

CONTENTS

- 1 Tech Powerhouse or Tech Threat? An Attitude Analysis of News Reports on Science and Technology in China by *China Daily* and *The New York Times*
Siyang Liu
- 16 A Comparative Study of *Oliver Twist* and *Jane Eyre* with the “Structure of Feeling” — The Contradiction Writing of the Victorian Era
Jing Hou, Lingling Xu, Tao Tao
- 24 A Conversation Analysis of Turn-Final *haobuhao* in Chinese Parent-Child Interaction
Shuyu Pan
- 33 Voice-Over as a Substitute for Visual Explanation in Short Videos
M. D. Reynolds
- 49 Composite Social Actions: Question-Intoned Repeat-Formatted Repair Initiation in Mandarin Conversation
Jian Guan

Tech Powerhouse or Tech Threat? An Attitude Analysis of News Reports on Science and Technology in China by *China Daily* and *The New York Times*

Siyang Liu¹

¹ School of Foreign Languages, Dalian Minzu University, Dalian 116600, China

Correspondence: Siyang Liu, School of Foreign Languages, Dalian Minzu University, Dalian 116600, China.

doi:10.63593/JLCS.2025.12.01

Abstract

This study investigates how science and technology (S&T) in China is represented through attitudinal resources by *China Daily* (CD) and *The New York Times* (NYT). 215 news articles from CD and 133 from NYT were analyzed according to a revised system of attitude within the appraisal framework. Results show that CD features overwhelmingly positive evaluation, highlighting the value of China's technologies and scientific research as well as China's tech strength. The competence of Chinese scientists in striving for innovation excellence is also foregrounded. By contrast, NYT displays both positive and negative attitudes in its reports. On the one hand, it emphasizes China's prowess and determination in developing high-tech industries as well as the utility of China's tech products. On the other hand, it depicts China's sci-tech endeavors as unethical and worrisome, and underlines the weakness and shortcomings of China's technologies. The significance of this study lies in its focus on the cross-cultural comparison of science news and the media representation of China's science and technology.

Keywords: attitude, science and technology in China, *China Daily*, *The New York Times*

1. Introduction

As the U.S. Treasury Department announced a new round of restrictions on China's technology investment in October 2024¹, the Sino-US technological competition has reached an unprecedented level. This sci-tech rivalry is now spilling over into the ideological realm, aggravating the discourse power competition between the two major countries around science and technology (Xie & Jiang, 2025). As carriers

of dominant ideologies, news media, especially mainstream media, is one of the main battlefields of the international discourse power game (Jungherr et al., 2019). In the field of international public opinion, American media have been relentlessly intensifying the construction of "China tech threat" narrative, while Chinese media stress that China's scientific and technological development is to benefit Chinese people and the whole world.²

¹<https://home.treasury.gov/news/press-releases/jy2687>

²<https://www.chinadaily.com.cn/a/202305/31/WS64769214a3107584c3ac304a.html>

As the backdrop of Sino-US tech competition is shaping media narratives, it is worthwhile to take a closer look at the representation of science and technology in China by mainstream news outlets from the two nations. *China Daily* (CD) and *The New York Times* (NYT), which stand out as two leading English-language newspapers from each side, are chosen as the news source of the study. The representation will be analyzed by examining evaluative language in the news reports, as it functions to “reflect the value system” which is “a component of the ideology” (Thompson & Hunston, 2000: 6). Of all the different approaches to analyzing evaluation in news discourse, the appraisal framework (Martin & White, 2005) has proved to be a powerful and effective tool (Huan, 2017; He & Caple, 2020) and is therefore adopted in this study.

The significance of this study lies in its focus on the cross-cultural comparison of science news and the media representation of China’s science and technology. The lack of cross-cultural analyses of science news reporting has been pointed out by Vestergaard and Nielsen (2016). Also, there has not been much scholarly attention to the representation of science and technology in China in spite of its frequent coverage in the news media. In view of this, the present study draws on the attitude system within the appraisal framework to explore how CD and NYT represent China’s science and technology through attitudinal resources. The underlying factors that might have influenced their representations will also be discussed.

This article consists of six sections. Section 2 introduces the context in which the study is situated. Section 3 briefly reviews previous literature on science news by focusing on ideology. Section 4 explains the data and methodology. Section 5 presents analysis results and findings. A discussion of the results and a conclusion are given in Section 6.

2. China’s S&T Endeavors and the US-China Tech Rivalry

In 2016, China released the Outline of the National Strategy of Innovation-Driven Development¹, which has set goals for the country to become an innovative nation by 2020, be in the front ranks of innovative countries by 2035, and a global scientific power by 2050. With

a strong determination to become a world leader in science and technology, China has implemented a series of strategic initiatives such as “Made in China 2025”, “Science and Technology Innovation 2030 Major Projects”, “Thirteenth Five-Year Development Plan for National Strategic Emerging Industries”. These state-led efforts have yielded results in a number of cutting-edge sci-tech areas, including information and communication technologies, artificial intelligence, aerospace engineering, biotechnology and so on². China has also rolled out measures to encourage the cultivation of sci-tech talents, strengthen protection of intellectual property rights, promote the commercialization of scientific achievements and expand international sci-tech cooperation. With these endeavors, China is no longer considered a technological follower but on its way to becoming an innovation powerhouse (Li et al., 2020). According to the Global Innovation Index released annually by the World Intellectual Property Organization, China has seen a steady rise of its ranking in the index, climbing from the 34th in 2012 to the 11th in 2022 (Dutta et al., 2022).

