

Original Research Article

A corpus-driven comparison of Chinese and American media representations of the Chinese animated film *Ne Zha 2*

Wei He

Shenzhen Campus of Jinan University, Shenzhen, Guangdong, 518053, China

Abstract: This article adopts a corpus-assisted discourse studies approach to analyze the representations of the Chinese animated film *Ne Zha 2* in one Chinese English-language media outlet (China Daily) and two U.S. media outlets (Deadline and Variety) between January and July 2025. The results show that Chinese and U.S. media share a consensus at the cognitive level, as both construct the film's meaning through dimensions such as box office performance, attributes of the animation industry, Chinese cultural resources, and markers of cultural identity. However, Chinese media are more inclined to present the film's commercial success through continuous quantitative indicators and industrial narratives, constructing *Ne Zha 2* as a symbol of the development of China's animation industry and cultural confidence. By contrast, U.S. media place greater emphasis on evaluative and comparative language, positioning the film within the frameworks of the global animation market and transnational circulation, and highlighting its market performance and industrial significance as a non-Hollywood production, thereby displaying different discursive construction orientations.

Keywords: corpus-driven discourse analysis; media discourse; film news; Chinese animation; cross-cultural representation; *Ne Zha 2*

1. Introduction

In recent years, China's film industry has grown rapidly in the domestic market and attracted increasing global attention. Research shows that, alongside expanding production scale, industrialization, and changes in policy and market environments, Chinese cinema has achieved growth in box office performance and genre diversity (Su, 2014; Wang et al., 2025; De Masi et al., 2025). In particular, Chinese animated films have advanced in technological sophistication, narrative capability, and commercial operation, becoming an important vehicle for promoting Chinese culture internationally (Curtin & Shah, 2010; Keane, 2013). Against this background, the Chinese animated film *Ne Zha 2* achieved major box office success and attracted wide attention in both Chinese and international media.

Media reporting is not simply an objective representation of events or cultural products, but a process through which reality is constructed through lexical choices, evaluative strategies, and discursive frameworks (Fairclough, 1995). Previous studies show that when reporting the same issue, media outlets often differ in values, ideological positions, and cultural cognition through variations in keyword selection, collocational patterns, and semantic prosody (Baker et al., 2008; Bednarek & Caple, 2012). Analyzing media language therefore helps uncover meaning-making mechanisms beneath the surface of news texts.

In cross-cultural contexts, representations of the same cultural product are especially significant. Previous research has shown that Chinese and Western media often differ in focal points, evaluative orientations, and discursive strategies when reporting China-related topics (Yang & Wang, 2023; Liu & Li,

2017; Huan, 2024). However, most studies focus on politics, environmental issues, or international conflict, while relatively little attention has been given to cultural products, especially Chinese animated films, in Chinese and U.S. media. Compared with hard news, reports on cultural products rely less on a clear problem–cause–solution structure and more on repeated emphasis of cultural meaning, market performance, and symbolic value.

From a methodological perspective, corpus-assisted discourse analysis offers an effective approach for examining media language. By analyzing high-frequency words, keywords, and co-occurrence patterns in large-scale textual data, researchers can reveal how meanings are naturalized through repeated use (Sinclair, 2004). Compared with purely qualitative discourse analysis, corpus methods reduce reliance on subjective judgment while still allowing contextual interpretation (Baker, 2006).

Against this background, this study examines news reports on *Ne Zha 2* from major Chinese English-language media and U.S. media by constructing comparable corpora. It focuses on how keywords, collocational patterns, and semantic prosody contribute to meaning construction in different media contexts. The study aims to demonstrate the usefulness of corpus-assisted discourse analysis for researching cultural products and to deepen understanding of discursive differences between Chinese and U.S. media.

To achieve the overall research objectives, this study will address the following two interrelated research questions.

(1) What similarities and differences can be identified in the discursive construction of *Ne Zha 2* in reports by Chinese and American media?

(2) How can these linguistic differences be explained in relation to the cultural positions and ideological backgrounds of different media contexts?

2. Literature review

2.1. Media discourse and news representation

Within discourse studies, the media are widely seen as an important site for constructing social meaning. Critical discourse analysis (CDA) argues that media discourse does not neutrally reflect reality. Through lexical choices, narrative structures, and evaluative practices, it represents events and social actors in particular ways, thereby implicitly constructing values and ideological positions (Fairclough, 1995). van Dijk (1998) further notes that ideology is often embedded in routine and taken-for-granted language practices, with news reporting serving as a key arena for ideological reproduction. Therefore, media language analysis concerns not only information, but also how linguistic forms shape social meaning.

This process is especially evident in news discourse. Although objectivity is regarded as a core journalistic value, studies show that news texts often contain evaluative orientations through vocabulary, syntax, and narrative focus (Fowler, 2013; Rubing & Sandaran, 2023). Richardson (2017) argues that even without explicit judgment, news discourse can guide audience interpretation through implicit linguistic strategies. As cultural and entertainment news has expanded, scholars have also recognized reporting on cultural products as an important site of discursive construction (Bednarek & Caple, 2012). In recent years, some studies have examined how films and other entertainment products are represented in news media. From perspectives such as news values, evaluative language, and cultural framing, scholars have explored how film news balances information dissemination with the construction of cultural meaning (Kristensen & From, 2012). These studies suggest that film news reporting is neither purely promotional nor simply

critical commentary, but rather follows the norms and strategies of news discourse. However, corpus-based discourse studies of film news remain limited.

