Research Article

Ruihua Zhao and Siyu Zhuang*

From Chinese Classic to South Korean Adaptation: A Multimodal Analysis of Transmedia Cultural Mediation in *Journey to the West*'s Animated and Television Versions

Abstract: This study examines how multimodal ensembles in *Journey to the West* adaptations collectively mediate cultural transmission across different audience demographics. Moving beyond conventional unimodal analyses, this study develops an integrated framework that simultaneously addresses linguistic, paralinguistic, and narrative dimensions in children's animation versus adult-oriented television adaptations. Through systematic comparison of the Little Fox animated series (targeting 6–12 years old) and the 1986 CCTV television series, we reveal how synchronized multimodal strategies, including lexical simplification (TTR = 0.28 vs. 0.35), syntactic adjustment (7.2 vs. 14.7 mean T-unit length), and narrative scaffolding (43.7 % omniscient narration) - create distinct ideological and cognitive mediation pathways. The multimodal analysis demonstrates three key mechanisms of audience adaptation: (1) naturalized cultural schematization through audiovisual anchoring in animation, where simplified vocabulary (82 % BNC top 1,000 words) aligns with visual cues to reinforce comprehension; (2) layered semiotic affordances in television adaptations, where complex syntax (1.12 clauses/sentence) interacts with cinematic techniques to preserve literary nuance; and (3) developmental scaffolding through strategic modality interplay, with animation's hybrid narration-dialogue structure (3.7-second turn-taking) offering cognitive support absent in television's dialogue-driven approach (2.4-second turns). These findings reveal how South Korean animation mediates Chinese cultural capital for transnational audiences, contributing to broader discussions on East Asian cultural power dynamics in globalized media.

*Corresponding author: Siyu Zhuang, Sun Yat-sen University, No.135, Xingang Xi Road, Guangzhou, China, E-mail: zhuangsy9@mail2.sysu.edu.cn. https://orcid.org/0009-0000-7640-5950
Ruihua Zhao, Sun Yat-sen University, No.135, Xingang Xi Road, Guangzhou, China, E-mail: zhaorh5@mail.sysu.edu.cn

1 Introduction

The interlingual adaptation of literary classics across media platforms necessitates a complex negotiation between linguistic transformation, audience considerations, and medium-specific affordances. This phenomenon finds particularly compelling manifestation in the divergent English adaptations of *Journey to the West*, where distinct versions crafted for child versus adult audiences illuminate critical questions about systematic variation in translation methodology relative to viewer demographics and presentational formats. The current investigation employs a purpose-built parallel corpus comprising two representative adaptations, the Little Fox animated series targeting 6–12 years old and the 1986 CCTV television series for general audiences, to address two fundamental research questions: What quantifiable systematic differences exist in lexical sophistication and syntactic structures between these audience-specific adaptations? How do the linguistic characteristics of animation voice-over narration versus television series dialogue reflect intrinsic medium-specific properties?

Contemporary scholarship in audiovisual translation studies (Bogucki and Deckert 2020) and media linguistics has increasingly emphasized the imperative for empirical examination of how audience age parameters and medium conventions collectively shape translation outcomes. While existing research has delineated broad principles of linguistic simplification in children's literature translation (McMartin and Van 2020), the precise linguistic mechanisms through which adaptations modulate complexity across age cohorts remain insufficiently explored. Similarly, although the distinctive narrative architectures of animation versus live-action drama have been thoroughly documented (Pallant 2021), their linguistic manifestations in translated versions have attracted surprisingly limited academic scrutiny.

This study aims to advance the field through three substantive contributions. First, it provides a rigorous, corpus-driven analysis of gradational variation in lexical and syntactic adaptation strategies along the child-to-adult audience spectrum. Additionally, it elucidates how medium-determined features, particularly the animation voice-over/television dialogue dichotomy, fundamentally condition translation decisions. What's more, by integrating quantitative linguistic profiling with qualitative discourse analysis, it establishes a novel analytical framework that bridges theoretical constructs from translation studies and media linguistics. The findings yield actionable insights for translation practitioners, media adaptation specialists, and educational content developers working across diverse platforms and audience segments.

2 Literature Review

The interplay between cross-media adaptation and linguistic adaptation constitutes a critical nexus in contemporary media studies, where textual transformations are increasingly shaped by technological affordances, cultural localization, and audience stratification. While existing scholarship has extensively mapped theoretical frameworks for analyzing cross-media narratives (Jenkins 2006; Kress & van Leeuwen 2001), few studies systematically quantify how linguistic features adapt to divergent audience needs across media formats. *Journey to the West*, as one of the most frequently adapted classical texts globally, offers a unique empirical lens to investigate these dynamics. Its multilingual translations and multimodal reinterpretations – spanning literature, film, animation, and gaming – provide a rich corpus for examining how audience-oriented strategies manifest in lexical, syntactic, and narrative dimensions.

2.1 Cross-Media Adaptation and Linguistic Adaptation

The study of cross-media adaptation has been profoundly shaped by theoretical frameworks from media studies, translation studies, and computational linguistics. Bolter and Grusin's (1999) remediation theory established the foundational premise that new media refashion older forms through dialectical processes of immediacy (transparent representation) and hypermediacy (self-conscious mediation). This concept has been extended by Jenkins (2006) through trans-media storytelling, which examines how narratives expand across platforms like films, games, and novels, each contributing distinct elements to a unified story-world. Ryan's (2015) work on digital interactivity further demonstrates how hypertext fiction disrupts linear narratives, suggesting that media-specific affordances fundamentally alter storytelling logic. In parallel, theory of adaptation emphasizes the creative reinterpretation inherent in cross-media transfers, arguing that adaptations are not derivative but transformative acts of "repetition with variation" (Hutcheon 2013).