China’s search for technology and innovation, however, has been considered worrying from the perspective of the United States which often takes China as a rival. On August 14, 2017, the White House issued the Memorandum for the United States Trade Representative, claiming that China’s laws, policies, practices, and actions related to intellectual property, innovation, and technology “would negatively affect American economic interests” and “undermine American manufacturing, services, and innovation”³. In a 35-page report released by The White House Office of Trade and Manufacturing Policy in June 2018, China was accused of engaging in forced technology and intellectual property transfer, state-supported technology acquisition, and cyber-enabled theft of technologies and IP⁴, which “challenged US strategic interests in national security and international order” (Kennedy & Lim, 2018: 561). Due to its intensifying concern over China’s innovation

¹http://www.china.org.cn/china/2016-05/20/content_38494125.htm

²https://english.cas.cn/newsroom/cas_media/202111/t20211104_290989.shtml

³<https://www.gpo.gov/fdsys/pkg/FR-2017-08-17/pdf/2017-17528.pdf>

⁴<https://trumpwhitehouse.archives.gov/wp-content/uploads/2018/06/FINAL-China-Technology-Report-6.18.18-PDF.pdf>

activities and technology ambitions, the Trump administration launched a tech war against China, with an intention to “diminish and postpone China’s challenge to the U.S.’ global hegemony by containing the development of China’s high-tech industries” (Sun, 2019: 199). The US government has since imposed continuous crackdown on China’s tech sector by blacklisting Chinese tech companies such as Huawei and ZTE from US investment and limiting their access to advanced American technology. In response, China took countermeasures against the US sanctions, while at the same time stepped up sci-tech self-reliance to break Washington’s scientific blockade.

The US-China tech rivalry is believed to be driven by the two countries’ divergence in social values, political/economic systems and ideologies (Mori, 2019; Sun, 2019; Wu, 2020). Holding an ideological bias towards China, the US sees China’s efforts to become a sci-tech power as “an existential threat to a presumption of American dominance in all aspects of technology” (Inkster, 2019: 109). US policymakers even put forward the notion of China’s tech-enhanced authoritarianism, which Sun (2019) has argued is a new form of politicization of economic and technological issues. As geopolitics and technology are increasingly intertwined, technology is in fact largely politicized and has become a more prominent element of the escalating rivalry between China and the US (Wu, 2020).

3. Ideology and Science News Reporting

It is widely acknowledged that news reports, as specific social and discursive practices, are not unbiased recording of social reality but rather influenced by various sociocultural and political factors (Fowler, 1991; Fairclough, 1992; Van Dijk, 1995). One major factor is ideology, which is a kind of “fundamental, axiomatic beliefs underlying the social representations shared by a group, featuring fundamental norms and values” (Van Dijk, 2009: 193). According to Van Dijk (2009), news discourses are imbued with ideologies and can contribute to the (re)production of ideologies in society.

Some scholars have pointed out that science news, which is usually assumed to be an objective representation of scientific reality, is in fact also influenced by the ideological standings of news producers, just like any other types of news. For instance, Carvalho (2007) argues that

different ideological cultures among British quality newspapers largely affect their representations of scientific knowledge on climate change. According to Carvalho (2007), the reporting by *The Times*, a news outlet with an ideological culture of neo-liberal capitalism, is characterized by skepticism about climate change and opposition to the continuation of policies and practices to address climate change. In contrast, influenced by a social democratic ideology, reports by *The Guardian* and *The Independent* tend to highlight the potential risks of climate change to environments and defend mitigating action in relation to climate change. The link between ideology and climate change reporting is also pointed out by Dirikx and Gelders (2010). A more recent study by Yu and Liu (2020) investigates how ideologies affect US news coverage on artificial intelligence (AI), and finds that Conservative-oriented reporting focuses on the shortcomings of AI and the potential threats it poses to society, whereas the Liberal-leaning reports foreground AI’s advantages. Such political standings in the media may give rise to politicization in science news reporting, with political topics, actions and actors featured more often than science-related ones in news articles (Chinn et al., 2020; Litvinenko et al., 2022). Other scholars find that non-scientific events such as political affairs or policy-making are likely to be triggers for science news (Elmer et al., 2008; Vestergaard & Nielsen, 2016).