2.2. Corpus-assisted discourse analysis and co-selection

The development of corpus linguistics has provided new methods for studying media discourse. Unlike traditional discourse analysis, which often focuses on individual texts, corpus-based approaches identify recurrent linguistic patterns in large datasets, revealing relatively stable meaning tendencies (Baker, 2006). By examining high-frequency words, keywords, as well as patterns of collocation and co-occurrence, researchers can examine which meanings are foregrounded and which expressions become normalized. This improves analytical systematicity and reduces excessive reliance on subjective interpretation.

Increasingly, corpus methods have been combined with discourse analysis, forming corpus-assisted discourse analysis. For example, Baker et al. (2008) used keyword analysis to examine British media coverage of immigration and refugees, revealing implicit discursive patterns. Baker and Vessey (2018) compared extremist texts in English and French, identifying similarities and differences in discursive construction. Likewise, Wang and Ma (2021) showed how corpus analysis can reveal recurring meaning patterns and ideological implications in media texts. Together, these studies demonstrate that corpus methods effectively connect macro linguistic patterns with micro-level contextual interpretation.

Within this framework, Sinclair's theory of co-selection offers an important perspective on how linguistic meaning is generated. Sinclair (1991) proposed that language use largely follows the "idiom principle," whereby words tend to occur in fixed or semi-fixed combinations, rather than being freely selected according to the "open-choice principle." Building on this insight, Sinclair (2004) further introduced co-selection, emphasizing that lexical meaning is not carried by individual items in isolation, but is gradually formed and extended through the simultaneous operation of multiple linguistic choices. By analyzing keywords, collocations, semantic preference, and semantic prosody, researchers can identify how seemingly neutral language conveys evaluation and stance (Stubbs, 2001). This provides a useful framework for analyzing meaning construction in media discourse.

2.3. Media representation of China and cultural products

Corpus-assisted discourse analysis has been widely used to study national images and social issues, particularly representations of China. Most studies focus on politics, economics, and international relations, examining how Western media construct China and the ideological orientations embedded in such representations (Ospina Estupinan, 2017; Yu et al., 2023). For instance, Wang, Coler, and Redeker (2023), compared coverage of "Made in China" in Chinese and U.S. newspapers, revealing differences in keyword usage, semantic collocation, and discursive strategies. These studies provide important insights into understanding discursive variation across media contexts; however, systematic corpus-based comparisons between Chinese and U.S. media remain limited.

Compared with political issues, cultural products have received far less corpus-based attention. Although Chinese animated films have gained increasing international visibility, few studies have examined how Chinese and U.S. media construct the same Chinese cultural product through news discourse, or compared convergences and divergences through keywords and co-occurrence patterns. Therefore, a corpus-assisted discourse analysis of Chinese and U.S. media coverage of *Ne Zha 2* is needed to deepen understanding of cross-cultural differences in media discourse.

3. Methodology

3.1. Data

This study is based on two topic-specific news corpora: The China Daily corpus (the CD corpus) and the U.S. media corpus (the US corpus). These corpora respectively represent China's official external communication discourse system and the discourse of the U.S. entertainment industry. Their construction follows a discursive function priority, prioritizing influential institutional media within each discourse system rather than formal symmetry of nation or media type.

The Chinese media corpus (CD) draws exclusively on China Daily, the most representative national English-language newspaper in China. It functions as an official channel for China's external communication and national narrative construction, making it widely used in research on national image and international communication (Liu, 2015, 2017). Its style is also relatively close to mainstream English news discourse, supporting cross-cultural comparability (Mody, 2010).

The U.S. media corpus (US) consists of reports from Variety and Deadline, two leading U.S. entertainment industry publications. Major general newspapers such as The New York Times were excluded due to limited coverage of *Ne Zha 2*. In contrast, entertainment industry media play a key role in defining industrial value, shaping industry consensus, and guiding professional discussion in the global film industry (Hesmondhalgh, 2018; Deuze, 2007). Their sustained coverage of *Ne Zha 2* covers box office, industrial significance, and international impact, indicating its integration into global entertainment discourse (Curtin & Sanson, 2016). Thus, these sources better capture dominant professional interpretations than fragmented general news coverage (Fairclough, 1995).

Data were collected from January to July 2025, covering pre-release, release, box office growth, and subsequent discussion stages, allowing observation of discourse evolution (Baker, 2006; Partington et al., 2013). Reports were selected according to unified inclusion criteria and randomly sampled within eligible texts. All articles had to explicitly mention *Ne Zha 2* in the headline or body. Non-news elements such as captions, copyright notes, and author information were removed to ensure comparability (Baker, 2006; Baker et al., 2008; McEnery & Hardie, 2011). A summary of the basic characteristics of the two corpora in terms of text numbers and lexical size is presented in **Table 1**.

Table 1. General information about the two corpora.

	<i>China Daily</i>	US news media	
		<i>The Variety</i>	<i>The Deadline</i>
Number of texts	45	29	15
Tokens	23,265	12,924	10,286
Total	23,265	23,210	

3.2. Methodology

This study adopts the framework of corpus-assisted discourse studies (CADS), which combines quantitative corpus-linguistic techniques with qualitative discourse analysis to link macro-level linguistic patterns with contextual interpretation (Baker et al., 2008; Partington, 2004, 2015). In recent years, CADS has been widely used in media discourse research to examine how social issues and cultural events are linguistically constructed in news contexts (Baker et al., 2008; Baker et al., 2013; Koteyko et al., 2010; Partington et al., 2013).