Within translation studies, the paradigm-shifting work redefined equivalence models by foregrounding linguistic adaptation strategies – simplification, cultural compensation, and audience localization – as dynamic processes shaped by target contexts (Baker 2018). Her analysis aligns with Díaz-Cintas and Remael's (2020) empirical findings on subtitling, where linguistic choices must reconcile source-text fidelity with screen-time constraints and visual synchronicity. The exploration of film dubbing similarly reveals the tension between cultural

authenticity and technical demands (e.g., lip-synchronization) (O'Sullivan 2011), a challenge magnified in translating culturally dense texts like *Journey to the West*. Yu's (2012) monumental English translation of this classic exemplifies such adaptive labor, employing footnotes and syntactic restructuring to bridge Buddhist allegories and modern Anglophone readerships. The corpus-based analysis of Yu's translation (Duan and Li 2024) further quantifies lexical simplification patterns, revealing systematic reductions in archaic vocabulary and complex clause embeddings.

The rise of corpus linguistics has enabled rigorous quantification of cross-media linguistic shifts. The pioneering corpus-based translation studies introduced metrics like lexical density and sentence-length distributions to compare source and target texts (Laviosa 2002). McEnery and Hardie (2012) expanded these methods to multimodal corpora, advocating for integrated analyses of visual and textual data in film and animation adaptations. However, applications to non-Western contexts remain sparse. Liu's (2023) examination of TV drama dialogues highlights colloquialization trends but confines its scope to a single medium. Meanwhile, Kress and van Leeuwen's (2001) multimodal discourse analysis provides a theoretical lens to dissect media-specific semiotic interactions (e.g., the visual-textual synergy of animation), though its application to East Asian adaptations is limited (Hu 2007; Feng et al. 2014).

Audience stratification has emerged as a critical variable in adaptation studies. Bednarek's (2010) analysis of TV drama dialogues demonstrates how linguistic features (e.g., slang, ellipsis) construct character identities and foster viewer empathy. Ensslin (2022) argues that children's animations employ vocabulary simplification and repetitive syntax to aid cognitive processing, while Pérez-González (2014) explores how fan-subbed media blend informal registers with cultural insider knowledge to cater to niche audiences. South Korean media adaptations occupy a unique interstitial position in East Asian cultural flows, simultaneously localizing content for global audiences while maintaining strategic proximity to source cultures. Studies of Korean animation exports (Jin 2016) identify a "pedagogical mediation paradigm," where linguistic simplification (e.g., reduced clause embedding) is paired with overt cultural glossing (e.g., narratorial explanations of the Buddhist cosmology of *Journey to the West*). This contrasts with Chinese adaptations, which often preserve cultural opacity as a marker of authenticity (Huang et al. 2018), and Japanese adaptations, which prioritize aesthetic fidelity through minimal intervention (O'Hagan 2022). Little Fox's approach exemplifies this South Korean model, leveraging the nation's expertise in cultural bridging, a legacy of its dual role as historical recipient of Chinese cultural influence and contemporary exporter of mediated content.

Despite these advances, few studies systematically compare linguistic strategies across media types (e.g., animation narration vs. TV drama dialogue) or quantify audience adaptation in non-Western contexts. For instance, Tian (2018) conducted an analysis of Chinese animated classics prioritizes visual aesthetics over language, while Zou (2020) proceeded a historical drama study focuses on cultural fidelity at the expense of syntactic complexity analysis.

2.2　Translation and Adaptation Studies of *Journey to the West*

The translation and adaptation of *Journey to the West* (Xiyouji) have sparked multidimensional scholarly debates across translation studies, media theory, and postcolonial criticism, revealing both the cultural richness of the text and the complexities of its cross-cultural dissemination. Anthony C. Yu's four-volume translation (1977–1983) is widely recognized as an academic benchmark, particularly for its meticulous handling of Buddhist and Taoist terminology. Wang and Humblé (2019) emphasize that Yu's translation strikes a balance between philological rigor and literary accessibility, preserving the allegorical depth of the original while catering to modern readership. In contrast, Arthur Waley's abridged Monkey (1942), though popular in the West, has been criticized for oversimplifying religious symbolism. Wang et al. (2023) argue that Waley's selective translation reduces intricate philosophical dialogues to mere adventure narratives, exemplifying what they term "cultural reductionism." The rise of digital humanities has introduced novel methodologies for analyzing translation patterns. Wang and Ping (2023) employ corpus linguistics to trace lexical shifts across six English translations from 1942 to 2018, revealing a transition from exoticizing terms like "demon" to culturally neutral descriptors such as "spiritual being," reflecting evolving Western perceptions of Chinese spirituality.

Cross-media adaptations constitute another critical research focus. Stephen Chow's film *Journey to the West: Conquering the Demons (2013)* exemplifies the postmodern hybridization of traditional narratives. Wei and Huang (2014) analyze how the film's absurdist humor and CGI-driven action sequences subvert the original's religious gravitas while paradoxically reinforcing its moral framework for global audiences. Conversely, Huang et al (2018) examine the international dissemination of the 1986 CCTV television series in Myanmar, proposing strategies to enhance China's cultural soft power through transnational media representation. Digital adaptations have opened new interpretive dimensions. The video game *Black Myth: Wukong (2024)*, praised for its hyper-realistic rendering of mythological landscapes, is critiqued by Xu (2024) for potentially reducing cultural heritage to consumable exoticism through its spectacle-driven aesthetics.

Postcolonial perspectives further complicate discussions of cross-cultural adaptation. Gao et al. (2024) explore ChatGPT's proficiency in translating culturally embedded symbols and implicit semantic elements, suggesting that AI-driven tools may unlock new possibilities for rendering classical literary texts into foreign languages while navigating linguistic and cultural nuances.

2.3 Studies of Audience-Oriented Linguistic Features Across Media

Contemporary research on audience-oriented linguistic features across media reveals the intricate adaptations language undergoes to accommodate diverse sociocultural and technological demands. In audiovisual translation, the seminal work of Díaz-Cintas and Remael (2020) demonstrates how streaming platforms enhance global accessibility through lexical simplification (e.g., replacing culture-specific idioms with paraphrases) and register shifts (e.g., formal to colloquial), as exemplified by Netflix's multilingual subtitling practices. These findings resonate with Hagström and Pedersen's (2022) analysis of dubbing strategies for children's animations, where syntactic compression and lexical repetition align with younger audiences' cognitive processing speeds. Gaming localization studies reveal similar complexities: O'Hagan (2022) proposes "adaptive ludolinguistics" in Final Fantasy series localizations, where dialogue complexity dynamically adjusts based on regional player demographics, while Mangiron (2021) argues that Japanese multimodal RPGs like Persona 5 retain honorific systems in English translations to preserve narrative authenticity for niche audiences.