Previous studies on the representation of S&T tend to highlight micro-level factors such as personal preferences of science journalists (Stocking & Holstein, 2009), journalistic norms (Boykoff & Boykoff, 2004), news values (Badenschier & Wormer, 2011), organizational arrangements (Guenther et al., 2019) and so on. The impact of ideological factors on the selection and representation of S&T in general is still under-explored. What’s more, prior research has focused on the representation of S&T by newspapers with different ideologies from the same sociocultural background, and comparative cross-cultural studies of science news are rare. For this reason, the present study will investigate the representation of S&T in China by CD and NYT, two mainstream newspapers from China and the US. Given the different socio-political environments in China and the US, which may lead to divergent ideological standpoints of Chinese and

American media, comparison between science news coverage in these two countries is likely to reveal meaningful findings. The representation will be analyzed through the lens of evaluation, as it functions to express a communal value system which in turn reveals the ideology of the society (Thompson & Hunston, 2000).

4. Data and Methodology

4.1 Data

This study collected news reports on science and technology in China published by *China Daily* (CD) and *The New York Times* (NYT) between 1 January 2013 and 31 December 2022. News reports by CD were collected from the ‘China-innovation’ section on its official website¹. As the number of news reports by CD is much higher than that of NYT, data sampling was conducted by adopting the constructed week sampling method. Constructed week sampling is a type of stratified random sampling technique popular in media studies in which the final sample represents all seven days of the week to account for cyclic variation of news content (Luke et al., 2011). As is recommended by Song and Chang (2012), one constructed week would allow reliable estimates of news content in a six-month population for daily newspapers in China, even for highly volatile variables. Therefore, we constructed two weeks for each year, obtaining 140 days’ reports between 2013 and 2022. After duplication removal, 215 news reports totaling 78,072 words were collected from CD. News reports by NYT were collected from the science and technology sections of its website² by using ‘China’ and ‘Chinese’ as the search terms. All the retrieved articles were read through carefully to exclude duplicates and irrelevant ones such as reports on S&T in other countries or reports on issues unrelated to S&T. This resulted in 133 news articles with 141,081 words in total. The collected news reports were then constructed into two corpora: the CD corpus and the NYT corpus. The composition of the data is shown in Table 1.

NYT	133	141,081
-----	-----	---------

4.2 Analytical Framework and Procedures

The collected reports were analyzed according to a revised system of attitude within the appraisal framework in systemic functional linguistics (SFL). Developed by Martin and White (2005), the appraisal framework consists of three systems: attitude, engagement, and graduation. Among them attitude is the kernel and is “concerned with our feelings, including emotional reactions, judgements of behavior and evaluation of things” (ibid: 35). The system of attitude is further divided into affect, judgement and appreciation. Affect construes emotional reactions and can be sub-categorized into un/happiness (e.g. happy/sad), in/security (e.g. confident/anxious), dis/satisfaction (e.g. pleased/angry), dis/inclination (e.g. willing/reluctant). Judgement deals with the evaluation of human behaviors according to social esteem and social sanction. Judgement of social esteem concerns normality (how unusual someone is), capacity (how capable someone is) and tenacity (how resolute someone is), while judgement of social sanction involves veracity (how truthful someone is) and propriety (how ethical someone is). Appreciation is the evaluation about things, which can be further categorized into reaction, composition and valuation. Reaction refers to the degree to which things catch our attention. Composition is about the internal structure of things, such as balance and complexity, and valuation is to do with our assessment of the social significance of the phenomenon. Attitude can be positive or negative, and it can be expressed directly or indirectly. If the attitude is construed directly through attitudinal expressions, it is “inscribed”. If it is expressed indirectly through lexical metaphors or ideational tokens, the attitude is “invoked”.

