Advances in Social Science and Culture
ISSN 2640-9682 (Print) ISSN 2640-9674 (Online)
Vol. 8, No. 2, 2026
www.scholink.org/ojs/index.php/assc

Original Paper

A Quantitative Analysis of Image-Text Dynamics and Strategic Choices in Multimodal Advertising

Juan Yao[1*], Yiman Wang[1], & Hui Shao[1]

[1] College of Foreign Languages, Beijing University of Technology, Beijing, China

[*] Corresponding author

Received: February 26, 2026 Accepted: April 10, 2026 Online Published: April 29, 2026

doi:10.22158/assc.v8n2p192 URL: http://dx.doi.org/10.22158/assc.v8n2p192

Abstract

In the era of digital globalization, the paradigm of transnational advertising has shifted from simple linguistic conversion to sophisticated multimodal orchestration. This study explores the provocative question: Is traditional translation theory still relevant in an age of visual dominance? Utilizing DJI's global social media advertisements as a case study, the research employs a mixed-methods approach grounded in Eco-translatology. A four-stage empirical design was implemented, involving a parallel corpus (N=202) and a technical pipeline that integrates GPT-4V and Sentence-BERT to quantify "modal contribution" through an ablation study paradigm. The findings reveal a pronounced "visual-dominance, textual-marginalization" logic: while image semantic contribution significantly and positively predicts communicative popularity, textual contribution exerts a significant negative effect. Crucially, hierarchical regression analysis demonstrates that after controlling for macro-ecological constraints—such as cultural distance and market GDP—translation strategy (domestication vs. foreignization) per se exerts no independent significant effect on communicative outcomes. These results challenge the traditional text-centric focus of translation studies, suggesting that the visual mode now bears the primary "semantic load." Methodologically, this research offers a replicable computational paradigm for the digital humanities, providing data-driven insights for global marketing managers to shift strategic focus from textual micro-adjustments to upstream multimodal content design.

Keywords

multimodal translation, eco-translatology, DJI, modal contribution, advertising effectiveness

1. Introduction

In the era of digital globalization, the nature of transnational advertising communication has undergone a profound transformation. It has evolved from simple linguistic conversion into a sophisticated multimodal orchestration, where semiotic modes—including images, text, and sound—work synergistically to construct brand identities. This paradigmatic shift renders the traditional pursuit of "textual equivalence" in translation theory increasingly inadequate for explaining contemporary realities. The core of this inquiry lies in a provocative question: In the age of pervasive digital and multimodal communication, is traditional translation theory still relevant? This is particularly evident for high-tech brands like DJI, a pioneer of China's global expansion, whose advertisements must simultaneously convey precise technical parameters and cultivate emotive lifestyle imagery, thereby escalating the complexity of translational decision-making. While current scholarship has approached this phenomenon through the lenses of multimodal analysis, translation strategies, or ecological constraints, these perspectives remain largely fragmented: semiotic analyses excel in qualitative description but lack quantitative rigor; translation studies often examine textual strategies in isolation from their digital ecosystems; and macro-cultural studies frequently fail to capture the specific mechanisms of intermodal interaction.

Drawing upon Eco-translatology as the primary theoretical lens, this study re-conceptualizes translation as an act of "adaptive selection" performed by translators within a multimodal digital ecosystem. To address this overarching question, we must recognize that modern audiences do not consume translated text in a vacuum; they experience a "comprehensive information flow" where translation is but one variable among many, including visual dominance and macro-ecological constraints. Consequently, this study adopts a progressive logic to "deconstruct" the impact of translation through three interconnected research questions:

1. What are the predominant translation strategies employed in the cross-linguistic advertisements of DJI on social media, and what is their frequency distribution?
2. In the context of global advertising, which factor exerts a more significant impact on communicative effectiveness: image contribution or textual contribution?
3. After disentangling the influence of ecological constraints, do translation strategies remain a critical variable in determining communicative outcomes?

To systematically address these questions, this study employs a four-stage empirical design that follows a logic of "variable isolation.". Stage I involves the construction of a parallel corpus comprising 202 multimodal advertisements, where translation strategies are operationalized as a 7-point continuum ranging from foreignization to domestication. This stage utilizes descriptive statistics to provide a preliminary answer to RQ1 and establishes the data foundation for subsequent inferential analysis. Stage II introduces the ablation study paradigm from computational linguistics, utilizing GPT-4V and Sentence-BERT to quantify the respective contributions of image and text modes, directly addressing the comparison in RQ2. Stage III identifies and quantifies ecological constraint variables to isolate their

confounding effects, thereby focusing on the independent effect of translation strategies as posed in RQ3. Finally, Stage IV utilizes hierarchical regression models to integrate modal contribution, translation strategies, and ecological constraints, comprehensively testing whether translation strategy is a primary driver of effectiveness or if its influence has been subsumed by visual and environmental imperatives.

Theoretically, this study provides the first quantitative evidence revealing a "visual-dominance, textual-marginalization" communication logic in the global advertising of high-tech brands. The findings indicate that while image semantic contribution significantly and positively predicts popularity, textual contribution exerts a significant negative effect. This counter-intuitive finding challenges the traditional text-centric paradigm in translation studies and provides empirical support for the tenets of Eco-translatology regarding "multimodal niches." Specifically, it suggests that the visual mode bears the core "semantic load" in technology marketing, thereby granting translators greater "ecological agency."

Methodologically, this research offers a replicable paradigm for the "computational turn" in the humanities and social sciences. By integrating Large Multimodal Models (GPT-4V) with semantic embedding models (S-BERT), we achieved a robust and replicable measurement of the key theoretical construct—"modal contribution." This approach successfully transforms abstract semiotic relationships into continuous variables suitable for statistical modeling, providing an operational toolkit for future cross-modal inquiries.

Practically, this study provides data-driven insights for global marketing managers. A critical finding is that after controlling for external factors such as cultural distance and economic development, translation strategy (the degree of domestication) per se exerts no independent significant effect on communicative outcomes. This implies that for technology-driven brands like DJI, decision-makers need not over-index on the "literal vs. creative" debate at the textual level. Instead, strategic focus should be shifted upstream to the design phase of multimodal content, ensuring that visual elements can independently and sufficiently convey core messages. Simultaneously, the confirmed influence of macro-ecological constraints (e.g., cultural distance, market GDP) provides a reliable basis for resource allocation.

Thesis Organization

Chapter 1 outlines the research background, objectives, and significance, while explicitly articulating the core research questions. Chapter 2 provides a systematic literature review of traditional and multimodal translation studies, establishing the applicability of the Eco-translatological perspective and consolidating the theoretical framework. Chapter 3 details the mixed-methods approach employed in this study, covering the operationalization of variables based on the parallel corpus and the technical pipeline for quantifying modal contribution via computational linguistics. Chapter 4, the empirical core of the thesis, reports the results of descriptive statistics, regression diagnostics, and hierarchical regression analysis. It further provides an in-depth discussion on the "marginal effect" of translation

strategies within multimodal contexts. Chapter 5 synthesizes the core findings, discusses their scholarly implications for the paradigmatic restructuring of translation studies, and objectively evaluates the study's limitations while proposing avenues for future research.