Social media platforms exhibit unique linguistic adaptation strategies. Androutsopoulos (2015) decodes Twitter/X hashtag campaigns, revealing how code-switching between English and vernacular languages (e.g., #BlackLivesMatter in multilingual contexts) maximizes cross-cultural resonance. Similarly, Virtanen and Lee (2022) document TikTok creators' strategic use of emoji clusters and syntactic truncation to optimize algorithmic visibility and audience engagement. Podcast research further highlights parasocial language techniques: Bozorgian and Shamsi (2022) emphasize the effectiveness of technology-enhanced listening input, such as podcasts, in addressing the scarcity of authentic, comprehensible materials in foreign language learning environments. Romero-Trillo (2019) demonstrates how learners acquire second-language prosodic features and pragmatic implications through blogs, thereby triggering sensory responses.

Multimodal discourse analysis provides critical insights into linguistic-visual synergies. Kress and van Leeuwen's (2001) framework illustrates how infographic designers integrate concise text with color-coded data visualizations to enhance information retention, a phenomenon corroborated by Han (2024) in YouTube

educational videos where speech rhythm coordinates with animated diagrams. Emerging studies on AI-generated content reveal novel adaptation mechanisms: Gao et al. (2024) demonstrate ChatGPT-4's ability to adjust narrative voice (formal vs. conversational) based on user prompts, while Cai (2024) critiques AI tools for excelling at repetitive tasks like text conversion and terminology lookup but failing to ensure logical coherence in translation, thus necessitating human translators to act as discerning "gatekeepers" with critical judgment.

In brief, while existing scholarship illuminates cross-media adaptation's theoretical and practical dimensions, studies mainly focuses on Western media (Jenkins 2006; Ryan 2015), rather than systematic analyses of Chinese literary adaptations like *Journey to the West,* methodological reliance on qualitative case studies (Zhao and Huang 2005; Pei 2024) hampers large-scale linguistic pattern detection across media, and the interplay between media specificity (e.g., animation's descriptive narration) and audience stratification (e.g., child vs. adult viewers) remains underexplored in corpus-driven research. This study may address these gaps through a corpus-based analysis of *Journey to the West* adaptations, interrogating how quantifiable linguistic divergences reflect targeted audience engagement and medium-specific constraints.

3 Methodology

This study employs a mixed-methods corpus-driven approach to investigate audience-oriented linguistic adaptation strategies in cross-media adaptations of *Journey to the West,* with a focus on quantifying lexical and syntactic divergences between animation and television adaptations while identifying medium-specific linguistic signatures. The research design integrates quantitative corpus linguistics, comparative statistical analysis, and qualitative discourse examination to systematically address two research questions: (1) systematic differences in linguistic complexity across audience-stratified adaptations and (2) medium-specific linguistic features shaped by narrative modes (voiceover vs. dialogue). The methodology is structured around four phases – corpus construction, data annotation, quantitative-qualitative analysis, and statistical validation – ensuring methodological transparency and replicability.

The self-built bilingual corpus comprises two primary components: (1) the Little Fox Animation Corpus, containing English subtitles from 108 episodes of *Journey to the West* animations explicitly targeting children aged 6–12, with 63,301 tokens, selected for its role as a culturally proximate non-Chinese mediator with glocalization strategies, measurable child-oriented standards, and standardized cross-cultural comparability for exporting Asian cultural content (Jin and Yoon 2017; Kim et al. 2018). (2) the 1986 CCTV TV Series Corpus, including original Chinese subtitles and official English subtitles from the 25-episode television series aimed at general audiences, with 89,031

tokens. The Little Fox animation's target demographic of 6–12-year-olds is explicitly confirmed by its distribution as EFL pedagogical material through graded reading programs (Little Fox Education 2023), with linguistic features systematically calibrated to this age group's comprehension abilities. While the 1986 CCTV series was broadcast as "general audience" programming, contemporary Chinese media contexts understood this classification as primarily serving adults and older youth, with children expected to access content through family mediation with adult explanation, unlike Little Fox's self-contained pedagogical design for independent child viewing (Wang et al. 2023). Metadata for both corpora document critical variables such as media type (animation vs. TV drama), audience demographics (age group, cultural familiarity), and narrative mode (voiceover narration in animation vs. dialogue-driven scripts in TV dramas). Data extraction was conducted using Subtitle Edit 4.0, followed by manual verification to ensure accuracy and consistency. Non-linguistic elements (e.g., sound effect descriptions, timestamps) were systematically removed to isolate textual data for analysis.

To enable granular linguistic analysis, the corpus underwent multi-layered annotation using computational and manual techniques. Part-of-speech (POS) tagging and syntactic dependency parsing were performed using Stanford CoreNLP (Manning et al. 2014) and spaCy (Honnibal et al. 2020), which provided automated identification of grammatical categories (nouns, verbs, adjectives) and syntactic structures (clause coordination, passive constructions). Manual annotation focused on culture-specific terms (e.g., Buddhist idioms) and medium-specific markers (e.g., voiceover segments in animation). XML markup was applied to encode metadata, including scene context (action, dialogue, narration) and audience targeting parameters (age-appropriateness, cultural localization level). Quantitative analysis prioritized three dimensions: lexical complexity, syntactic complexity, and medium-specific divergence. Lexical complexity was operationalized through type-token ratio (TTR), lexical density (ratio of content words to function words), and frequency of culture-bound terms. Syntactic complexity metrics included mean sentence length (measured in words per sentence), clause coordination index (C/M unit ratio), and passive voice frequency. Statistical comparisons between the animation and TV corpora were conducted using independent *t*-tests and ANOVA in R (R Core Team 2023), with significance thresholds set at *p* < 0.05. To visualize patterns, Tableau generated heatmaps contrasting lexical-syntactic profiles across media types.