Accordingly, the study integrates keyword, concordance, and collocation analysis with discourse interpretation. It examines both linguistic features and the interaction between language and social context, including stance and value judgment (Fairclough, 1995; Fowler, 1991; van Dijk, 1998).

Corpus analysis was conducted using Sketch Engine. Keyword lists were obtained by comparing each corpus with a large general English reference corpus, English Web 2021. Keywords are lexical items whose frequency in a given study corpus is statistically significantly higher than in the reference corpus (Baker, 2006). Keyword lists were ranked by keyness scores, with higher scores indicating stronger statistical significance.

For each corpus, the top 100 lexical keywords were extracted. Technical noise items and incomplete forms were manually removed. The remaining keywords were semantically categorized. This study adopts a corpus-driven analytical approach in which categories emerge from the data rather than prior theoretical frameworks (Tognini-Bonelli, 2001).

Representative keywords were then selected for collocation and concordance analysis. Collocates were identified within a span of five words to the left and right of the search term to reveal semantic tendencies. Concordance lines were used to examine actual usage patterns and avoid overgeneralization based solely on statistical results (Baker, 2006). As Morley and Bayley (2009) note, CADS depends on iterative comparison between keyword patterns and contextual discourse analysis.

Keyword categories followed three steps. First, themes emerging from the two corpora were cross-compared to assess consistency and overlap. Second, shared and corpus-specific keywords were identified, with attention to semantically related lexical items. Finally, categories were refined through contextual analysis. The results are presented in a matrix-style table showing shared keywords and corpus-specific salient keywords across thematic categories (see **Table 2**).

This study focuses on categories containing relatively large numbers of keywords or high cumulative keyword frequencies, treating them as core themes with dominant discursive functions. If a theme shows high and comparable relative frequencies in both corpora, it is treated as a shared theme; if it shows a markedly higher frequency in one corpus, it is treated as a salient theme specific to that corpus. Findings on shared and salient themes are presented in the following section.

4. Findings

This section first introduces the common categories. It then presents the more salient findings from the CD corpus, followed by Section 4.3, which outlines the more prominent results from the US corpus.

4.1. Common categories

As shown in **Table 2**, one category with relatively high frequencies in both the CD and US corpora is box office performance and market scale. It is especially prominent in the CD corpus (relative frequency = 9.67), while remaining important in the US corpus (4.39). This suggests that both Chinese and U.S. media tend to frame *Ne Zha 2* through market performance and commercial value. In the CD corpus, this category includes *highest-grossing*, *box-office*, *presale/pre-sale*, *yuan*, *maoyan*, and *blockbuster*; in the US corpus, it includes *box-office*, *highest-grossing*, *top-gross*, *surpass*, and *blockbuster*. Commercial success is therefore repeatedly constructed as a key dimension of the film's meaning. This supports film industry research showing that box office functions as a quantifiable evaluative mechanism and a core driver of media narratives (De Vany, 2003; Elberse, 2007; Vogel, 2020; Hesmondhalgh, 2018).

In the CD corpus, *box-office* is the most central item in this category. Its collocations frequently occur with concrete economic indicators and ranking-related terms such as *revenue*, *record*, *chart*, and *success*, forming a semantic field of "scale–record–performance." This anchors the film's value within a quantifiable and comparable framework. The corresponding concordance lines further illustrate this pattern, as in Examples (1) and (2).

(1) China's total box-office revenue for the holiday reached a new high, with "*Ne Zha 2*" contributing a significant share.

(2) The film has grossed nearly 15.87 billion yuan, becoming one of the highest-performing titles in the domestic market.

These contexts indicate that the media do not simply report figures, but embed the film within an industrial narrative based on totals, records, and market position. By contrast, in the US corpus, evaluative labels such as *highest-grossing*, *top-gross*, and *blockbuster* appear frequently, but *box-office* does not form stable collocational patterns under the current threshold settings. This suggests that U.S. media tend to compress economic processes into outcome-oriented labels rather than develop sustained narratives around specific indicators. Such practice aligns with conventions of entertainment journalism, where evaluation is often achieved through rankings or titles (Tryon, 2013; Vogel, 2020).

The second common category that is prominent in both the CD and US corpora concerns the animation industry and production context. In terms of keyword distribution, this category maintains relatively high salience in both corpora (CD relative frequency = 8.73; US = 5.77), indicating that the media generally interpret *Ne Zha 2* as a significant case within the animation industry rather than an isolated text. Both *animation* and *animated* enter the keyword sets, suggesting attention to industrial, generic, and textual dimensions. Previous studies note that media discussions of animation often involve genre production, industrial structure, and market expansion rather than form alone (Wells, 2002; Furniss, 1998; Lent, 2001; Lobato, 2019).

Table 2. Categories common to both CD and US corpora^A.

Category	Keywords in CD corpus	total frequency	relative frequency	Keywords in US corpus	total frequency	relative frequency
Box office performance and market scale	highest-grossing, box-office, presale, pre-sale, yuan, maoyan, blockbuster	225	9.67	box-office, highest-grossing, top-gross, surpass, blockbuster	102	4.39
Chinese cultural tradition and mythology	mythology, mythological, taiji, fengshen, wukong, jiaozi, yanyi	73	3.14	mythology, mythological, wuxia, chinese-language	3	0.13
Animation industry and production context	animation, animated, wofactory, enlight	203	8.73	animation, animated, non-hollywood, moviegoer	134	5.77
National and cultural identity markers	chinese, chinatown	319	13.71	chinese, chinese-language, chinatown	22	0.95

^AOnly representative keywords with keyness values ≥ 10.83 ($p < 0.001$) are listed. Categories with lower keyword counts or marginal frequencies are excluded to maintain analytical focus.