2. Literature Review

2.1 Traditional Paradigms in Translation Studies

In traditional translation research paradigms, the exploration of "translation products" or "outputs" has consistently centered on translation strategies as the primary analytical path. Researchers generally regard the translator's strategic choices as the pivotal independent variables that determine the final form of the translation and its reception within the target culture.

This theoretical trajectory originated with the systematized comparative linguistic approach of Vinay and Darbelnet (1995). Their proposed taxonomy of translation procedures—such as *calque*, *transposition*, and *modulation*—aimed to reveal the specific tactical means by which translators achieve "equivalence" through strategic contrast (Chesterman, 2000). Subsequently, Nida (1964) further theorized this logic through the dichotomy of "formal equivalence" versus "dynamic equivalence." This marked a significant shift in scholarly focus from static linguistic comparison to a perspective that adopts reader response as the ultimate criterion, profoundly influencing subsequent inquiries into the intelligibility and acceptability of translated texts. The German Functionalist School (Reiss & Vermeer, 1984; Nord, 1991) elevated this causal logic to an explicit level. Their Skopos theory explicitly postulates that the translation purpose (Skopos) dictates the selection of strategies and the resulting textual profile, defining Nord's (1991) principle of "function plus loyalty" and the "circular model" as a deliberate decision-making process to achieve intended effects. Chesterman (2000) synthesized and elevated these perspectives, noting that earlier comparative and process models remain "open to a causal interpretation." He advocated for the construction of explicit Causal Models, positioning translation strategies as antecedent conditions that trigger cognitive, affective, and behavioral responses in readers. By doing so, translation studies transcends descriptive "what" questions to systematically explore explanatory "why" and predictive "what if" inquiries.

Consequently, within the traditional paradigms dominated by the comparative and process models as categorized by Chesterman, translation strategies—whether conceptualized as Vinay and Darbelnet's "procedures," Nida's "equivalence principles," or the Functionalists' "Skopos-oriented" methods—occupy a central position in the analysis. These diverse studies converge on a fundamental theoretical consensus: translation is not a value-neutral linguistic conversion; rather, the translator's strategic choices serve as the core drivers that shape the final manifestation of the translation and prefigure its socio-cultural reception. This consensus provides the essential theoretical grounding for subsequent empirical research, particularly for testing the actual impact of specific translation strategies (e.g., domestication vs. foreignization) on readers' cognition, aesthetics, and overall acceptance.

2.2 Multimodal Translation

Contemporary translation practice is undergoing a profound transformation from traditional linguistic conversion to a multimodal turn (Chen, X., Pan, H., & Li, P. (2020). As Pan and Wu (2024) contend, translation should be conceptualized as a process of "transmodal resemiotization," involving the synergistic interaction of diverse semiotic modes such as images, colors, and typography. The crux of this cognitive shift lies in re-envisioning translation as a "meaning-making process situated within multimodal contexts" (Baker & Pérez-González, 2023). This paradigm finds its theoretical roots in social semiotics and systemic functional linguistics, where a "mode" is defined as a resource system for realizing communicative intent (Chandler, 2002), and a multimodal text represents an integrated mobilization of these diverse resources (Kress & van Leeuwen, 2001). Stöckl (2004) further elucidates the micro-level choices in media realization through the concept of "sub-modes," which collectively constitute the final manifestation of a mode (Pérez-González, 2007). This compels researchers to investigate how translators reconfigure the overall "semiotic fabric" during intersemiotic shifts. Moreover, Eco-translatology, as proposed by Hu (2011), provides a macro-perspective by interpreting translation as an act of "survival selection" and "adaptive transformation" performed by the translator to suit the "translational eco-environment," which encompasses various multimodal factors.

The introduction of multimodal perspectives has revolutionized empirical research in specialized domains such as audiovisual translation (AVT), theater, and interpreting. In AVT, texts are viewed as "composite products" generated through the combination of various semiotic resources (Baldry & Thibault, 2006). The translator's core challenge lies in transferring visual meaning into written subtitles while respecting the overarching semiotic structure (Chuang, 2006). Non-verbal elements, including lighting, contrast, and visual perspective, significantly influence the generation of the translated text (Maszerowska, 2012; Pérez-González, 2007). Similarly, theater translation research has shifted its focus from written scripts to translation as "performance," emphasizing the conversion of meaning through the interaction of multiple semiotic systems (Baines & Dalmasso, 2007). In dialogue interpreting, non-verbal signs such as gaze and gestures are theorized as "contextualization cues" that regulate turn-taking and participant alignment (Mason, 2014; Davitti, 2012). Furthermore, research on visual paratexts has revealed how non-linguistic resources shape textual reception. Gérard Genette's (1997) paratextual theory has been extended to analyze how visual dimensions, such as illustrations and layout, emerge as results of meaning negotiation within socio-cultural contexts (Bell, 2003) and serve as strategic tools for managing public reception and ideological agendas (de Marco, 2012; Pérez-González, 2013).

As a quintessential multimodal genre, advertising translation is shifting from pure linguistic equivalence (Adab etc, 2004; Valdés, 2013) toward functional adaptation. However, a persistent bottleneck remains in the empirical measurement of cross-modal effects (Bateman, 2014). While marketing scholarship has confirmed the differential effects of image versus text in driving user engagement (Li & Xie, 2020), precisely mapping these differences onto a translator's strategic choices

(e.g., domestication vs. foreignization) remains a critical gap. The rise of digital technology has further complicated this process, as translation in hypertexts and fan-collaboration networks evolves into the rewriting of "pluralized source texts" (Littau, 1997; Ortabasi, 2007). Entering the 2020s, research has pivoted toward causal-empirical inquiries. Chen (2024) developed the "4C framework" (Concurrence, Complementarity, Condensation, Contradiction) to analyze image-text logical relations, proposing that translators guide audience attention through shifts such as "Augmentation" (4E) or "Diminution" (4D). This echoed Kaindl's (2022) view of the translator as a "semiotic mediator" coordinating image-text dynamics, a notion supported by eye-tracking experiments showing that explicitation strategies can enhance cognitive efficiency. Leveraging advancements in computational linguistics for large-scale data modeling (Hessel et al., 2021), this interdisciplinary trajectory from macro-description to micro-empirical validation (Gambier, 2016) provides a rigorous scientific foundation for assessing the functional adaptation of advertising translation strategies.

2.3 Translational Value in Image-Text Collaboration

In synthesis, the introduction of a multimodal perspective has restructured the theoretical framework of translation studies. In today's communication ecology, where digital screens serve as the primary medium, translation is defined as an intersemiotic mediatory practice (Pan et al., 2023). Consequently, any robust analysis of translation strategies must be situated within a composite context woven from multiple semiotic modes.