Qualitative discourse analysis complemented quantitative findings by examining how narrative modes shape linguistic adaptation. Voiceover narration in the Little Fox animations was analyzed using Biber's (1988) multidimensional framework to identify register features (e.g., informational vs. involved discourse), while TV drama dialogues were assessed for pragmatic strategies (e.g., turn-taking patterns, speech act distribution). Cultural localization strategies were scrutinized

through close reading of key terms, such as the translation of culturally specific concepts (e.g., retention of "Guanyin" vs. domestication as "Goddess of Mercy"). The integration of quantitative and qualitative methods ensures a holistic understanding of how audience stratification and media affordances jointly shape linguistic adaptation.

4 Findings

Through a rigorous comparative analysis of the Little Fox animated series corpus and the 1986 CCTV television series corpus, this study systematically examines the quantitative impact of medium-specific constraints and target audience considerations on the English translation strategies employed in *Journey to the West* adaptations across four key dimensions: lexical, syntactic, narrative, and pragmatic features. As demonstrated in Table 1, the analysis reveals statistically significant gradational differences in lexical complexity (all p < 0.001) that extend beyond surface-level metrics to reflect deliberate linguistic adaptation strategies tailored to distinct audience demographics.

The lexical complexity metrics exhibit clear bifurcation patterns. The animated corpus displays significantly lower Type-Token Ratio (TTR = 0.28) compared to the television version (TTR = 0.35, p < 0.001), indicating strategic lexical simplification for

Table 1: Lexical complexity of the corpora.

Measurement indicator	Animation	TV series	p–value
Type–token ratio	0.28	0.35	<0.001
Lexical density (%)	54.1	61.3	<0.001
Proportion of content words (%)	58.7 ± 2.1	67.2 ± 1.8	<0.001
Coverage of the 1,000 most frequent words (%)	82	68	<0.001
Average word length	4.1	5.3	<0.001
Flesch–kincaid index	85.2 (very easy)	72.6 (average)	<0.001
Lexical difficulty index (LFP)	62.4	78.9	0.001

LFP (Lexical Frequency Profile) is calculated via VocabProfile. This table presents multiple measurement indicators to compare the lexical complexity between animation and TV series corpora. The Type–Token Ratio, which measures the diversity of vocabulary, shows a value of 0.28 for animation and 0.35 for TV series, with a p–value < 0.001, indicating a significant difference. Lexical Density, representing the proportion of content words in the text, is 54.1 % for animation and 61.3 % for TV series, also with a p–value < 0.001. The Proportion of Content Words (%) is 58.7 ± 2.1 for animation and 67.2 ± 1.8 for TV series, again showing a highly significant difference (p–value < 0.001). The Coverage of the 1,000 Most Frequent Words (%) is 82 for animation and 68 for TV series, with a significant p–value < 0.001. The Average Word Length is 4.1 for animation and 5.3 for TV series, and the Flesch–Kincaid Index, which assesses reading ease, is 85.2 (Very Easy) for animation and 72.6 (Average) for TV series, both with p–values < 0.001. The Lexical Difficulty Index (LFP) is 62.4 for animation and 78.9 for TV series, with a p–value of 0.001. These p–values collectively demonstrate statistically significant disparities in lexical complexity between the two corpora.

younger audiences. Frequency analysis conducted via AntConc 3.5.8 shows that content words (nouns, verbs, adjectives, adverbs) constitute only 58.7 % (± 2.1 %) of the 63,301-token animated corpus, while function words (prepositions, conjunctions, articles) account for 41.3 %, a distribution that aligns with established principles of child language acquisition by enhancing syntactic predictability and textual cohesion. In contrast, the 89,031-token television corpus manifests an inverse distribution with 67.2 % (± 1.8 %) content words and 32.8 % function words, reflecting its prioritization of semantic density (animation: 54.1 % vs television: 61.3 %, $p < 0.001$) to accommodate richer cultural information and more complex conceptual content.

Advanced lexical profiling using Wordsmith 7.0 further elucidates these adaptations. The animated corpus demonstrates 82 % coverage by the BNC top 1,000 word list, with a mean frequency of 148.7 occurrences per 10,000 words and minimal Academic Word List (AWL) representation (1.2 %) – characteristics indicative of a tightly controlled lexical system designed for L2 learners. Conversely, the television corpus shows greater lexical sophistication with 68 % BNC top 1,000 coverage, lower mean frequency (89.4/10,000 words), and substantially higher AWL presence (4.7 %), preserving the source text's literary nuances. The Lexical Frequency Profile (LFP) analysis confirms this pattern, with the television corpus scoring 26.5 % higher (78.9 vs 62.4, $p = 0.001$) on lexical difficulty metrics. Morphological analysis reveals the animated corpus' words average 4.1 letters compared to the television version's 5.3 letters ($p < 0.001$), while Flesch-Kincaid readability scores (animation: 85.2 ["very easy"] vs television: 72.6 ["standard"]) correspond to approximately two grade levels' difference in reading difficulty (Kincaid et al. 1975), providing robust empirical evidence of audience-driven adaptation strategies.

As quantitatively demonstrated in Table 2, the L2 Syntactic Complexity Analyzer (Lu 2010) reveals fundamentally divergent syntactic architectures between the corpora.

The Little Fox animation manifests prototypical "oral simplification" (Tavakoli and Foster 2008), with markedly shorter T-units (7.2 words), lower clause density (0.38/clause), and reduced subordinating conjunction frequency (12.3/1 k words). Its 4.1 % passive construction rate and inflated coordinate conjunction usage (23.5/1 k words, 1.8 × the TV series) collectively establish a syntactically transparent system favoring paratactic progression – a design strategically reducing cognitive load for child audiences. Conversely, the CCTV television corpus exhibits "literate complexity" (Biber and Gray 2010; Qin and Kong 2018) through extended T-units (14.7 words), dense subordination (1.12 clauses/sentence), and frequent passive voice (11.3 %). The 28.6/1 k words subordinator frequency and 2.4–3.1 × greater adverbial/relative clause usage construct an embedded syntactic network that meticulously preserves the source text's classical cadence while conveying intricate spatiotemporal relationships and psychological depth.

Table 2: Cross-modal linguistic feature.