In the US corpus, the collocations and concordance lines of animation illustrate this clearly. The term frequently co-occurs with items such as market, release, and *blockbuster*, making "animation" a nodal concept linking the film to market expansion, distribution systems, and cross-regional circulation, as illustrated in Examples (3)–(5).

(3) Animation expanded to a further 12 markets following its strong domestic performance.

(4) Animation was used for the theatrical release of Chinese blockbuster "*Ne Zha 2*".

(5) The animation blockbuster "*Ne Zha 2*" continues to shatter records and expand its international reach.

These contexts interpret the success of *Ne Zha 2* as part of broader animation industry development rather than an isolated textual exception (Curtin, 2004; Lobato, 2019). This pattern also corresponds with the high relative frequency of animation-related categories in the CD corpus. Although company names (such as *wofactory* and *enlight*) do not appear symmetrically in the US corpus, animation as an industrial genre remains a shared interpretive framework (Wells, 2002; Furniss, 1998).

The third shared category concerns Chinese cultural tradition and mythology. In terms of frequency distribution, this category is more prominent in the CD corpus (relative frequency = 3.14) and includes a range of specific cultural proper nouns, whereas in the US corpus it is mainly composed of more general items such as *mythology* and *mythological*. This reflects different ways of mobilizing cultural resources: Domestic media tend to use detailed symbols, while overseas media rely more on generalized cultural labels for background information (Iwabuchi, 2002; Higson, 2000; Straubhaar, 2007). Collocational patterns of *mythology* in the US corpus show frequent co-occurrence with modifiers such as *Chinese*, serving to specify cultural origin. Examples (6) and (7) further illustrate its backgrounding function.

(6) The story draws on Chinese mythology, reworking familiar legends for contemporary audiences.

(7) Rooted in Chinese mythology, the film has also shattered box office records worldwide.

In these contexts, mythology functions less as a topic of extended discussion than as a recognizable cultural frame accompanying commercial narratives (Iwabuchi, 2002; Higson, 2000).

The fourth common category involves national and cultural identity markers. This category has the highest relative frequency in the CD corpus (13.71). Although it is less salient in the US corpus (0.95), it nevertheless appears consistently. Relevant keywords include *chinese* and *chinatown* in the CD corpus, and *chinese*, *chinese-language*, and *chinatown* in the US corpus. Although identity references appear in both corpora, their functions differ. The CD corpus more often links them to narratives of domestic cultural subjectivity and industrial development, whereas the US corpus tends to use them as descriptive labels of language or cultural attributes. The unilateral appearance of *chinese-language* in the US corpus suggests that language functions as a classificatory device for non-English texts. This aligns with studies of media representation and othering strategies (van Dijk, 1998; Wodak, 2001; Machin & Mayr, 2012). Concordance lines in the US corpus also frequently juxtapose identity markers with commercial contexts, as shown in Examples (8) and (9).

(8) "*Ne Zha 2*" leads Chinese New Year box office charts as audiences return to cinemas.

(9) The Chinese-language film continues to attract attention following its strong domestic performance.

Such juxtapositions render cultural identity not merely background information, but an integral dimension of the film's meaning construction alongside market performance and industrial framing (Machin & Mayr, 2012; van Dijk, 1998).

In sum, the common categories identified in the CD and US corpora show that media constructions of *Ne Zha 2* mainly rely on four recurring interpretive frameworks: commercial performance, animation industry positioning, cultural resources, and identity markers. The main differences lie not in whether these frameworks are present, but in the linguistic resources and organizational strategies used. For example, the

CD corpus forms a more stable indicator-based discourse through the collocations and numerical narration associated with *box-office*, whereas the US corpus shows a greater tendency to compress and organize information through outcome-oriented labels and classificatory markers (such as *highest-grossing*, *top-gross*, and *chinese-language*).

4.2. CD categories

As shown in **Table 3**, two thematic categories in the CD corpus have noticeably higher frequency and semantic concentration than in the US corpus, namely Market and industry structure and Festivals and release events. These categories do not form equally salient keyword sets in the US corpus, indicating that CD is more inclined to organize reporting on *Ne Zha 2* through structured industry- and event-based frameworks.

Table 3. Categories more frequently referenced in CD.

Category	Keywords in CD corpus	Total frequency and relative frequency
Market and industry structure	market, industry, cinema, theatrical, distribution	43 1.85
Festivals and release events	festival, release, screening, holiday	12 0.52

Within the Market and industry structure category, the relevant keywords occur a total of 43 times in the CD corpus (relative frequency = 1.85), making it one of the most prominent corpus-specific categories in CD. This category is organized around items such as *market*, *industry*, *cinema*, *theatrical*, and *distribution*, collectively pointing to the operational mechanisms and structural configuration of the film industry. Collocation analysis shows that *market* and *industry* in the CD corpus frequently co-occur with words denoting scale, structure, and institutional conditions, thereby embedding the film's success within a broader industrial narrative. Through this pattern, *Ne Zha 2* is presented not only as a commercially successful film, but also as a key case in the development of China's film market and animation industry.