Table 1. Data composition of the study

Corpus	Number of news reports	Total words
CD	215	78,072

¹<https://www.chinadaily.com.cn>

²<https://www.nytimes.com>

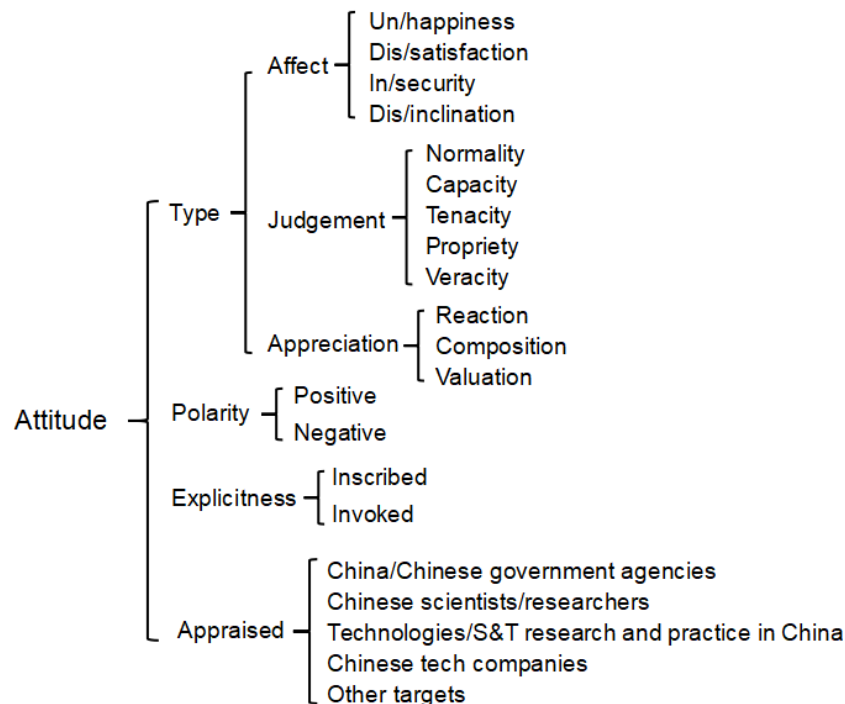


Figure 1. The revised system of attitude applied in this study

Apart from the type, polarity and explicitness of attitude, a new category of “appraised” was added to the framework for further specifying the person, thing or activity that are judged and appreciated, given that we intend to investigate the discursive use of attitudinal resources in representing science and technology in China. As we coded the data, some recurrent targets of evaluation were detected. Based on these targets, we specified “appraised” into five categories: China/Chinese government agencies, technologies/S&T research and practice in China, Chinese scientists/researchers, Chinese tech companies and other targets. As for affect, which deals with people’s emotional response and is not a prominent attitude category in science news, we decided to code the type and polarity of affect only. In example (1), positive valuation of China’s technology, i.e. the counter-drone system, is directly inscribed through the attitudinal adjective “useful”, while the fact stated in example (2) invokes negative judgement of China’s capacity in producing home-made optical network technology.

- (1) “The system is *useful* in airport, border control, drug and smuggling suppression, and security work in nuclear plants and large public gatherings,” Shi added.
- (2) And the new Chinese supercomputer *is still*

based on American-made optical network technology that is used to connect the thousands of microprocessor chips.

The collected data was imported into the UAM CorpusTool¹ and manually annotated according to the revised framework presented in Figure 1. The data was coded by two assistants familiar with the attitude system separately and inter-coder reliability revealed a Krippendorff’s alpha score of 0.899. All cases of disagreement were revisited by the author, and discussions were held until a final decision was made. During this process, *The Language of Evaluation* (Martin & White, 2005) and *Handbook for Analyzing Chinese and English Appraisal Meanings* (Peng, 2015) were frequently consulted for help. After the analysis, UAM can perform automatic summarization and retrieval of the analysis results.

5. Findings

The distribution of attitudes constructed in the CD and NYT corpus in terms of polarity and type is summarized in Table 2. It can be seen that the reports by CD display an unbalanced distribution of positive and negative attitudes: almost all the attitudes are positive (97%), and few negative aspects of the scientific events are

¹<http://www.corpustool.com/index.html>

mentioned (3%). Positive capacity and valuation are the most prominent. Among the three types of attitude, appreciation is the most frequently represented (54%), closely followed by judgement (41%), while affect shows very low percentages (5%). In contrast, reports by NYT combine both positive and negative attitudes. A somewhat surprising finding is that more than half of the attitudes expressed in NYT reports are positive (61%), among which positive capacity, tenacity and valuation are highlighted.

NYT contains many more negative attitudes than CD does (39% and 3% respectively), with a prime focus on negative propriety, capacity, valuation and insecurity. 72% of the attitudes constructed in NYT reports fall into the category of judgement, far beyond the proportions of appreciation and affect (18% and 10% respectively). The prominent attitude features and their realizations in the reports by CD and NYT are discussed in the following.

Table 2. Distribution of attitudes in the corpus

		CD		NYT	
		Positive	Negative	Positive	Negative
Affect	Happiness	+12	-0	+3	-5
	Satisfaction	+1	-0	+2	-5
	Security	+11	-3	+7	-84
	Inclination	+20	-0	+21	-3
	Sum	N=47 (+44, -3), 5%		N= 130 (+33, -97), 10%	
Judgement	Normality	+0	-0	+5	-14
	Capacity	+278	-14	+337	-96
	Tenacity	+65	-0	+179	-4
	Propriety	+62	-0	+50	-222
	Veracity	+0	-0	+0	-3
	Sum	N=419 (+405, -14), 41%		N=910 (+571, -339), 72%	
Appreciation	Reaction	+15	-3	+16	-4
	Composition	+4	-1	+4	-5
	Valuation	+533	-10	+156	-50
	Sum	N=566 (+552, -14), 54%		N=235 (+176, -59), 18%	
Total		+1001 (97%)	-31 (3%)	+780 (61%)	-495 (39%)
		1,032 (100%)		1,275 (100%)	

5.1 CD's Representation of S&T in China

5.1.1 China's Useful Technologies and Valuable Scientific Research: Positive Valuation

In terms of the distribution of attitudes, the most remarkable feature of the CD corpus is its highlighting of positive valuation, which accounts for more than 50% of all the positive attitudes (see Table 2). Positive valuation mainly concerns technologies and scientific research in China, and is frequently realized by emphasizing the performance or function of technological products. This is illustrated through examples (3)-(5) where the robots, the icebreaker and the satellite are presented as

capable of performing complicated tasks, thus constructing positive appreciation of the technology. In some other cases, positive valuation is inscribed through positive adjectives such as "advanced", "innovative", "effective", "intelligent", "cutting-edge" and expressions containing superlatives such as "the biggest and best of its kind", "the fastest ground vehicle in the world", "the largest and most sensitive radio telescope".