Feature	Animation	TV series	Ratio (TV/animation)
Syntactic			
Mean T-unit length	7.2	14.7	2.0:1
Subordinate clauses	0.38	1.12	2.9:1
Passive constructions (%)	4.1	11.3	2.8:1
Narrative			
Narration proportion (%)	43.7	0	–
Dialogue tempo (second)	3.7	2.4	1:1.5
Pragmatic			
Directive speech (%)	19.3	3.5	5.6:1
Cultural adaptation (%)	93.2	35.7	1:2.6

All comparisons significant at p < 0.001 (two-tailed). This table presents a detailed analysis of various cross–modal linguistic features between animation and TV series. Under the syntactic category, the Mean T–unit length is 7.2 for animation and 14.7 for TV series, with a ratio of 2.0:1 for TV series to animation, indicating that the average length of T–units in TV series is twice that in animation. The number of Subordinate clauses is 0.38 for animation and 1.12 for TV series, with a ratio of 2.9:1, showing a nearly three–fold increase in TV series. The percentage of Passive constructions is 4.1 % for animation and 11.3 % for TV series, with a ratio of 2.8:1. In the narrative category, the Narration proportion (%) is 43.7 for animation and 0 for TV series, with no applicable ratio as the value for TV series is zero. The Dialogue tempo (in seconds) is 3.7 for animation and 2.4 for TV series, with a ratio of 1:1.5, suggesting a relatively faster dialogue tempo in TV series. Under the pragmatic category, the percentage of Directive speech is 19.3 % for animation and 3.5 % for TV series, with a ratio of 5.6:1, meaning directive speech is much more prevalent in animation. The percentage of Cultural adaptation is 93.2 % for animation and 35.7 % for TV series, with a ratio of 1:2.6. All comparisons are significant at p < 0.001 (two–tailed), indicating strong statistical differences between the linguistic features of animation and TV series across these categories.

4.1 Narrative Modality and Pragmatic Orientation

The animation's hybrid narration-dialogue structure (Toolan 2001), with 43.7 % narratorial exposition, employs syntactically simplified utterances (22.6 % shorter than dialogue portions) with concentrated psych-verb placement (78.5 % in narration) – a dual mechanism externalizing mental states for developing viewers. Its 3.7-second turn-taking intervals and didactic framing contrast sharply with the television version's pure dialogic mode featuring rapid 2.4-second exchanges and 4.2 × more hedging devices, demanding inferential engagement while enhancing dramatic authenticity.

Pragmatic analyses uncover diametric adaptation priorities: The animation's pedagogical orientation surfaces through frequent directives (19.3 % vs 3.5 %) and comprehension checks (12.7/1 k words), alongside extensive cultural domestication (93.2 %). The television version prioritizes sociolinguistic fidelity through honorifics (8.7 %, 6.4 × animation) and foreignization (64.3 % culture items), with 3.8 × more modal verbs reconstructing hierarchical power dynamics.

4.2 Medium-Specific Syntactic Profiling

Detailed profiling reveals the animation's narrator favors truncated utterances (mean 8.2 words) with elementary clause coordination (C/M ratio 1.3), while the television dialogue employs extended turns (11.6 words) and complex subordination (C/M 2.1). As Table 2 delineates, passive voice – a key formality marker – appears minimally in animation (0.7 % verbs) but proliferates in ritual television discourse (4.2 %, e.g., "By Jade Emperor's decree, you shall be punished"). These patterns manifest not as isolated adaptations to audience or medium, but as emergent properties of their dynamic interaction. Animation's syntactic reduction (8.2 words, C/M 1.3) operates within a self-reinforcing system where: (1) the medium's iconic representation (exaggerated gestures, hyper-articulated mouth movements) compensates for linguistic brevity through visual amplification; (2) child viewers' developing working memory aligns with animation's shot-length constraints (3.7s average); and (3) simplified syntax enables faster visual sequencing, creating a feedback loop that sustains comprehension. By comparison, television's syntactic complexity (11.6 words, C/M 2.1) similarly emerges from interdependent factors: (1) live-action's continuous takes (8.5s average) permit clause embedding without visual discontinuity; (2) adult viewers' developed theory of mind leverages actors' micro-expressions to parse nested structures; and (3) the camera's physical constraints necessitate formal devices (e.g., 4.2 % passive voice) to convey off-screen power relations. This tripartite model, where medium affordances, audience capacities, and narrative requirements co-adapt, reveals syntactic choices as ecological adaptations to each format's unique communicative ecosystem. These systematic variations empirically demonstrate how audience design and medium affordances jointly govern linguistic adaptation strategies.

5 Discussion

The empirical findings emerging from this large-scale comparative analysis of the Little Fox animated series and CCTV television series adaptations of *Journey to the West* present a complex picture of how linguistic adaptation strategies are systematically modulated by the interplay between target audience considerations and medium-specific constraints. At the lexical level, the dramatic 26.5 % disparity in Lexical Frequency Profile scores between the animated corpus (62.4) and television corpus (78.9) cannot be adequately explained through simple reference to word frequency differences alone, but rather reflects fundamentally divergent approaches to vocabulary selection rooted in distinct communicative purposes. The animated

version's rigorous limitation to high-frequency vocabulary (82 % coverage by BNC top 1,000 words) creates what might be termed a "pedagogically optimized" lexical environment where the 98.14 % text coverage threshold identified by Laufer and Nation (1995) as necessary for unassisted comprehension is carefully maintained through strategic repetition of core items (mean frequency = 148.7 occurrences per 10,000 words), while simultaneously excluding potentially problematic low-frequency lexemes that might disrupt the young L2 learner's reading fluency (Schmitt et al. 2011). This lexical simplification strategy extends beyond mere word selection to encompass morphological dimensions as well, with the animated corpus' words averaging just 4.1 letters in length compared to the television version's 5.3 letters, a statistically significant difference (p < 0.001) that reflects conscious avoidance of polysyllabic Latinate derivatives in favor of shorter Germanic-based words that are more easily decoded by developing readers. The television adaptation, by contrast, demonstrates a markedly different lexical philosophy through its intentional incorporation of lower-frequency vocabulary (32 % beyond the 1,000 most common words), including numerous culture-specific terms like "arhat" (罗汉) and "asura" (阿修罗) that preserve the source text's religious and philosophical nuances, albeit at the cost of increased lexical difficulty that would likely render the text inaccessible to the animation's target demographic. What proves particularly revealing is not merely the existence of these lexical differences, but their precise quantification and systematic patterning, which collectively suggest an almost mathematical precision in how audience design parameters translate into concrete linguistic choices, a finding that challenges more intuitive, less empirically grounded models of adaptation strategy selection that have dominated much of the audiovisual translation literature to date.

