Achard, R. Dirven, and F. Ruiz de Mendoza Ibáñez (Eds.), *Cognitive linguistics: Current Applications and Future Perspectives* (pp. 379–402). Mouton de Gruyter.
- [10] Zhang, D. (2009) On a Synthetic Theoretical Framework for Multimodal Discourse Analysis. *Foreign Languages in China*, 6(1), 24–30.
- [11] Perego, E. and Díaz-Cintas, J. (2012) Subtitling: Theory, Practice and Research. *Perspectives*, 20(4), 479–487.
- [12] Ramos Pinto, S. and Mubarak, A. (2019) Multimodal Corpus Analysis of Subtitling: The Case of Non-Standard Varieties. *Target*, 32(3), 389–419.
- [13] Li, X. (2023) Cultural Loss and Compensation in the English Subtitling of Chang'an. *Journal of Chinese Film Studies*, 3(2), 45–60.
- [14] Zhang, Y. (2024) Translating the Sublime: A Functionalist Approach to Poetry in Chang'an. *Translation and Interpreting Studies*, 19(1), 88–105.
- [15] Aixelá, J. F. (1996) Culture-Specific Items in Translation. In R. Álvarez and M. Carmen-África Vidal (Eds.), *Translation, Power, Subversion* (pp. 52–78). Multilingual Matters.
- [16] Perego, E. (2009) The Codification of Non-Verbal Information in Subtitled Texts. *The Translator*, 15(2), 257–280.