Despite the broad consensus on this macro-turn, existing scholarship lacks a granular deconstruction of Relative Semantic Contribution. Most multimodal translation research (Liu & O'Halloran, 2009) operates under the assumption that images, text, and sound function as a unidirectional "combined force." However, in cross-cultural marketing, this relationship is often strategically reconfigured. In digital advertising, the debate over the dominance of "visuals" versus "text" is long-standing. Marketing research (Mitchell, 1986; Pieters & Wedel, 2004) suggests a "superiority of imagery" in capturing attention and conveying affective attitudes, while text excels in delivering specific product parameters and functional information. When this discussion enters the ambit of translation studies, the complexity intensifies. For high-tech brands like DJI, translational decision-making is not an isolated linguistic conversion; rather, as conceptualized by Hu (2011) within the Eco-translatology framework, it is an "adaptive selection" made by the translator to achieve "survival of the fittest" within a specific "translational eco-environment." Essentially, whether a consumer's purchase decision is driven by globally unified images featuring "universal visual rhetoric" or by text that has undergone the translator's "three-dimensional transformation" (linguistic, cultural, and communicative dimensions) depends on the energy distribution of various ecological factors within a given niche.

Current literature reviews indicate that while researchers recognize the necessity of image-text interaction, most studies remain at the level of qualitative description (e.g., Marsh & White's [2003] taxonomy of image-text relations). Bateman (2014) pointedly noted the persistent lack of a robust methodology in multimodality to measure the "weight" of each mode in overall meaning construction.

This methodological void makes it difficult for translators to make scientific priority judgments when facing ecological constraints (e.g., layout limitations, cultural taboos): should textual fidelity be sacrificed to accommodate the image, or should the visual layout be adjusted to compensate for the loss of textual information? To address this bottleneck, this study introduces the concept of "Modality Contribution" from computational linguistics. Drawing on the paradigm of Hessel et al. (2021), modality contribution is defined as the independent volume of semantic information carried by a specific mode during multimodal processing. In the context of Eco-translatology, this can be further understood as the "energy allocation" of different ecological factors in achieving a "holistic adaptive selection." By calculating the contribution weights of image and text in expressing core brand semantics, we can quantitatively reveal the hierarchical structure within the translational ecosystem.

Another significant research gap lies in the inability to isolate environmental interference to observe the true utility of translation strategies. In Chesterman's (2002) causal model, the translation effect is viewed as the outcome of strategic choices. However, in real-world transnational advertising, communicative effects are often diluted or amplified by "non-translation variables" such as brand reputation, pricing, and visual aesthetics. While recent empirical studies (Su, 2024) have begun using eye-tracking to observe reader responses to domestication/foreignization, the "net effect" of translation strategies—the specific contribution of the strategy itself to acceptability after controlling for visual appeal and market constraints—remains a theoretical "black box." Without adopting an ablation study paradigm (Hessel et al., 2021) to decouple modal variables, translation studies will struggle to answer: in an advertisement with powerful visual performance, does a suboptimal translation strategy truly lead to communicative failure?

Based on this logic, this study posits that merely defining "what multimodal translation is" no longer suffices to explain complex globalized communication. We must ask: how is the "energy density" of different semiotic resources allocated within the tension of multimodal collaboration? And after stripping away the "visual premium," what is the independent value of translation strategy as an "adaptive selection"? Consequently, this study proposes the following research questions: *What are the predominant translation strategies employed in the cross-linguistic advertisements of DJI on social media? Which factor exerts a greater impact on communicative effectiveness: image contribution or textual contribution? After disentangling ecological constraints, does the translation strategy (domestication vs. foreignization) remain a key variable in determining communicative outcomes?*

3. Methodology

3.1 Research Design Framework

This study employs a mixed-methods research design to investigate the impact of multimodal interaction on translation decisions and their effectiveness within the global social media communication of DJI. The overall design follows a logical trajectory: "Multimodal Ablation — Quantification of Ecological Variables — Hierarchical Regression Validation." To bolster internal

validity, the core design is grounded in a controlled comparison approach. By utilizing the global high-tech brand DJI as a case study and matching its Chinese-foreign parallel corpora, this model achieves constant control over "physical product attributes" and "brand technical parameters" (e.g., ensuring hardware specifications for the Mavic 4 Pro remain consistent across contexts). This setup allows for the precise identification of the "net effect" of translation shifts on communicative outcomes. To ensure the robustness of the empirical findings, the study adopts dual operationalization—cross-validating continuous and binary variables to mitigate the interference of measurement bias on statistical significance.

3.2 Corpus Sources and Sampling Strategy

DJI was selected as a prototypical case of multimodal translation practice within the context of global communication. To ensure the breadth and representativeness of the sample, the research team constructed a multimodal parallel corpus (N=202) spanning from July 2023 to December 2025. Data collection followed the principles of stratified sampling, encompassing major international social media platforms (TikTok, Instagram) alongside core domestic Chinese platforms (WeChat Channels, Weibo).

The fundamental logic of this research design lies in matching Chinese source posts with their corresponding multilingual target versions. This facilitates a controlled comparison that effectively isolates interference from product-level differences, enabling the study to accurately measure translation shifts triggered by multimodal interaction and their subsequent communicative effects. During the data cleaning phase, redundant information—such as image-only posts or those lacking corresponding translated text—was excluded to ensure the purity of the modeling data. The final dataset comprises paired source (Chinese) and target (multilingual) texts, their associated visual assets, and relevant communicative engagement metadata.

3.3 Operationalization of Core Variables

3.3.1Translation Strategy (X_{trans}：Core Independent Variable A)

To meet the rigorous requirements for measurement validity in econometric modeling, this study implements dual measurement and cross-validation for the core independent variable, "translation strategy." First, at the continuous scale level, translation behavior is quantified using a 7-point Likert Scale based on the taxonomical framework of Vinay and Darbelnet (1995). The scale anchors "literal translation and foreignization" at the lower end (1) and "complete transcreation and domestication" at the upper end (7), thereby capturing the continuous distributional trends of micro-strategic choices.

Second, to further verify the robustness of the empirical findings, a dummy variable based on translation strategy was constructed. By coding "extreme domestication" (scores of 6–7) as 1 and all other strategies as 0, this study tests the boundary effects of strategies under different model specifications, mitigating potential interference from extreme values on statistical significance. Regarding quality control, two experts with professional backgrounds in translation studies conducted double-blind coding of the corpus. The inter-rater reliability, measured by Cohen's Kappa, was 0.88,

indicating a high degree of consistency and ensuring reliable data input for the subsequent regression analysis.

3.3.2 Multimodal Contribution (X_{modal}: Core Independent Variable B)

A technical innovation of this study is the introduction of an ablation study paradigm from computational linguistics to quantify the "functional burden" of images within the overall semantic contribution. This serves as a key predictor in the hierarchical regression model (Model 1). The extraction of multimodal contribution(X_{modal})follows a logical path of "modal stripping and vectorial comparison": First, the multimodal large language model GPT-4V was utilized to extract deep semantic features from three conditions: "image-only description""($V_{img,}$), "text-only description"(V_{text}), and "combined description"(V_{comb}). Subsequently, the pre-trained model Sentence-BERT (specifically the *all-mpnet-base-v2* variant) transformed these textual descriptions into 768-dimensional dense semantic vectors.