Moving beyond lexical considerations to syntactic architecture, the data reveal equally profound and carefully calibrated differences in how information is packaged and presented across the two media forms. The animated corpus' preference for shorter T-units (averaging just 7.2 words compared to the television version's 14.7 words) reflects what psycholinguistic research has identified as the cognitive constraints of working memory capacity in child learners, specifically the "7 ± 2" information chunking threshold proposed by Miller (1956), suggesting that the adaptors have intuitively or deliberately structured their syntax to align with developmental cognitive limitations (In'nami et al. 2022). This syntactic simplification is further evidenced by the animation's heavy reliance on coordinate conjunction (23.5 instances per 1,000 words, 1.8 times more frequent than in the television version), which creates what Halliday (1989) describes as "paratactic transparency" - a linear, additive sentence structure that maximizes processing ease while sacrificing the informational density achievable through more complex embedding strategies. The television adaptation, unconstrained by such developmental considerations, freely employs sophisticated hypotactic structures

including center-embedded relative clauses (constituting 17.4 % of all subordinate clauses) and left-dislocated adverbial phrases (23.1 % of subordination), which while posing greater processing challenges, enable the preservation of classical Chinese narrative rhythms and the intricate interweaving of multiple narrative threads that characterize the original literary work. Particularly noteworthy is the television version's significantly higher frequency of passive constructions (11.3 % versus the animation's 4.1 %), which serves not only as a marker of formal register but also as a crucial mechanism for maintaining the original text's thematic emphasis on fate and predestination through constructions like "the scriptures are guarded by divine law" that position characters as recipients rather than initiators of action. The systematic nature of these syntactic differences strongly suggests that media adaptors operate with an implicit (and sometimes explicit) understanding of how grammatical complexity interacts with audience comprehension, and that their choices reflect careful calibration rather than arbitrary simplification or complexification. The interdependent nature of audience and medium factors becomes particularly evident when examining how each adaptation leverages its format's unique affordances to address viewer capabilities. The animation's hybrid narration-dialogue structure (43.7 % narration) exemplifies this synergy: the medium's conventional use of voiceover (Pallant 2021) is strategically deployed to scaffold child viewers' comprehension through explicit mental state explanations (e.g., Sun Wukong felt proud of his trick), while simultaneously compensating for animation's limited ability to convey subtle psychological cues through facial expressions alone. Conversely, the television series' exclusive reliance on dialogue capitalizes on live-action's strength in conveying subtext through actors' performances, a approach viable for its adult audience's stronger inferential abilities but potentially challenging for children. This demonstrates how what might appear as simple audience adaptations are in fact carefully mediated through the distinctive storytelling toolkit of each medium.

The narrative structure differences between the two adaptations prove equally revelatory when examined through the lens of contemporary narrative theory. The animated version's hybrid format combining third-person omniscient narration (constituting 43.7 % of the total text) with character dialogue creates what might be termed a "scaffolded narrative" environment where the narrator assumes responsibility for not only advancing the plot but also explicitly interpreting character psychology (78.5 % of mental state verbs like "think" and "know" occur in narration rather than dialogue), spatial relationships ("the monkey king flew over the mountains"), and temporal connections ("the next day, they continued their journey") that in the television version must be inferred by viewers from dialogue and contextual clues alone. This narrative explicitness, while potentially reducing interpretive richness, serves the crucial pedagogical function of modeling comprehension strategies for young viewers who are still developing the cognitive apparatus necessary for making such inferences

independently. The television adaptation's exclusive reliance on dialogue (100 % of text, with zero narration) represents not merely a different storytelling technique but what might more properly be considered a different epistemology of narrative understanding, one that requires viewers to construct mental models of character motivation, temporal sequence, and causal relationship through careful attention to conversational implicature, turn-taking patterns (averaging just 2.4 s between speaker changes compared to the animation's 3.7 s), and what Sperber and Wilson (1986) term "ostensive-inferential communication" where meaning must be actively reconstructed rather than passively received (Carston 2021). The implications of this narrative divide extend beyond mere presentational differences to touch upon fundamental questions about how different media forms construct knowledge and shape interpretive habits in their audiences, with the animated version fostering what might be called "guided literacy" and the television version cultivating "critical viewing competencies" that may have lasting consequences for how audiences engage not just with these particular texts but with narrative media more broadly. The interplay of linguistic simplification and cultural annotation in Little Fox's adaptation also reflects South Korea's industrial pragmatism in cultural exports. The animation's frequent use of metanarrative commentary (e.g., 'In Chinese culture, the dragon symbolizes good luck') mirrors strategies seen in Korean-translated manhwa (Jin and Yoon 2017), where culturally specific concepts are explicitly framed for non-native audiences. This differs from both the CCTV series' implicit cultural encoding (e.g., untranslated honorifics like Dashi 大师) and Western adaptations' tendency toward cultural substitution (e.g., Waley's replacement of Guanyin with 'the Goddess of Mercy'). Such choices underscore how South Korean adaptations recontextualize Chinese cultural capital, not by diluting its distinctiveness but by rendering it legible within global children's media conventions.

6 Conclusions

This study has systematically investigated the linguistic adaptation strategies in the English translations of *Journey to the West* across animated and television adaptations, employing a multidimensional analytical framework encompassing lexical, syntactic, narrative, and pragmatic dimensions. The findings reveal that audience-specific design and medium constraints lead to quantifiable and systematic linguistic variations, offering new insights into cross-media translation theory. These findings also underscore that audience considerations and medium constraints operate not as independent variables but as mutually shaping forces in cultural adaptation. The animation's lexical simplification (82 % BNC top 1,000 words) works precisely because the medium provides visual reinforcement, just as the television

series' syntactic complexity (14.7-word T-units) is sustainable due to its audience's greater tolerance for delayed payoff in live-action formats.