This orientation is evident in the concordance lines. For example, *market* is often used to describe the film's position and influence within the overall film market, emphasizing its significance for the industry landscape (see Example (10)).

(10) The film's success reflects the growing maturity of China's domestic film market and its capacity to support large-scale animated productions.

At the same time, *industry* more frequently appears in discussions of the development and structural transformation of the animation industry, linking the film's success to broader processes of industrial upgrading and market maturation, as shown in Example (11).

(11) The animation industry has benefited from improved production capacity and a more diversified theatrical distribution system.

Through such contexts, the CD corpus elevates the commercial performance of an individual film into an explanation of market mechanisms and industrial systems, thereby constructing a highly institutionalized industrial discourse (Hesmondhalgh, 2018; Curtin, 2004). This practice resonates with previous findings that Chinese media reporting on cultural industries tends to adopt structured and macro-level narrative perspectives.

By contrast, while items such as *market* or *industry* appear in the US corpus, they do not form stable high-frequency categories or clear collocational patterns. This suggests that CD is more likely to interpret the film through industrial and market structures, whereas the US corpus is less likely to situate it

systematically within institutional discussions.

The second category more salient in the CD corpus is Festivals and release events. Keywords in this category occur 12 times in CD (relative frequency = 0.52). Although the absolute frequency is not high, the semantic concentration is strong, centering on items such as *festival*, *release*, *screening*, and *holiday*. Together, these items point to temporal nodes, exhibition contexts, and festive settings related to the film's release, reflecting an event-oriented narrative tendency.

Collocational and concordance analyses show that these keywords are often linked to release arrangements, holiday schedules, and viewing scenarios, repeatedly situating the film within symbolically meaningful temporal frameworks. For instance, *festival* commonly emphasizes the film's exhibition and circulation value in festive or event-based contexts, as illustrated in Example (12), while *holiday* links the film's release to moments of collective viewing, reinforcing socio-cultural significance (see Example (13)). In this way, the CD corpus does not merely report release timing, but draws on festival and release narratives to connect the film's success to broader structures of social time (Couldry, 2003; Dayan & Katz, 1992).

(12) The film was screened at several domestic animation festivals, highlighting its cultural significance beyond box-office performance.

(13) Released during the Spring Festival holiday, the film quickly became a popular choice for family audiences.

Although items such as *release* occasionally appear in the US corpus, their usage is largely backgrounded and does not form keyword clusters or stable semantic fields comparable to those in the CD corpus. This further highlights CD's emphasis on release events and temporal nodes, showing attention not only to market outcomes but also to how the film is embedded in institutional and social events.

Overall, the more salient categories in the CD corpus cluster around market structure, industrial systems, and release events. This suggests that CD is more inclined to interpret the significance of *Ne Zha 2* through macro-level industrial and event-based frameworks rather than solely through evaluation of the individual text.

4.3. US categories

As shown in **Table 4**, several thematic categories emerge in the US corpus that are not prominent in the CD corpus. These categories mainly concern film reviews and evaluative discourse, global circulation and overseas markets, and comparisons with Hollywood and the global animation system. Their relatively high frequencies and stable semantic fields indicate that US media tend to organize reporting on *Ne Zha 2* through evaluative, comparative, and transnational circulation frameworks.

Table 4. Categories more frequently referenced in the US corpus.

Category	Keywords in US corpus	Total frequency and relative frequency
Critical reception and reviews	review, critic, praised, reception, audience	112 3.91
Global distribution and overseas markets	international, global, overseas, release	74 2.58
Comparison with Hollywood and global animation	Hollywood, Disney, Pixar, comparable	53 1.85

Within Critical reception and reviews category, the relevant keywords occur a total of 112 times in the US corpus (relative frequency = 3.91), making it one of the most prominent corpus-specific categories in the US data. This category is mainly organized around items such as *review*, *critic*, *praised*, *reception*,

and *audience*, highlighting the film's reception and evaluative orientation. Collocation analysis shows that *review* and *critic* in the US corpus frequently co-occur with words expressing evaluative stance and aesthetic judgment, anchoring the film's meaning within frameworks of professional criticism and audience response. This orientation is particularly evident in the concordance lines. *Review* often appears in generalized assessments of the film's overall quality and style, emphasizing its positioning within review contexts, as shown in Example (14); meanwhile, *critic* is more commonly used to introduce professional judgments, situating the film within established critical discourse systems (see Example (15)). In this way, the US corpus tends to construe *Ne Zha 2* as a cultural text to be evaluated and interpreted, rather than primarily as an industrial or institutional phenomenon (Gray, 2010; Staiger, 2000).

(14) Documentary Short Film "Death by Numbers" by Kim A. Snyder received strong attention in reviews following its release.

(15) Winner of the inaugural Audience Award of Directors' Fortnight at Cannes, critics highlighted the film's visual ambition and narrative scale.

The second salient category in the US corpus is Global distribution and overseas markets. Keywords in this category occur 74 times (relative frequency = 2.58) and mainly include *international*, *global*, *overseas*, and *release*, pointing to the film's transnational circulation pathways and overseas market performance. Unlike the CD corpus, which emphasizes domestic market structures, the US corpus tends to situate the film within global circulation networks. Collocational and concordance analyses show that *international* and *overseas* frequently co-occur with references to distribution scope, market expansion, and cross-regional exhibition, repeatedly describing the film as "entering" or "expanding into" new markets. For instance, *international* is often used in concordance lines to highlight the film's circulation potential in overseas contexts, as shown in Example (16), while *overseas* further embeds the film within a global circulation framework that takes Western mainstream exhibition spaces as reference points, emphasizing symbolic recognition and institutional visibility in non-domestic markets (see Example (17)). This pattern aligns with studies of transnational circulation and recontextualization of cultural products (Lobato, 2019; Curtin, 2004).