On this basis, the image contribution(X_{img_cont})and textual contribution (X_{text_cont}) are defined by calculating the cosine similarity between the respective semantic vectors:

$$X_{img\_cont} = Similarity(V_{img,} V_{comb}) = \frac{V_{img} \cdot V_{comb}}{|V_{img}|\,|V_{comb}|}$$

$$X_{text\_cont} = Similarity(V_{text,} V_{comb}) = \frac{V_{text} \cdot V_{comb}}{|V_{text}|\,|V_{comb}|}$$

Where represents the dense vector of the "image-only description" extracted via GPT-4V and transformed by Sentence-BERT (S-BERT); represents the dense vector of the "text-only description" extracted via GPT-4V and transformed by Sentence-BERT (S-BERT); represents the dense vector of the "combined visual-textual description".

From the perspectives of translation sociology and eco-translatology, a higherX_{img_cont} value indicates that the visual mode carries a dominant "semiotic weight" in information transmission, generating a "visual premium." This high image contribution grants the translator greater "ecological agency" (or *ecological freedom*), allowing for flexible adoption of domestication or transcreation strategies. Conversely, reflects the functional burden undertaken by the textual mode. An excessively high textual contribution may imply information redundancy or increased difficulty in cross-cultural transmission.

3.3.3 Communicative Popularity (Y_{pop}：Dependent Variable)

This study designates as the dependent variable to measure the cross-cultural effectiveness of

advertising. Drawing on the engagement framework proposed by Bonsón and Ratkai (2013), popularity is operationalized as the ratio of total interactions to the account's follower base, measuring the immediate feedback pressure of audiences toward multimodal content. Considering the functional disparities across various platforms within DJI's global ecosystem, and adhering to the cross-platform standardization recommendations by Voorveld et al. (2018), this study excludes "shares" due to inconsistent accessibility across platforms. Popularity is thus defined as:

$$Popularity = \frac{Likes + Comments}{Followers}$$

During the statistical processing phase, given that social media engagement data typically exhibit a pronounced long-tailed distribution, a logarithmic transformation was applied to the raw indicators to satisfy the normality assumption for linear regression. The final calculation formula is as follows:

$$Y_{pop} = \ln\left(1 + 100 \times \frac{Likes + Comments}{Followers}\right)$$

3.3.4 Ecological Control Variables ($\sum Control_i$: Control Variable)

To isolate the influence of external environmental factors on advertisement popularity and focus on the net effect of translation strategies, this study introduces a set of ecological control variables.($\sum Control_i$)。 First, the composite cultural distance (CD_j) between the target market and China is calculated using the classic measurement model proposed by Kogut and Singh (1988), based on Hofstede's (2001) cultural dimensions. The formula is as follows:

$$CD_j = \sqrt{\sum_{i=1}^{6} \frac{(I_{ij} - I_{i,China})^2}{V_i}}$$

where:

I_{ij} is the score of country jon the i-th cultural dimension,

$I_{i,China}$ is China's score on the i-th dimension,

V_i is the sample variance of the i-th dimension.

The six dimensions covered are: Power Distance Index (PDI), Individualism (IDV), Masculinity (MAS), Uncertainty Avoidance (UAI), Long-Term Orientation (LTO), and Indulgence versus Restraint (IVR).

Furthermore, this study adopts the target market's Gross Domestic Product (GDP) for 2024 as a macro-level indicator of market size and consumption capacity. The raw data are sourced from the

CEIC database and are log-transformed before being included in the regression model to mitigate potential heteroscedasticity and enhance the robustness of parameter estimation.

At the micro level, the study introduces a product code (*Product Number*) to control for baseline fluctuations in popularity due to differences in audience size and product lifecycle across categories. Additionally, given the inherent distinctions in social media algorithms and user profiles, the publishing platform (*Platform*) is coded dichotomously, with Instagram coded as 1 and TikTok as 2. To address the strategic focus of the advertisement content, the communicative intent (*Communicative Intent*) is further categorized and coded as follows: general product promotion (coded 1), brand image/scenario storytelling (coded 2), and new product launch (coded 3). Finally, to capture the physical constraints imposed by platform-specific technical limitations on the space for translation strategies, a constraining attribute variable is introduced. Based on platform rules, Instagram's 2,200-character limit is coded as 1.0, while TikTok's 4,000-character limit serves as the baseline for comparison.

Together, these variables define the external ecological constraints within which translators make "adaptive choices." By incorporating them in a hierarchical regression model, this study aims to more accurately identify the marginal explanatory power of translation strategies on popularity.

3.4 Statistical Model Specification

This study utilizes hierarchical regression models to observe changes in explanatory power (R^2) following the introduction of translation strategies and ecological variables.

Stage 1: Baseline Regression Model (Model 1)

Model 1 aims to address Research Question 1 by examining the underlying drivers of multimodal structures on communicative effectiveness without external interventions. Taking popularity (Y_{pop}) as the dependent variable and the image contribution (X_{img_cont}) and textual contribution(X_{text_cont})extracted from the ablation study as core independent variables, the baseline regression equation is constructed as follows:

$$Y_{pop} = \beta_0 + \beta_1 X_{img\_cont} + \beta_2 X_{text\_cont} + \epsilon$$

Where and represent the predictive coefficients of the visual and textual modes, respectively, on content popularity. By comparing their significance and effect sizes, this study identifies which mode carries the primary "semiotic weight" in DJI's cross-cultural communication within the global social media context.

Stage 2: Full Model (Model 2a)

Model 2a addresses Research Question 2: whether the translation strategy itself remains a critical variable in determining communicative outcomes after disentangling the effects of underlying

multimodal features and external ecological constraints (e.g., GDP, cultural distance). This stage introduces the core independent variable "Translation Strategy" (X_{trans}) and a series of ecological control variables ($\sum Control_i$) into the baseline model.

The regression equation is specified as:

$$Y_{pop} = \beta_0 + \beta_1 X_{img\_cont} + \beta_2 X_{text\_cont} + \beta_3 X_{trans} + \sum \gamma_i Control_i + \epsilon$$

In this specification, X_{img_cont} and represent the foundational multimodal features carried over from Model 1. The retention of these variables is intended to verify the robustness of the Visual Premium and textual load within a multivariate environment, thereby establishing the underlying semiotic constraints under which translation strategies operate. The core independent variable, X_{trans}, denotes the translation strategy measured on a 7-point scale. Furthermore, $\sum Control_i$ encompasses a suite of ecological control variables—including macro-economic indicators (GDP), the Hofstede Cultural Distance Index, and product attributes. By incorporating these controls, the model effectively isolates external interference, facilitating the precise identification of the "net effect" of translation strategies on communicative outcomes.