- (3) With the hands, robots *can grasp and place different objects steadily and accurately* after judging the shape and size of the item.
- (4) Incorporating state-of-the-art design and

advanced equipment, the icebreaker *is capable of operating in any ocean*.

- (5) “Through that mission, we improved the recoverable satellite platform. Currently, our satellite *is able to return 250 kg payloads back to Earth from space*,” Zhao said.

Another strategy to realize positive valuation is to emphasize the utility of Chinese technologies by highlighting their application in solving problems and improving people’s livelihood. For instance, CD has reported on how

temperature screening robots are used to assist the epidemic control in China, AI monitoring platform to protect wildlife, Beidou-backed app to help herdsman track cattle in pastures, meteorological satellites to advance weather forecast, etc. In addition, technologies or scientific research are frequently collocated with the verb “help*” to construct positive valuation. Some concordances with “help*” as the node are presented in Figure 2, where the subjects of “help*” are highlighted in red.

1	thereafter, said Duan. In-orbit 3D printing is effective in	helping	with space station repair and maintenance and is essential
2	to that of Mars. Our space biology research will also	help	us understand possible life forms on Mars,” Wang said.
3	China’s overseas medical assistance mission, Eyebright’s products have	helped	people in many countries, including Cambodia, Sudan and Congo,
4	planet, algae might be our pioneer,” Wang said. New technology	helps	researchers decrease emissions at coal plants Researchers at Zhejiang
5	and recovery technologies at sea. The seaborne launch technology will	help	China provide launch services for countries participating in the
6	the whim of the user, he said. The technology can	help	in life-threatening missions as well as in rehabilitation
7	vice chairman of Nuctech. In 2016, a just-upgraded Nuctech equipment	helped	Australian Border Security seize 254 kg of cocaine and 104 kg
8	me to work at a chip company.” Smart tech system	helps	communities ward off virus China’s tech powerhouses are
9	s Yanchang oilfield, Shaanxi TV reported on Monday. The technology	helped	discover new oil reserves in an oilfield that has
10	contributions to the astronomic research in the world.” New technology	helps	find over 800 m tons of oil The deep oil
11	a Shanghai hospital. With the 5G network, the robots can	help	medical staff carry out tasks involving consultations, disinfection, cleaning
12	rely on isolated solar power and hydropower grid. Unique tech	helps	resolve water shortage in Beijing Looking at the Chanping
13	of fireworks a day - one box every 15 seconds. High-tech	helps	spot distracted driving in SW China city An intelligent

Figure 2. Some concordances with “help*” as the node

- (6) “The instrument *is a good start and can provide experience for the future development of a larger coronagraph*,” said CAS academician Fang Cheng.
- (7) Other recent scientific and technological breakthroughs include the launch of the Xihe solar observation satellite and..., which *have significantly bolstered* China’s quantum computational advantage.

Apart from positive performance and utility, technologies and sci-tech research in China are also evaluated positively in terms of their social significance. As illustrated in examples (6) and (7), technologies and scientific breakthroughs in China are represented as valuable in promoting the future development of scientific research in a certain technical field and in contributing to China’s efforts to become a technological powerhouse.

5.1.2 A Powerful China and Capable Chinese Scientists: Positive Capacity

Table 2 shows that the second predominant

attitude type in the CD corpus is positive capacity, which mainly concerns China/Chinese government agencies and Chinese scientists/researchers. One way to realize positive capacity is to emphasize China’s status as an emerging force in science, technology and innovation, or as a world sci-tech power. This is largely achieved through appositive nominal phrases referring to China as “a scientific superpower in the making”, “front-runner”, “leader”, “an undeniable actor in spearheading global innovation”, etc. Another way is to highlight the capability of China in achieving scientific and technological successes. This can be illustrated through the following examples.

- (8) The feat proved that China *had acquired “full ocean depth capability”*, meaning that it *can carry out routine manned expeditions and scientific exploration at any depth in any ocean*.
- (9) Last year, China *made many influential breakthroughs* in basic research, ranging from creating the world’s first synthesized

single-chromosome yeast to the launch of the China Spallation Neutron Source, Wang said.

As can be seen in these examples, China is presented as a powerful and leading tech powerhouse, capable of making its homegrown innovation and mastering cutting-edge technologies. In addition, expressions emphasizing the advancement and firstness of Chinese technologies such as “the world’s first meteorological satellite in dawn-dusk orbit” and “the world’s largest single-dish radio telescope” are often mentioned in the corpus, which potentially invokes a positive judgement of China as it has accomplished something other countries haven’t. In the reports by CD, positive capacity is also represented through Chinese scientists/researchers’ faculty and creativity in developing new technologies and achieving breakthroughs in key technical fields. They are depicted as offering new solutions, carrying out scientific tests, and reporting latest research progress in many disciplines such as medicine, physics, astronomy, material science and information technology. Some examples are presented below.