[16] [...] appeared to ease some censorship restrictions, with two Hollywood films ending up in the Top 10 of the international box office.

[17] He described booking the film at London's iconic Odeon Leicester Square cinema, "where really only major U.S. studios are looking usually at James Bond and 'Wicked'," as an "inspiring moment" for Chinese cinema overseas.

The third category is Comparison with Hollywood and global animation, with keywords occurring 53 times in the US corpus (relative frequency = 1.85), including *Hollywood*, *Disney*, *Pixar*, and *comparable*. The salience of this category indicates that US media frequently draw on established Hollywood and global animation reference systems to interpret the film's meaning. These keywords are commonly used to compare the film with familiar industry benchmarks, providing audiences with accessible evaluative coordinates. For example, *Hollywood* often functions as a comparative framework, positioning the film within the global animation production system, as illustrated in Example (18), while references to *Disney* or *Pixar* further reinforce this comparative logic (see Example (19)). Through such referential narration, the film is incorporated into a Hollywood-centered global cultural hierarchy, with its meaning defined through comparison (Higson, 2000; Elsaesser, 2005).

[18] [...] lucrative for cinemas during the year in China, so its performance this week doesn't

necessarily signal a return to full Hollywood dominance.

[19] The film became the No. 2 animated movie ever globally, ranking behind Disney titles and placing as the No. 11 film overall.

Overall, the more salient categories in the US corpus cluster around evaluation, comparison, and global circulation. This suggests that US media are more inclined to treat *Ne Zha 2* as an object to be reviewed, compared, and positioned within the global cultural system, while engaging less systematically with industrial structures or institutional events. This further highlights the different meaning-construction pathways adopted by the CD and US corpora.

5. Discussion

Existing corpus-assisted discourse studies have mainly examined politics, the environment, and conflict, with less attention to cultural products, especially Chinese animated films in cross-cultural media contexts. By analyzing representations of *Ne Zha 2* in Chinese and U.S. media, this study extends corpus-based discourse research and provides evidence for understanding cultural positioning and ideological operations in news discourse.

The findings show that Chinese media (CD) and U.S. media (US) share several stable interpretive frameworks in reporting *Ne Zha 2*. Across both corpora, similar semantic resources are repeatedly used in relation to commercial performance, the animation industry, and Chinese cultural resources and identity. The film is thus embedded within broader narratives of market, industry, and cultural affiliation. This supports previous research that film news reporting is shaped by industrialized and institutionalized evaluative mechanisms (De Vany, 2003; Hesmondhalgh, 2018).

At the same time, the two corpora differ in linguistic organization and evaluative orientation. Chinese media tend to legitimate the film's success through quantitative indicators such as box office figures, market scale, and industrial milestones, presenting achievement as a verifiable industrial outcome. By contrast, U.S. entertainment media rely more on condensed outcome-oriented labels (e.g. *highest-grossing*, *top-gross*, *blockbuster*), using comparative language to position the film within the market. This reflects different conventions for defining and reporting success (Tryon, 2013; Vogel, 2020).

These differences relate to the distinct cultural positions of the two media systems. Chinese media place the success of *Ne Zha 2* within narratives of animation industry development, market maturation, and national cultural progress, contributing to broader discourses of cultural confidence and the international "going out" of Chinese culture (Keane, 2013; Curtin & Shah, 2010).

By contrast, U.S. media, located in a mature global animation market, have no structural need to frame the film's success as a domestic breakthrough. Instead, *Ne Zha 2* is represented as a notable market phenomenon from an "other" cultural system, whose significance is assessed through global circulation, transnational comparison, and existing industry standards. This reflects a center-media perspective in which value is confirmed through position within a global hierarchy (Higson, 1989; Iwabuchi, 2002). Accordingly, U.S. media focus more on the film's global market position than on its significance for Chinese animation.

Overall, there is no fundamental opposition between Chinese and U.S. media in whether *Ne Zha 2* is noticed or valued. The main difference lies in how shared themes are linguistically organized. Chinese media embed the film's success within a national industrial narrative, whereas U.S. media rely more on evaluative, comparative, and global circulation frameworks. This supports Baker and Vessey's (2018) view that cross-cultural discursive differences often arise less from topics themselves than from the adaptation of

linguistic resources to different audiences and communicative contexts.

Several limitations should be noted. The corpus includes only one Chinese English-language outlet and two U.S. entertainment publications, which limits representativeness. In addition, the analysis focuses mainly on textual patterns rather than non-textual factors such as editorial processes, audience positioning, or policy contexts. Future research could expand data sources, include more countries or languages, and combine corpus analysis with interviews or multimodal approaches.

Practically, the findings suggest that interpretations of the same cultural product are not symmetrical across media contexts. Cultural export is therefore not simply a matter of content translation, but also requires understanding the evaluative frameworks and discursive conventions of target contexts. Corpus-assisted discourse analysis can help identify such differences and support cross-cultural communication strategies.

6. Conclusions

This study examined how the Chinese animated film *Ne Zha 2* was discursively constructed in Chinese and U.S. English-language media through a corpus-assisted discourse analytical approach. By comparing coverage from China Daily, Deadline, and Variety, it identified shared patterns and divergent strategies in media meaning-making, and considered how these patterns are shaped by broader cultural and institutional contexts (McEnery & Hardie, 2012; Bell, 1991).