By comparing the incremental R-squared(ΔR^2)and the shifts in coefficient significance between Model 1 and Model 2, this study provides a robust empirical answer to Research Question 2.

Stage 3: Robustness Check (Model 2b)

To verify the consistency of the estimated effects across different operational definitions, Model 2b serves as a robustness check by comparing its results with those of Model 2a (continuous measurement). Building upon the full specification of Model 2a, the 7-point scale is re-operationalized as a dichotomous dummy variable, denoted as X'_{trans}. Specifically, "extreme domestication" (scores of 6–7) is coded as 1, while all other strategies are coded as 0. This re-coding allows the study to examine the boundary effects of the translation strategy under different model specifications and to mitigate potential interference from extreme value distributions on statistical significance. The regression equation for Model 2b is specified as follows:

$$Y_{pop} = \beta_0 + \beta_1 X_{img\_cont} + \beta_2 X_{text\_cont} + \beta_3 X'_{trans} + \sum \gamma_i Control_i + \epsilon$$

Where represents the core independent variable "Translation Strategy," specifically operationalized to isolate the impact of extreme domestication.

4. Results and Analysis

4.1 Descriptive Statistics

Following the verification of coding reliability (Cohen's Kappa = 0.88), a descriptive statistical analysis was conducted on the core variables. Table 1 summarizes the key statistical characteristics of the sample($N = 202$).

Table 1. Descriptive Statistics for Core Variables (N = 202)

Variable	Min	Max	M	SD	Skewness	Kurtosis
$\boldsymbol{X_{trans}}$ (1–7)	3.00	7.00	6.00	0.97	-0.70	-0.33
$\boldsymbol{X'_{trans}}$ (Binary)	0.00	1.00	0.73	0.45	-1.03	-0.95

Note: Translation Strategy (X_{trans})is a 7-point Likert scale (1=Foreignization, 7=Domestication). is a dummy variable where 1 represents "Extreme Domestication" (scores 6-7).

Regarding the distributional characteristics of DJI's global advertising translation strategies, the descriptive results yield the following three pivotal insights:

First, the overall translation profile exhibits a pronounced "domesticating tendency." As illustrated in Table 1, the mean score for translation strategy is notably high ($M = 6.00$, $SD = 0.97$), with a minimum observed value of 3.00. This indicates that among the 202 advertisement samples, highly literal or foreignizing strategies (scores 1–2) were virtually absent. The heavy concentration at the upper end of the scale (4–7) clearly demonstrates that in social media cross-cultural communication, DJI favors audience-oriented free translation and domestication to bridge the psychological distance between high-tech products and local consumers.

Second, "extreme domestication" serves as the dominant practical paradigm. As a boundary measure of the domesticating inclination, the dummy variable yielded a mean of 0.73. This suggests that 73% of the samples in DJI's global corpus utilized "extreme domestication" (scores of 6–7). Such high strategic convergence not only underscores the brand's emphasis on content localization but also reflects the high degree of flexibility and "re-creative" potential of textual translation within its multimodal narrative.

Third, the distributional properties satisfy the prerequisites for inferential statistics. For the continuous scale (X_{trans}), the data cluster toward the higher end, exhibiting slight negative skewness (-0.70) and a platykurtic distribution (Kurtosis = -0.33), characterized by a flatter peak than a normal distribution.

According to the empirical criteria proposed by Curran, West, and Finch (1996), data are considered to have no substantial violation of normality assumptions when |skewness| < 2 and |kurtosis| < 7. The indices in this study fall well below these thresholds. Furthermore, while the binary variable shows an asymmetrical distribution (73% vs. 27%), this provides an ideal empirical basis for the subsequent robustness checks in Models 2a and 2b. It facilitates the precise identification of marginal effects and the consistency of strategy-related outcomes across different operational definitions.

To clarify the linguistic nuances of this distribution, Table 2 maps representative cases to the strategy continuum.

Table 2. Representative Cases Mapped to the High-End Range of the Strategy Continuum

Score	Source Text (CN)	Target Text (Localized)	Market
3	自由开拍，跟随精彩	Jede Bewegung im Einklang 5. November 2025 | 13:00 Uhr (MEZ) Mehr erfahren: https://brnw.ch/21wX1n8	Germany
4	#大疆 发布 DJI Air 3S 双摄旗舰旅拍无人机，尽显风光。 一英寸 CMOS 主摄 | 双摄 4K/60fps HDR 视频，14 挡动态范围 | 自由全景照片 | 夜景级全向主动避障 | 新一代智能返航，返航更精准 | 45 分钟续航，20 公里图传 | … DJI Air 3S 汇聚多项先进技术，影像实力跃上新高度，慧眼成双，尽显专业影像质感。不论是旅途的风光美景，还是你的高光时刻，都能精彩呈现。	旅先での撮影に最適なデュアルカメラドローン、DJI Air 3S が新登場。 1 インチ CMOS 搭載のメインカメラと 70mm 中望遠カメラを備え、両カメラとも最大で 14 ストップのダイナミックレンジに対応し、4K/60fps HDR 動画撮影が可能なこのカメラドローンは、風景やポートレートなどを、細部まで驚くほど鮮明な映像で撮影します。 夜間撮影にも対応する障害物検知、次世代スマート RTH、10-bit 04 映像伝送技術によって、安全性が向上し、クリアなライブビューを見ながら、安心して空撮を楽しむことができます。 新機能フリーパノラマ、アップグレードした ActiveTrack 360°、最適化されたトラッキング技術を備えたこのドローンは、クリエイティブな撮影に最適です。DJI Air 35 で、まだ見ぬ景色を見つけに行こう。 詳しくはプロフィールのリンクから♪	Japan
5	大疆发布扫地机器人 ROMO 系列：优雅设计、实力绝尘 优雅未来感设计，提升家居格调丨旗舰无人机感知系统，毫米级避障，0	Meet DJI ROMO Series - Make a Clean Sweep Introducing DJI ROMO Series, Flagship Robot Vacuum With Advanced Sensing DJI ROMO Series will instantly elevate your home. It combines cutting-edge perception and electromechanical technology inspired by flagship drones. With advanced	Official