- (10) Hong Kong medicine scholars *develop innovative targeted delivery system* for treating bone tumor.
- (11) In December, the journal Science reported that Chinese scientists *have created the world’s first photonic quantum computer prototype*, called Jiuzhang...

As illustrated in examples (10) and (11), positive capacity of scientists is predominantly expressed through a material process with Chinese scientists as the Actor and scientific discoveries as the Goal, where verbal groups such as “have developed, have created, realized, achieved, innovated” are often used to suggest their competence.

5.2 NYT’s Representation of S&T in China

5.2.1 A Competitive Yet Incapable China: Positive & Negative Capacity

As shown in Table 2, positive capacity is foregrounded of the attitudinal resources in the reports by both CD and NYT, taking up 27% and 26% respectively. Similar to the strategies used in the CD corpus, positive capacity of China/Chinese government in the NYT corpus highlights the country’s competence and competitiveness in S&T. This is mainly realized

by 1) using positive labels to refer to China and 2) representing China as taking the lead and making progress in the sci-tech field. This can be illustrated through examples (12) to (14). It is worth noting that about 40% percent of positive capacity concerns China’s success in the field of aerospace, probably because China has achieved many breakthroughs in exploring outer space in recent years.

- (12) China Extends Lead as *Most Prolific Supercomputer Maker*
- (13) China, Not Silicon Valley, *Is Cutting Edge* in Mobile Tech
- (14) In December 2020, it [China] *became the first country in about four decades to bring back lunar rocks and soil*, amassing several pounds of samples, experts said.

There are 72 instances of positive capacity pointing to Chinese tech companies. Most tech companies judged positively are elite tech giants active in the fields of A.I. and information technology, such as Huawei, Baidu, Alibaba, ZTE and IFlyTek. These tech companies are usually mentioned in NYT reports along with descriptions of their importance and high position (see example (15)). In other cases, they are cast as successful and competitive in developing innovative technologies and achieving technological advances (see example (16)).

- (15) The company [Huawei], *China’s leading maker of telecommunications equipment and smartphones*, has found itself at the center of...
- (16) Two years before Microsoft did, Baidu, the Chinese internet search company, *created software capable of matching human skills at understanding speech*.

Different from CD’s overwhelming highlighting of positive capacity, NYT also contains negative capacity in its reports. Most negative capacity is realized by depicting China’s lagging behind the developed countries in core sci-tech areas and its reliance on advanced foreign technologies. Examples (17)-(19) are three representative instances indicating that the technology gap between China and Western countries remains large.

- (17) China *looks to the West for much of its technology*.
- (18) China still *lags in important areas*. Its most powerful, high-end servers and

supercomputers often *rely in part on American technology*.

- (19) China's technology boom, it turns out, has been largely *built on top of Western technology*.

In addition, NYT often states that China imports chips and semiconductors in huge quantities from other countries, to invoke negative evaluation that China is unable to produce its own high-end chips and semiconductors. In other instances, NYT highlights China's failure in its scientific and technological endeavors (see examples (20)-(22)).

- (20) ...it [China] *failed to* turn scientific knowledge into usable technology.

- (21) ...most notably, China *cannot make* reliable engines or avionics, he said.

- (22) Military analysts say China *has long tried to replicate* foreign drone designs.

Examples like these contribute to constructing a negative image of a failing China unable to realize high-quality tech development and self-dependent innovation.

5.2.2 Useful but Imperfect Chinese Technologies: Positive & Negative Valuation

It can be seen from Table 2 that valuation is the most prominent subtype of appreciation constructed in the reports by NYT. Different from the reports by CD where almost all instances of appreciation are positive, those by NYT contain both positive and negative valuations of technologies and S&T research/practice in China. A close reading of the instances shows that positive valuation, which accounts for about 67% of the instances of appreciation, mainly functions to highlight the value of China's high-tech products, especially those in the fields of information and communication, such as cellphones, supercomputers, drones, AI software. Example (23) represents the usefulness and importance of drones in China for their wide range of applications.

- (23) Drones in China *inspect power lines, survey fires and disaster zones, spray crops, and monitor air pollution around factories*. In some remote areas, they *have delivered packages*.

- (24) These and other technical limits mean that China's two aircraft carriers *will not be able to range out to sea for as long as American carriers, nor will the jets that fly from them be*

able to carry as much munitions and fuel as American carrier fighters.

Almost 21% of the instances of appreciation fall into the subcategory of negative valuation. Expressions of negative valuation center upon the worthlessness and limitations of technologies in China. Example (24) evaluates China's aircraft carriers as less capable than American carriers in terms of performance. Apart from describing what China's technological products cannot do through verbal groups (e.g. failed to, will/would not be able to), the negative valuation is also often realized through adjectives (e.g. imperfect, impractical, vulnerable, not powerful enough, ineffective, unable) and nouns (e.g. shortcomings, flaws, limitations).