At the level of thematic selection, the findings show clear convergence between Chinese and U.S. media. Across both corpora, *Ne Zha 2* is mainly framed through box office performance, its position within the animation industry, references to Chinese cultural resources, and markers of cultural or national identity. Rather than being treated as an autonomous cultural text, the film is consistently embedded within market, industrial, and classificatory narratives. This supports the view that film news discourse relies on institutionalized evaluation mechanisms through which cinematic value becomes measurable and communicable across media systems (Vogel, 2020).

At the same time, the most salient differences lie not in what is reported, but in how meaning is linguistically organized. Chinese media tend to construct the film's success through extended quantitative indicators and industry references, presenting commercial achievement within a broader trajectory of domestic industry development. U.S. entertainment media, by contrast, rely more on condensed evaluative labels and comparative formulations to position the film within the global animation market. These differences reflect distinct conventions for reporting and legitimizing success rather than fundamentally different assessments of the film itself (Bell, 1991; Harrington, 2008).

Taken together, the findings suggest that entertainment news is an important site for negotiating cultural position and symbolic value. In the Chinese media context, reporting on *Ne Zha 2* translates the performance of an individual cultural product into an indicator of national cultural and industrial capacity. In the U.S. media context, the film is more often constructed as a noteworthy case from an external cultural system whose significance is established through global circulation, comparison, and classification. Cross-cultural differences in representation are therefore better understood as differences in discursive orientation and evaluative logic than as simple contrasts in thematic attention (Harrington, 2008).

Several limitations should be noted. The corpus is limited to one Chinese English-language newspaper and two U.S. entertainment publications, so the findings mainly apply to comparable professional news contexts. In addition, the analysis focuses on textual patterns and does not address non-textual factors such

as editorial processes, audience reception, or policy intervention. The conclusions should therefore be understood as grounded in linguistic evidence from news texts rather than as comprehensive explanations of media production (Baker et al., 2008).

Future research could expand the range of media sources, languages, and national contexts to test the broader applicability of these patterns. It could also combine corpus-assisted discourse analysis with interviews, ethnographic methods, or multimodal analysis to better understand how discursive patterns interact with institutional practices and audience interpretations.

About the author

Wei He (2005.03-), female, Han nationality, from Huizhou, Guangdong Province. She is currently an undergraduate student majoring in Business English.

References

- [1] Baker P, Gabrielatos C, McEnery T. (2013). Sketching Muslims: A corpus driven analysis of representations around the word 'Muslim' in the British press 1998–2009. *Applied Linguistics* 34(3): 255-278.
- [2] Baker, P. (2006). Using Corpora in Discourse Analysis. Continuum. <https://doi.org/10.5040/9781350933996>.
- [3] Baker, P., & Vessey, R. (2018). A corpus-driven comparison of English and French Islamist extremist texts. *International Journal of Corpus Linguistics*, 23(3), 255-278.
- [4] Baker, P., Gabrielatos, C., Khosravini, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & society*, 19(3), 273-306.
- [5] Bednarek, M., & Caple, H. (2012). News discourse.
- [6] Bell, A. (1991). The language of news media (pp. 84-85). Oxford: Blackwell.
- [7] Couldry, N. (2003). Media rituals: A critical approach. Routledge.
- [8] Curtin, M. (2004). Media capitals: Cultural geographies of global TV. *Television after TV: Essays on a medium in transition*.
- [9] Curtin, M., & Sanson, K. (2016). Precarious creativity. University of California Press.
- [10] Curtin, M., & Shah, H. (Eds.). (2010). Reorienting global communication: Indian and Chinese media beyond borders. University of Illinois Press.
- [11] Dayan, D., & Katz, E. (1992). Media events: The live broadcasting of history. harvard university press.
- [12] De Masi, V., Di, Q., Li, S., & Song, Y. The Ascendancy of Chinese Animation: An Analysis of Market Performance, Production Innovation, and Policy Support.
- [13] De Vany, A. (2003). Hollywood economics: How extreme uncertainty shapes the film industry. Routledge.
- [14] Deuze, M. (2007). Media work. Polity.
- [15] Elberse, A. (2007). The power of stars: Do star actors drive the success of movies?. *Journal of marketing*, 71(4), 102-120.
- [16] Elsaesser, T. (2005). European cinema: Face to face with Hollywood (p. 566). Amsterdam University Press.
- [17] Fairclough, N. (1995). Media discourse (pp. 9-14). London: Edward Arnold.
- [18] Fowler, R. (2013). Language in the News: Discourse and Ideology in the Press. Routledge.
- [19] Furniss, M. (1998). Art in motion: animation aesthetics. Indiana university press.
- [20] Gray, J. (2010). Show sold separately: Promos, spoilers, and other media paratexts. New York University Press.
- [21] Harrington, S. (2008). Popular news in the 21st century Time for a new critical approach?. *Journalism*, 9(3), 266-284.
- [22] Hesmondhalgh, D. (2018). The cultural industries.
- [23] Higson, A. (1989). The concept of national cinema. *Screen*, 30(4), 36-47.