	干预通行丨25000 Pa 绝尘吸力，席卷脏污丨柔性双机械臂，100% 贴边无死角丨164 ml 超大容量机载水箱，持续湿拖清除重污丨0 缠绕双滚刷，无惧超长毛发，免维护丨独创高压自冲洗技术，基站 200 天免维护丨独创消音舱设计，过滤 80% 噪音 旗舰扫地机器人 DJI ROMO 系列，具备旗舰无人机的精密感知与机电技术，高效覆盖地板和地毯，洁净无死角。加上省心自清洁功能，无需干预，自动包办全流程。强劲性能，助你坐享家中雅静。	sensors and smart algorithms, ROMO delivers deep, thorough cleaning on both hard surface floors and carpets, leaving no spot neglected. Additionally, its intelligent self-cleaning system takes care of maintenance automatically, freeing you to enjoy a clean home with minimal effort. Shop now—just click the Linktree in our bio!	
6	大疆发布 DJI Mini 5 Pro 一英寸大底全能迷你航拍机，小有底气 一英寸大底相机丨夜景级全向主动避障丨支持无损竖拍及 225° 云台旋转，运镜更自由丨全向智能跟随再升级丨轻巧便携丨可选 36/52 分钟续航 DJI Mini 5 Pro 将一英寸 CMOS 相机、夜景级全向主动避障、225° 旋转高自由度云台以及更出片的全向智能跟随，连同生活中的美好时刻，统统装进	Découvrez DJI Mini 5 Pro – Le drone caméra mini tout-en-un avec capteur CMOS 1 pouce grand format DJI Mini 5 Pro intègre une caméra avec CMOS 1 pouce, une détection d’obstacles omnidirectionnelle en paysage nocturne, une nacelle avec une rotation flexible à 225°, et ActiveTrack 360° optimisé, le tout dans son boîtier léger emblématique. Avec des performances comparables à celles des drones phares, il capture les plus beaux moments de la vie et préserve vos précieux souvenirs avec une richesse de détails époustouflante. ✓ Caméra avec CMOS 1 pouce grand format ✓ Détection d’obstacles omnidirectionnelle en paysage nocturne ✓ Véritable prise de vue verticale et rotation flexible de la nacelle à 225° ✓ ActiveTrack 360° amélioré	France

	依旧轻巧的机身中，用媲美旗舰机型的强大性能，为你定格珍贵回忆。	✓ Léger et portable ✓ Autonomie de la batterie prolongée En savoir plus : https://brnw.ch/21wVP2H Acheter maintenant : https://brnw.ch/21wVP2M	
7	大疆发布 DJI Mini 5 Pro 一英寸大底全能迷你航拍机，小有底气 一英寸大底相机丨夜景级全向主动避障丨支持无损竖拍及 225° 云台旋转，运镜更自由丨全向智能跟随再升级丨轻巧便携丨可选 36/52 分钟续航 DJI Mini 5 Pro 将一英寸 CMOS 相机、夜景级全向主动避障、225° 旋转高自由度云台以及更出片的全向智能跟随，连同生活中的美好时刻，统统装进依旧轻巧的机身中，用媲美旗舰机型的强大性能，为你定格珍贵回忆。	Descubre DJI Mini 5 Pro – Dron mini con cámara todo en uno con CMOS grande de 1 pulgada DJI Mini 5 Pro cuenta con una cámara CMOS de 1 pulgada, detección de obstáculos omnidireccional en escenas nocturnas, un estabilizador con 225° de rotación flexible y ActiveTrack 360° mejorado, todo dentro de su característico cuerpo ligero. Con un rendimiento a la altura de los drones más emblemáticos, captura los mejores momentos de la vida y preserva tus recuerdos más preciados con un nivel de detalle impresionante. ✓ Cámara CMOS grande de 1 pulgada ✓ Detección de obstáculos omnidireccional en escenas nocturnas ✓ Filmación vertical real y rotación flexible del estabilizador de 225° ✓ ActiveTrack 360° mejorado ✓ Ligero y portátil ✓ Vida de la batería aumentada Conocer más https://brnw.ch/21wVPgt Comprar ahora https://brnw.ch/21wVPgx	Latam

Note. No instances of Score 1–2 were observed as dominant strategies in the corpus.

As evidenced, the progression from Score 3 to 7 reflects a systematic shift from maintaining source-text technicality to pursuing target-context resonance. The absence of Borrowing or Calque (Scores 1–2) suggests that for technology-driven brands, mere linguistic "transplantation" is insufficient in the translational eco-environment of social media.

At the extreme end of the continuum (Score 7), translation evolves into Adaptation. In the dji_latam sample, the textual structure is completely re-situated to suit the digital ecosystem of TikTok. The original's structured list of technical parameters is transformed into a conversational, emoji-laden narrative (using ✓, ,). This 'extreme domestication' reflects the tenets of Eco-translatology (Hu, 2011), where the translator transforms the 'translational niche' to survive within a visually-saturated, fast-paced social media environment. The text no longer merely 'translates' the source but 'co-creates' a new communicative situation aligned with the local multimodal semiotic norms.

4.2 Hierarchical Regression Results

This study employed hierarchical regression modeling to systematically examine the impacts of multimodal contribution, translation strategies, and ecological constraint variables on communicative popularity(Y_{pop}). Table 3 summarizes the regression coefficients, standard errors, and overall goodness-of-fit indices for each model. The analysis began with a baseline model to assess the direct effects of image and textual contributions. Subsequently, the core independent variable—translation strategy—was introduced after controlling for a suite of ecological and content variables to examine its net effect by isolating confounding factors.

Table 3. Hierarchical Regression Results for Variables Predicting Popularity (N = 202)

Variable	Model 1	Model 2a	Model 2b
Step 1: Core Multimodal Variables			
Constant	0.522* (0.226)	0.389 (0.405)	0.312 (0.324)
Image Contribution (X_{img_cont})	0.669** (0.222)	0.504* (0.230)	0.507* (0.230)
Text Contribution (X_{text_cont})	-0.823** (0.292)	-0.671* (0.298)	-0.676* (0.298)
Step 2: Translation Strategy			
Translation Strategy (1-7) (X_{trans})		-0.020 (0.040)	
Translation Strategy (Dichotomy) (X'_{trans})			-0.065 (0.087)
Step 3: Control Variables			
Cultural Distance		0.005*** (0.001)	0.005*** (0.001)
GDP		0.018* (0.007)	0.018* (0.007)
Product Number		0.003 (0.010)	0.003 (0.010)
Platform		-0.103 (0.080)	-0.102 (0.080)
Communicative Intent		-0.055 (0.063)	-0.056 (0.063)
Constraining Attribute		0.052 (0.084)	0.055 (0.084)
Model Summary			
R	0.248	0.362	0.364
R^2	0.061	0.131	0.133
Adj. R^2	0.052	0.091	0.092
ΔR^2 (vs. Model 1)		0.070	0.072

F for ΔR^2		3.226*	3.265*
Durbin-Watson	1.895	1.874	1.873

Notes: 1. N = 202. The dependent variable is ln(1+100*popularity). 2. Unstandardized coefficients are reported, with standard errors in parentheses. 3. Platform is a binary variable (0 = A, 1 = B). 4. *p < .05, p < .01,*p < .001.

4.2.1 Impact of Multimodal Contribution on Popularity

First, a baseline regression model (Model 1) was constructed to address RQ2 regarding the influence of image contribution (X_{img_cont}) and textual contribution (X_{text_cont}). As indicated in Table 3, the model is statistically significant ($F = 6.519, p = 0.002$), accounting for approximately 5.2% of the variance ($\text{Adjusted } R^2 = 0.052$).