5.2.3 A Determined China and the Ambitious Chinese Government: Positive Tenacity

Positive tenacity is also a prominent feature in the reports by NYT. Positive markers of tenacity are mainly deployed to represent China/Chinese government agencies as determined to achieve self-reliance and prosperity in key technical fields and aspirant to become a global leader in core sci-tech areas. For example, China is frequently represented as striving to expand its tech capabilities in such instances as "China Seen in Push to Gain Technology Insights", "the country's bid to be at the forefront of quantum research", "China's drive to become a leading exporter of military equipment", etc. The two words "ambition(s)" and "ambitious" appear 68 times altogether in the NYT corpus. Many of these instances refer to China's technology ambitions and its ambitious programs (e.g. Made in China 2025) in areas like space, chip industry and AI, which help to construct positive judgement of China's tenacity. In addition, the reports by NYT provide information on Chinese governmental efforts in promoting high-tech development, especially its heavy investment in cutting-edge tech areas, which indicates China's resolve to rise as a technological powerhouse. This can be seen from example (25).

- (25) The government in Beijing *has poured the equivalent of billions of dollars* into new projects in order to catch up with the West in producing original research...

Apart from illustrating China's dedication and determination to become a scientific power, positive tenacity is also achieved by

representing the country as careful and patient, reaching its goal in a step-by-step way and avoiding risky technological leaps.

5.2.4 An Unethical and Irresponsible China: Negative Propriety

The most salient feature of the negative attitudes constructed in NYT reports is negative propriety, which occupies 45% of all the negative attitudes (see Table 2). Most negative propriety serves to negatively judge China and Chinese government agencies, and it is mainly achieved by 1) focusing on China's involvement in unethical S&T activities and misuse of advanced technologies, and 2) representing China as failing to meet responsible standards in sci-tech practice. For instance, China is often accused of "stealing" intellectual property and valuable foreign technological know-how from American firms, "transferring the science and technology of the developed world" via illegal methods and forcing foreign tech companies to hand over their latest technologies so as to support the development of local companies. These criticisms tend to be expressed quite explicitly through negative lexis such as "steal", "copy", "theft", "illicitly", etc. Examples (26) and (27) are two representative cases.

(26) Foreign governments and companies also say that China's technological advances have too often come from *illicitly copying foreign advances* — increasingly through Internet hacking.

(27) China *has sought technologies to block American surveillance and communications satellites*, and its major investments in cybertechnology — and probes and *attacks on American computer networks* — are viewed by American officials as a way to both *steal intellectual property* and prepare for future conflict.

In addition, China's misuse of advanced technologies is also frequently described in the NYT corpus. When reporting on China's surveillance technology, NYT portrays China as "a digital totalitarian state" and the Chinese government as repressive and inhuman in using techno-authoritarian tools to monitor its people, especially the minorities. This can be illustrated in the following two examples.

(28) China is ramping up its ability to *spy on its nearly 1.4 billion people* to new and disturbing levels, giving the world a blueprint for how to build *a digital*

totalitarian state.

(29) Now, documents and interviews show that the authorities are also *using a vast, secret system of advanced facial recognition technology to track and control the Uighurs*, a largely Muslim minority.

In some other cases, the negative consequences of China's wrongdoings are emphasized, such as the expanding digital surveillance will "violate privacy", lead to "a future of tech-driven authoritarianism", and usher in "a new era of automated racism", etc. When reporting on China's space exploration, apart from emphasizing the country's capability, NYT also frequently reports on China's irresponsibility in dealing with its space debris, criticizing China of deliberately launching mammoth rockets to orbit and allowing them to fall out of control to Earth where they may.

5.2.5 Concerns over China's Sci-Tech Activities and Plans: Insecurity

While positive attitudes outnumber negative ones in judgement and appreciation, the opposite is true in affect in the NYT corpus. The most prominent feature of affect in NYT reports is insecurity, which takes up 65% of the total. A close examination of all the instances shows that most of the insecurity is triggered by China's scientific and technological endeavors, its tech capabilities and ambitions, Chinese technologies and tech giants. Expressions realizing insecurity that occur 5 times and more in the NYT corpus are listed in Table 3. Table 3 shows that insecurity is largely related to concerns and worries.

Table 3. Expressions realizing insecurity in the NYT corpus

Expressions of insecurity	No.
Concern/-s/-ed/-ing	29
Worry/-ies/-ied/-ing/-isome	21
Fear/-s/-ful	10
Wary/-iness	6
Total	66

In such cases, the emoters, i.e., participants who experience the emotion, are largely related to the US and western countries. This is illustrated in the following examples.

(30) *Concern Grows in US Over China's Drive to*

Make Chips

- (31) Western companies *fear* that the Made in China policy could be used to justify government demands to hand over their latest technology...
- (32) Despite its benign name, China's Jade Rabbit rover could kindle *anxieties* among some American politicians and policy makers...