- [24] Higson, A. (2000). The limiting imagination of national cinema. *Cinema and nation*, 63-74.
- [25] Huan, C. (2024). China opportunity or China threat? A corpus-based study of China's image in Australian news discourse. *Social Semiotics*, 34(5), 808-825.
- [26] Iwabuchi, K. (2002). *Recentering globalization: Popular culture and Japanese transnationalism*. Duke University Press.
- [27] Keane, M. (2013). *Creative industries in China: Art, design and media*. John Wiley & Sons.
- [28] Koteyko, N., Thelwall, M., & Nerlich, B. (2010). From carbon markets to carbon morality: Creative compounds as framing devices in online discourses on climate change mitigation. *Science Communication*, 32(1), 25-54.
- [29] Kristensen, N. N. (2012). Lifestyle journalism: Blurring boundaries. In *Lifestyle journalism* (pp. 25-40). Routledge.
- [30] Lent, J. A. (Ed.). (2001). *Animation in Asia and the Pacific*. Indiana University Press.
- [31] Liu M (2015) Scapegoat or manipulated victim? Metaphorical representations of the Sino-US currency dispute in Chinese and American financial news. *Text & Talk* 35(3): 337–357.
- [32] Liu M (2017) "Contesting the cynicism of Neoliberalism": A corpus-assisted discourse study of press representations of the Sino-US currency dispute. *Journal of Language and Politics* 16: 7–28.
- [33] Liu, M., & Li, C. (2017). Competing discursive constructions of China's smog in Chinese and Anglo-American English-language newspapers: A corpus-assisted discourse study. *Discourse & Communication*, 11(4), 386-403.
- [34] Lobato, R. (2019). *Netflix nations: The geography of digital distribution*. New York University Press.
- [35] Machin, D., & Mayr, A. (2012). *How to do critical discourse analysis: A multimodal introduction*.
- [36] McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- [37] Mody B (2010) *The Geopolitics of Representation in Foreign News*. Lanham, MD: Lexington Books.
- [38] Morley J and Bayley P (eds) (2009) *Corpus-Assisted Discourse Studies on the Iraq Conflict: Wording the War*. New York; London: Routledge.
- [39] Ospina Estupinan, J. D. (2017). The coverage of China in the Latin American Press: Media framing study. *Cogent Arts & Humanities*, 4(1), 1287319.
- [40] Partington A (2004) Corpora and discourse, a most congruous beast. In: Partington A, Morley J and Haarman L (eds) *Corpora and Discourse*. Bern: Peter Lang, pp. 11–20.
- [41] Partington A (2015) Corpus-assisted comparative case studies of representations of the Arab world. In: Baker P and McEnery T (eds) *Corpora and Discourse Studies: Integrating Discourse and Corpora*. New York: Palgrave Macmillan, pp. 220–243.
- [42] Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: theory and practice in corpus-assisted discourse studies (CADS)*. (Studies in corpus linguistics; No. 55). John Benjamins Publishing Company.
- [43] Richardson, J. E. (2017). *Analysing newspapers: An approach from critical discourse analysis*. Bloomsbury Publishing.
- [44] Rubing, G., & Sandaran, S. C. (2023). A critical discourse analysis of news discourse on in the times. *International Journal of Academic Research in Business and Social Sciences*, 13(1), 968-984.
- [45] Sincalir, J. (1991). *Corpus, Concordance, Collocation*.
- [46] Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge.
- [47] Staiger, J. (2000). *Perverse spectators: The practices of film reception*. NYU Press.
- [48] Straubhaar, J. D. (2007). *World television: From global to local*. SAGE publications.
- [49] Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. John Wiley & Sons.
- [50] Su, W. (2014). Cultural policy and film industry as negotiation of power: The chinese state's role and strategies in its engagement with global hollywood 1994–2012. *Pacific Affairs*, 87(1), 93-114.
- [51] Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam/Philadelphia, PA: John Benjamins. <https://doi.org/10.1075/scl.6>.
- [52] Tryon, C. (2013). *On-demand culture: Digital delivery and the future of movies*. Rutgers University Press.

- [53] Van Dijk TA (1998). *Ideology: A Multidisciplinary Approach*. London: SAGE.
- [54] Vogel, H. L. (2020). *Entertainment industry economics: A guide for financial analysis*. (10th ed.). Cambridge: Cambridge University Press.
- [55] Wang, C., Kerry, L., & Gao, X. (2025). *Resilience and Reflections: A Study of Chinese Cinema in the Wake of Global Challenges and Socio-Political Shifts*. *SAGE Open*, 15(3), 21582440251357163.
- [56] Wang, G., & Ma, X. (2021). Were they illegal rioters or pro-democracy protestors? Examining the 2019–20 Hong Kong protests in *China Daily* and *The New York Times*. *Critical Arts*, 35(2), 85-99.
- [57] Wang, Y., Coler, M., & Redeker, G. (2023). A contrastive study on the representation of "Made in China" in Chinese and US newspapers. *Frontiers in Communication*, 8, 1129376.
- [58] Wells, P. (2002). *Animation: Genre and authorship* (Vol. 13). Wallflower Press.
- [59] Wodak, R. (2001). The discourse-historical approach. *Methods of critical discourse analysis*, 1, 63-94.
- [60] Yang, M., & Wang, Z. (2023). A corpus-based discourse analysis of China's national image constructed by environmental news in *The New York Times*. *Humanities and Social Sciences Communications*, 10(1), 1-12.
- [61] Yu, Y., Tay, D., & Yue, Q. (2023). Media representations of China amid COVID-19: A corpus-assisted critical discourse analysis. *Media International Australia*, 191(1), 73-87.