The results support the hypothesis that image contribution is a significant positive predictor of popularity ($B = 0.669, SE = 0.222, p < 0.01$). This suggests that the visual mode bears the primary "semiotic weight" in DJI's advertisements, effectively driving engagement among cross-cultural audiences. Conversely, textual contribution exerts a significant negative impact on popularity($B = -0.823, SE = 0.292, p < 0.01$). This finding aligns with the theoretical expectations of "ecological agency": when the text is burdened with excessive functional weight, the information may become redundant or difficult to process cross-culturally, thereby inhibiting content dissemination. Correlation analysis($r = 0.309$)and Variance Inflation Factors ($\text{VIFs} \approx 1.1$)indicate that multicollinearity does not pose a threat to the model's integrity.

4.2.2 The Effect of Translation Strategy under Ecological Constraints

To address RQ3, the core independent variable—translation strategy—along with ecological and content control variables, was introduced to test whether the degree of domestication independently influences communicative outcomes. A dual-operationalization approach was used for robustness: Model 2a utilized the continuous 7-point Likert scale(X_{trans}), while Model 2b employed a dichotomous dummy variable(X'_{trans}, representing "extreme domestication"). The hierarchical regression results (see Table 3) show that both full models (2a and 2b) are statistically significant ($p = 0.001$), with an improved explanatory power over the baseline ($\Delta R^2 = 0.07$). Notably, the effects of image and textual

contributions remained robust in direction and significance across all models, reinforcing the conclusion that multimodal structure serves as the foundational driver.

However, the hypothesis regarding translation strategy was not supported. Whether measured as a continuous variable(Model 2a: $\mathrm{B} = -0.020, \mathrm{p} = 0.614$)or a dummy variable(Model 2b: $\mathrm{B} = -0.065, \mathrm{p} = 0.455$), the direct effect of translation strategy on popularity was statistically non-significant. This finding indicates that within the context of a high-tech brand like DJI, the choice between domestication and foreignization at the textual level is not a decisive factor for cross-cultural success. Instead, effectiveness is primarily driven by the multimodal architecture and external ecological constraints.

The analysis of control variables revealed key external drivers. Cultural Distance showed a strong significant positive effect ($\mathrm{B} = 0.005, \mathrm{p} < 0.001$), suggesting that DJI's content gains higher popularity in markets with greater cultural disparity from China—potentially due to the appeal of "exoticism" or a standardized premium brand image. Additionally, the GDP of the target market was a significant positive predictor ($\mathrm{B} = 0.018, \mathrm{p} < 0.05$), aligning with market size expectations. Other variables, such as product count, platform, and communicative intent, yielded no significant effects.

4.2.3 Statistical Robustness and Effect Size Diagnostics

o address the non-significant findings regarding translation strategies in the regression models, this study conducted a post-hoc power analysis. First, the effect size (f^2) was evaluated based on the criteria proposed by Cohen (1988). Calculated from the incremental R-squared ($\Delta \mathrm{R}^2 = 0.07$) of Model 2a, the effect size for the full model relative to the baseline (Model 1) was $f^2 = 0.08$. This value falls between a small effect (0.02) and a medium effect (0.15), indicating that the model possesses the requisite sensitivity to capture even subtle variable contributions.

The calculation formula is as follows:

$$f^2 = \frac{\Delta \mathrm{R}^2}{1 - R^2_{full}}$$

Secondly, with a significance level of and a sample size of $N = 202$, the post-hoc statistical power of this study exceeded 0.85. This value is well above the conventional threshold of 0.80 recommended by Cohen (1988), confirming that the model possesses sufficient sensitivity to detect potential effects. This evidence ensures that the non-significance of translation strategies was not a result of insufficient

statistical power; rather, it empirically reveals that within a multimodal context characterized by a high "visual premium," strategic adjustments at the purely textual level yield a minimal marginal contribution to content popularity. Additionally, the Durbin-Watson statistic (≈ 1.88) confirmed that the model is free from significant residual autocorrelation.

5. Conclusion and Discussion

5.1 Summary of Empirical Findings

Through a four-stage empirical design, this study systematically addressed the pre-defined research questions. First, descriptive statistics confirm a pronounced domesticating tendency in DJI's global social media communication, with a mean score of 6.00 and 73% of samples reaching the level of "extreme domestication." This reflects the high degree of "adaptive selection" performed by translators within a multimodal digital ecosystem.

Second, the regression models reveal the underlying logic driving communicative popularity. Image contribution exhibited a significant positive predictive effect($B = 0.669, p < 0.01$), whereas textual contribution showed a significant negative effect($B = -0.823, p < 0.01$). This contrast directly addresses RQ2, confirming that the visual mode bears the core "semiotic weight" in DJI's advertisements.

Finally, the net effect test for the core variable (RQ3) found that after disentangling ecological constraints—such as cultural distance and target market GDP—the translation strategy (domestication vs. foreignization) per se exerted no significant independent influence on communicative outcomes($p > 0.05$). Post-hoc statistical power analysis($\text{Power} > 0.85$)further validates the robustness of this conclusion, revealing that in the presence of a potent "visual premium," the effects of strategic adjustments at the purely textual level are significantly diluted.

5.2 Implications for Theory and Practice

The empirical findings of this study offer a fresh perspective for expanding translation studies and provide practical insights for the global communication of Chinese brands. Theoretically, this research transcends the traditional preoccupation with the binary opposition of intralingual strategies (domestication vs. foreignization). It suggests that a translator's decisions are, to an extent, an adaptive response to a digital ecosystem composed of imagery, technical attributes, market economics, and cultural boundaries. In multimodal contexts, translation should perhaps not be viewed as an isolated textual conversion, but rather as a regulatory variable within an ecosystem. This prompts a shift in research focus from "how to translate" to "the weight allocation of translation within the holistic construction of multimodal meaning." Practically, this study observes an effect dilution of translation

strategies in high-tech communication contexts. When the visual mode generates a strong Visual Premium and anchors the core semantics, the marginal contribution of textual strategy adjustments remains limited. This universality of visual semiotics effectively delineates the boundaries of traditional translation strategies. For global brands like DJI, we suggest that enterprises allocate resources toward visual content design with "cross-cultural intuition." Furthermore, since the independent effect of translation strategy is non-significant in specific models, brand managers should look beyond micro-level linguistic adjustments. Instead, they should ensure that translated content achieves logical self-consistency with macro-ecological factors—such as cultural distance and economic levels—thereby exploring a paradigm shift from "linguistic localization" to "ecological adaptation."

5.3 Limitations

While this study utilizes rigorous statistical diagnostics to ensure the reliability of its conclusions, several avenues for further exploration remain. This inquiry primarily focused on the directionality of translation strategies (domestication/foreignization). Future research should introduce more granular variables, such as linguistic style, rhetorical devices, or emotional granularity, to investigate whether translation exerts non-linear moderating effects on audiences at a micro-level. Regarding the depth of audience perception, the current "popularity" metric reflects communicative outcomes. Subsequent studies could utilize eye-tracking or neuromarketing methods to further deconstruct the psycho-interactive mechanisms between visual modes and textual translation from the perspective of cognitive processing.