Methodologically, the integration of corpus-based quantitative analysis (e.g., lexical frequency profiling, syntactic complexity metrics) with qualitative discourse analysis has provided a robust approach to identifying adaptation patterns. This dual-method framework not only enhances the precision of linguistic feature extraction but also captures the dynamic interplay between language simplification and cultural fidelity. Theoretically, this study contributes to audiovisual translation studies by establishing measurable thresholds for audience-oriented adaptations. For instance, the identification of 98 % lexical coverage as critical for child audiences (Laufer and Nation 1995) and the syntactic "7 ± 2" cognitive load limit (Miller 1956) offer actionable guidelines for future adaptations. These findings deepen our understanding of how linguistic choices mediate between source-text authenticity and target-audience accessibility. Practically, the study highlights the importance of tailored strategies: simplified lexico-syntactic structures for pedagogical contexts (e.g., children's animations) and complex, culturally resonant language for literary preservation (e.g., classic literature adaptations). Such distinctions may be vital for translators, educators, and media producers aiming to balance educational objectives with cultural integrity.

References

Androutsopoulos, J. 2015. "Networked Multilingualism: Some Language Practices on Facebook and their Implications." *International Journal of Bilingualism* 19 (2): 185–205.

Baker, M. 2018. *In Other Words: A Coursebook on Translation*, 3rd ed. London: Routledge.

Bednarek, M. 2010. *The Language of Fictional Television: Drama and Identity*. London: Bloomsbury.

Biber, D. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D., and B. Gray. 2010. "Challenging Stereotypes About Academic Writing: Complexity, Elaboration, and Explicitness." *Journal of English for Academic Purposes* 9 (1): 2–20.

Bogucki, Ł., and M. Deckert, eds. 2020. *The Palgrave Handbook of Audiovisual Translation and Media Accessibility (Palgrave Studies in Translating and Interpreting)*. Basingstoke: Palgrave Macmillan.

Bolter, J. D., and R. Grusin. 1999. *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.

Bozorgian, H., and E. Shamsi. 2022. "Autonomous use of Podcasts with Metacognitive Intervention: Foreign Language Listening Development." *Journal of Applied Linguistics* 32 (3): 459–83.

Cai, L.. 2024. "Logical Coherence in Translation: A "Blind Spot" for AI Tools." *Chinese Translators Journal* 45 (04): 182–7.

Carston, R. 2021. "Relevance Theory: Recent Developments, Current Challenges and Future Directions." *Journal of Pragmatics* 182 (4): 33–47.

Díaz Cintas, J., and A. Remael. 2020. *Subtitling: Concepts and Practices*, 1st ed. London: Routledge.

Duan, X., and F. Li. 2024. "A Study on the English Translation and Communicative Value of *Journey to the West*." *Modern Linguistics* 12 (7): 640–6.

Ensslin, A. 2022. *Language in the Digital Era: Challenges and Perspectives*. Cambridge: Cambridge University Press.

Feng, D., D. Zhang, and K. O'Halloran. 2014. "Advances and Frontiers in Multimodal Discourse Analysis." *Contemporary Linguistics* 16 (1): 88–99.

Gao, R., Y. Lin, N. Zhao, and Z. G. Cai. 2024. "Machine Translation of Chinese Classical Poetry: A Comparison Among ChatGPT, Google Translate, and DeepL Translator." *Humanities and Social Sciences Communications* 11 (1): 1–10.

Hagström, H., and J. Pedersen. 2022. "Subtitles in the 2020s: The Influence of Machine Translation." *Journal of Audiovisual Translation* 5 (1). https://doi.org/10.47476/jat.v5i1.2022.195.

Han, X. 2024. "Cultural Heritage in Digital Entertainment: A Case Study of Black myth: Wukong and its Implications for the Chinese Gaming Industry." *Research and Commentary on Humanities and Arts* 2 (6). https://doi.org/10.18686/rcha.v2i6.4727.

Honnibal, M., I. Montani, S. Van Landeghem, and A. Boyd. 2020. "spaCy: Industrial-Strength Natural Language Processing in Python." *Journal of Open Source Software* 5 (52): 2606.

Hu, Z. 2007. "Theoretical Foundations and Research Methods of Multimodal Discourse Analysis." *Foreign Language Research* (5): 1–10.

Huang, Q., Y. Cui, and C. Zhang. 2018. "Enhancing National Image Through Transnational Television: A Case Study of the 1986 CCTV *Journey to the West* in Myanmar." *International Communication Gazette* 80 (5): 456–73.

Hutcheon, L. 2013. *A Theory of Adaptation*, 2nd ed. London: Routledge.

In'nami, Y., Y. Hijikata, and R. Koizumi. 2022. "Working Memory Capacity and l2 Reading: A Meta-Analysis." *Studies in Second Language Acquisition* 44 (2): 381–406.

Jenkins, H. 2006. *Convergence Culture: Where Old and New Media Collide*. New York: NYU Press.

Jin, D. Y. 2016. *New Korean Wave: Transnational Cultural Power in the Age of Social Media*. Urbana: University of Illinois Press. http://www.jstor.org/stable/10.5406/j.ctt18j8wkv.

Jin, D. Y., and T. J. Yoon. 2017. "The Korean Wave: Retrospect and Prospect: Introduction." *International Journal of Communication* 11: 2241–9.

Kim, S. J., A. Song, G.-L. Lee, and A. Bach. 2018. "Using Animated Folktales to Teach Cultural Values: A Case Study with Korean-American Bilingual Kindergartners." *Journal of Research in Childhood Education* 32 (3): 295–309.

Kincaid, J. P., R. P. Fishburne Jr., R. L. Rogers, and B. S. Chissom. 1975. *Derivation of New Readability Formulas for Navy Enlisted Personnel. Research Branch Report 8 – 75*. Millington, TN: Naval Technical Training Command.

Kress, G., and T. van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.

Laufer, B., and P. Nation. 1995. "Vocabulary Size and use: Lexical Richness in L2 Written Production." *Applied Linguistics* 16 (3): 307–22.

Laviosa, S. 2002. *Corpus-Based Translation Studies: Theory, Findings, Applications*. Amsterdam: Rodopi.