5.4 Conclusion

Drawing on the lens of Eco-translatology, this study empirically observes the structural squeezing of translation strategy effects by the multimodal digital ecology. The findings indicate that in DJI's global social media advertisements, the "visual premium" of images and external ecological constraints constitute the foundational logic of communicative effectiveness, while intralingual strategy adjustments exhibit diminishing marginal contributions. This conclusion not only challenges the "text-centric" orthodoxy in traditional advertising translation but also aligns with the burgeoning trend of multimodal and big-data-driven research paradigms within the discipline. It provides an empirical basis for internationalizing Chinese brands to strategically balance modal allocation and translational choices.

References

Adab, B., & Cristina, V. (Eds.). (2004). Key debates in the translation of advertising material. *St. Jerome*.

Baines, R., & Dalmasso, F. (2007). A text on trial: The translation and adaptation of Adel Hakim's Executeur 14. *Social Semiotics*, *17*(2), 229-257.

Baker, M., & Pérez-González, L. (2023). *Translation and interpreting[M]//The Routledge handbook of*

applied linguistics (pp. 230-242). Routledge.

Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis: A multimodal toolkit and coursebook*. University of Toronto Press.

Bateman, J. (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge.

Bell, M., Barnard, J., & McKenzie, D. F. (2003). *Mise-en-page, illustration, expressive form[M]//A History of the Book in Britain*.

Bonsón, E., & Ratkai, M. (2013). A set of metrics to assess stakeholder engagement and social legitimacy on a corporate Facebook page. *Online information review*, *37*(5), 787-803.

Chandler, D. (2002). *The basics*. London, UK: Routledge.

Chen, Z. (2024). *Subtitling as multimodal representation: A corpus-based experimental approach to text-image relations*.

Chen, X., Pan, H., & Li, P. (2020). The multimodal turn in translation studies: Current status and future prospects. *Foreign Language Studies*, (02), 80-87.

Chesterman, A. (2002). Semiotic modalities in translation causality. *Across Languages and Cultures*, *3*(2), 145-158.

Chesterman, A. (2000). Teaching strategies for emancipatory translation. *Benjamins translation library*, *38*, 77-90.

Chuang, Y. T. (2006). Studying subtitle translation from a multi-modal approach. *Babel*, *52*(4), 372-383.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences New York* (Vol 54, pp. 77-155). NY: Academic.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological methods*, *1*(1), 16.

Davitti, E. (2012). *Dialogue interpreting as intercultural mediation integrating talk and gaze in the analysis of mediated parent-teacher meetings*. The University of Manchester (United Kingdom).

De Marco, M. (2012). *Audiovisual translation through a gender lens*. Amsterdam: Rodopi.

Gambier, Y. (2016). Translations| rapid and radical changes in translation and translation studies. International Journal of Communication, *10*, 20-20.

Genette, G. (1997). *Paratexts: Thresholds of interpretation*. Cambridge University Press.

Hessel, J., Holtzman, A., Forbes, M. et al. (2021). C*lipscore: A reference-free evaluation metric for image captioning[C]//Proceedings of the 2021 conference on empirical methods in natural language processing*, 7514-7528.

Hofstede, G. (2001). Culture's consequences: Comparing values, behaviors, institutions and organizations across nations. *International Educational and Professional*.

Hu, G. S. (2011). Research on Ecological Translation Studies Grows Silently. *China Social Sciences News*, 2011-11-17(005).

Kaindl, K. (2022). *The Centrality of the Margins. Transfiction and Bordering Approaches to Theorizing*

Translation: Essays in Dialogue with the Work of Rosemary Arrojo.

Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of international business studies*, *19*(3), 411-432.

Kress, G. R., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*.

Li Y, Xie Y. Is a picture worth a thousand words? An empirical study of image content and social media engagement[J]. Journal of marketing research, 2020, 57(1): 1-19.

Littau, K. (1997). Translation in the age of postmodern production: From text to intertext to hypertext[C]. *Forum for Modern Language Studies*, *33*(1), 81-96.

Liu, Y., & O'Halloran, K. L. (2009). Intersemiotic texture: Analyzing cohesive devices between language and images. *Social semiotics*, *19*(4), 367-388.

Marsh, E. E., & Domas, W. M. (2003). A taxonomy of relationships between images and text. *Journal of documentation*, *59*(6), 647-672.

Mason, I. (2014). Role, positioning and discourse in face-to-face interpreting[M]//Interpreting and translating in public service settings. *Routledge*, 52-73.

Maszerowska, A. (2012). Casting the light on cinema–how luminance and contrast patterns create meaning. *MonTi: Monografías de Traducción e Interpretación*, (4), 65-85.

Mitchell, A. A. (1986). The effect of verbal and visual components of advertisements on brand attitudes and attitude toward the advertisement. *Journal of consumer research*, *13*(1), 12-24.

Nida, E. A. (1964). *Toward a science of translating: With special reference to principles and procedures involved in Bible translating*. Brill Archive.

Nord, C. (1991). Scopos, loyalty, and translational conventions. *Target. International Journal of Translation Studies*, *3*(1), 91-109.

Ortabasi, M. (2007). Indexing the past: Visual language and translatability in Kon Satoshi's Millennium Actress. *Perspectives*, *14*(4), 278-291.

Pan, F., & Wu, X. (2024). Mapping intellectual structures and research trends of translation studies: A bibliometric analysis from 2007 to 2021. *Perspectives*, *32*(4), 736-757.

Pan, L., Wu, X., Luo, T. et al. (2023). Multimodality in translation studies (Vol 10). London & New York: Routledge.

Pérez-González, L. (2007). Appraising dubbed conversation: Systemic functional insights into the construal of naturalness in translated film dialogue. *The Translator*, *13*(1), 1-38.

Pérez-González, L. (2013). Co-creational subtitling in the digital media: Transformative and authorial practices. *International Journal of Cultural Studies*, *16*(1), 3-21.

Pérez-González, L. (2007). Intervention in new amateur subtitling cultures: A multimodal account. *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, 6.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of marketing*, *68*(2), 36-50.

Reiss, K., & Vermeer, H. J. (1984). Groundwork for a general theory of translation. *Tubingen: Niemeyer*, 101.

Stöckl, H. (2004). In between modes. *Perspectives on multimodality*, *6*, 9.

Su, R. (2024). *The impact of foreignising translation strategies on the reception of contemporary Chinese short fiction: Empirical evidence from psychometric instruments and eye tracking*. Macquarie University.

Valdés, C. (2013). *Advertising translation[M]//The Routledge handbook of translation studies* (pp. 303-316).

Vinay, J. P., & Darbelnet, J. (1995). *Comparative stylistics of French and English*.

Voorveld, H A. M., Van Noort, G., Muntinga, D. G. et al. (2018). Engagement with social media and social media advertising: The differentiating role of platform type. *Journal of advertising*, *47*(1), 38-54.