Little Fox Education. 2023. *Animated Stories for English Learners*. https://www.littlefox.com.

Liu, W. 2023. "A Multimodal Discourse Analysis on Subtitle Translation of the Knockout: A Chinese Political-Themed TV Drama." *International Journal of Languages, Literature and Linguistics* 9 (6): 437–40.

Lu, X. 2010. "Automatic Analysis of Syntactic Complexity in Second Language Writing." *International Journal of Corpus Linguistics* 15 (4): 474–96.

Mangiron, C. 2021. "Found in Translation: Evolving Approaches for the Localization of Japanese Video Games." *Arts* 10 (1): 9.

Manning, C., M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky. 2014. "The Stanford CoreNLP natural Language processing toolkit." In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 55–60. Association for Computational Linguistics.

McEnery, T., and A. Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

McMartin, J., and J. Van Coillie. 2020. *Children's Literature in Translation: Texts and Contexts (Translation, Interpreting and Transfer)*. Leuven: Leuven University Press.

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63 (2): 81–97.

O'Hagan, M. 2022. "Indirect Translation in Game Localization As a Method of Global Circulation of Digital Artefacts: A Socio-Economic Perspective." *Target: International Journal of Translation Studies* 34 (2). https://doi.org/10.1075/target.00007.oha.

O'Sullivan, C. 2011. *Translating Popular Film*. Basingstoke: Palgrave Macmillan.

Pallant, C. 2021. *Animation: Critical and Primary Sources*. London: Bloomsbury Academic.

Pei, L. 2024. "Transmedia Narrative Research on Literary-Adapted Films from the Perspective of Media Convergence." *Film Literature* (4): 72–5. https://doi.org/10.3969/j.issn.0495-5692.2024.04.014.

Pérez-González, L. 2014. *Audiovisual Translation: Theories, Methods and Issues*. London: Routledge.

Qin, H. W., and L. Kong. 2018. "The Ways in Which Translated Language Influences Original Language: An Analysis Based on the Structural Complexity of Chinese." *Journal of Foreign Languages* 41 (5): 15–26.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Romero-Trillo, J. 2019. "Prosodic Pragmatics and Feedback in Intercultural Communication." *Journal of Pragmatics* 151 (3): 91–102.

Ryan, M.-L. 2015. *Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity in Literature and Electronic Media*. Baltimore: Johns Hopkins University Press.

Schmitt, N., X. Jiang, and W. Grabe. 2011. "The Percentage of Words Known in a Text and Reading Comprehension." *Modern Language Journal* 95 (1): 26–43.

Sperber, D., and D. Wilson. 1986. *Relevance: Communication and Cognition*. Oxford, UK: Blackwell Publishing.

Tavakoli, P., and P. Foster. 2008. "Task Design and Second Language Performance: The Effect of Narrative Type on Learner Output." *Language Learning* 58 (2): 439–73.

Tian, X. 2018. "The Cross-Cultural Journey of "Chinese Stories": Animated Films, "National Identity," and Classical Literature Adaptations." *Hundred Schools in Arts* 34 (2): 157–62. 239.

Toolan, M. 2001. "Narrative: A Critical Linguistic Introduction (2nd ed.)." *Journal of Literary Semantics* 30 (2): 158–160.

Virtanen, T., and C. Lee. 2022. "Face-Work in Online Discourse: Practices And Multiple Conceptualizations." *Journal of Pragmatics* 195 (3): 1–6.

Wang, F. (Robin), and P. Humblé. 2019. "Readers' Perceptions of Anthony Yu's Self-Retranslation of *the Journey to the West*." *Perspectives* 28 (5): 756–76.

Wang, B. H., and Y. Ping. 2023. "What's New in the New English Translation of *Journey to the West*? – Preliminary Findings Based on a Corpus-Metric Stylistic Study." *Journal of Yanshan University (Philosophy and Social Science Edition)* 24 (3): 9–17+F0002.

Wang, F., K. Liu, and P. Humblé. 2023. "Reframing the Narrative of Magic Wind in Arthur Waley's Translation of *Journey to the West*: Another Look at the Abridged Translation." *Humanities and Social Sciences Communications* 10 (876). https://doi.org/10.1057/s41599-023-02397-0.

Wei, T. J., and J. L. Huang. 2014. "The Postmodernity of Stephen Chow's Films in *Journey to the West*: Conquering the Demons." *Movie Literature* (24): 96–7.

Yu, Anthony C. 2012. *The Journey to the West*, Vol. 1–4. Chicago: University of Chicago Press.

Zhao, S., and J. Huang. 2005. "Linguistic Features and Translation Techniques In Audiovisual Translation." *Journal of Hefei University of Technology (Social Sciences)* 19 (3): 133–7.

Zou, Y. 2020. "Transmedia Narrative Research on Chinese Traditional Literary Works in the Context of Media Convergence." *Journal of News Research* 11 (14): 220–1.

Bionotes

Ruihua Zhao
Sun Yat-sen University, No.135, Xingang Xi Road, Guangzhou, China
zhaorh5@mail.sysu.edu.cn

Ruihua Zhao is an associate professor of English at Sun Yat-sen University, China, and Ph.D. from Hong Kong Polytechnic University. Her interest includes systemic functional linguistics, discourse analysis and ecolinguistics. She is the author of *Expansion of Eco-discourse Analysis Towards the International Relationships: A case study on China's Arctic position* (2023) and *Review of Current Studies in Chinese Language and Discourse* (2020); coeditor (with Huang, G. W.) of *Harmonious Discourse Analysis: Ecolinguistics through Chinese Culture and Philosophy* (2021) and *Harmonious discourse analysis: Approaching peoples' problems in a Chinese context* (2021).

Siyu Zhuang
Sun Yat-sen University, No.135, Xingang Xi Road, Guangzhou, China
zhuangsy9@mail2.sysu.edu.cn
https://orcid.org/0009-0000-7640-5950

Siyu Zhuang is a master candidate in the School of Foreign Languages at Sun Yat-sen University, China. Her interest is both in systemic functional linguistics, multiomodal discourse analysis and cross-cultural translation studies.