6. Discussion and Conclusion

In representing S&T in China, reports by CD feature overwhelmingly positive evaluation, highlighting the value of China's technologies and scientific research, and the competence of China/Chinese researchers in achieving sci-tech progress. By contrast, NYT contain both positive and negative attitudes in its reports. On the one hand, it emphasizes China's prowess and determination in developing high-tech industries as well as the utility of China's tech products. On the other hand, it depicts China's sci-tech endeavors as unethical and worrisome, and underlines the weakness and shortcomings of China's technologies. We propose that the different representations of S&T in China by CD and NYT might have been influenced by the newspaper's ideological and political stances and the ongoing tech war between China and the US.

As an official English-language news outlet of the Chinese government, CD is said to be the "mouthpiece for the Party in its efforts to communicate with the wider world" (Chen, 2004: 700). It can be said that CD's reports reflect the ideology and political stance of the Party vis-à-vis China's science and technology. The Communist Party of China and the Chinese government attach great importance to sci-tech innovation and see it as the primary driving force for socio-economic development. Since 2012, China has implemented an innovation-driven development strategy to improve the quality and efficiency of economy. This may be responsible for CD's overwhelming attention to the positive aspects, especially positive valuation, of China's science and technology.

Meanwhile, the Chinese government also puts much emphasis on China's technological progress and achievements, which is a consistently highlighted part in the annual government work report. According to President Xi Jinping's speech at the opening session of the

20th CPC National Congress, China has witnessed successes on multiple sci-tech fronts, made breakthroughs in some core technologies in key fields and joined the ranks of the world's innovators¹. This is reflected in our findings that CD tends to underline national scientific developments, as can be seen through by the high proportion of positive capacity in its reports. The highly positive representation by CD echoes prior research that there is a tendency for Chinese news media to foreground positivity when reporting on domestic events (Li & Zhu, 2020).

NYT's positive representation of China and China's technologies, however, is contrary to previous findings that US news coverage favors a negative tone when it comes to stories on China (Liss, 2003; Peng, 2004). We argue that such unusual positivity is partly a result of the ideological tradition in the US to construct its rival as a threat (Yuan & Fu, 2020). With a hegemonic and cold-war mentality, the US government has deemed China's advances and ambitions in strategic technologies as a long-term, fundamental challenge to American global tech leadership (Kania, 2018). China's plans to become a leader in high-tech innovation, the "Made in China 2025" initiative in particular, are believed to be substantially threatening to the global hegemony of the US (Wu, 2020). This ideology is reflected in NYT's coverage, evident in the following two examples where the link between China's positive capability and the threat it brings is manifested.

- (33) Although a latecomer by decades to space exploration, China *is quickly catching up*, experts say, and *could challenge the United States for supremacy* in artificial intelligence, quantum computing and other fields.
- (34) The Chinese *are getting good at building* these computers, and *it's a competitive issue now for U.S. industry and national security*.

Therefore, NYT's emphasis on China's competence and resolve in developing high-tech industries probably serves to reinforce the pre-established stereotypical perceptions of China's technological rise as a threat by the US in an implicit way and hence trigger worries. Yet, it must also be acknowledged that, in the past decade China does have witnessed a dramatic technology boom and become a

¹ <https://english.news.cn/20221025/8eb6f5239f984f01a2bc45b5b5db0c51/c.html>

pioneer in many key areas like manned spaceflight, artificial intelligence and 5G research, a boom which is so obvious that the American media cannot turn a blind eye to it.

Besides, the US administration also believes that certain S&T practices by China present threats to US national security interests and are inconsistent with American values. In 2021, the US Commerce Department's Bureau of Industry and Security took actions against China's technology entities in supercomputing, biotechnology, quantum computing and other areas, claiming that they are using US technology to support China's destabilizing military modernization efforts or enabling human rights abuses in Xinjiang¹. This political bias against China's S&T has permeated the coverage by NYT, which contains many expressions of negative propriety and insecurity. Its focus on negative propriety of China is in line with the typical frames of representing China by the US media, such as anti-communism (Wang, 1991), accusations of human rights abuses and stories of political suppression (Lee, 2002), as can be seen in examples (28) and (29) provided above. In addition, previous studies have shown that negative images of China have enabled the justification of an increasingly hard-line approach to China (Ooi & D'Arcangelis, 2017). Therefore, NYT's highlighting of China's negative propriety may serve to rationalize US actions to crack down on China's high-tech industries and grasp global tech hegemony amid the ongoing tech competition with China.

Facing the pressure from the US, China has taken a series of countermeasures and been forced into a tech war. In response to the Western-hyped "China tech threat", China's state actors stress that China's technological development is aimed at making lives better for the Chinese people and has not harmed the interests of other countries. "It does not target anyone, still less pose any threat," according to Mao Ning, the spokesperson for the Chinese

Foreign Ministry². Thus, by painting China's S&T as beneficial rather than threatening, reports by CD serve to illustrate China's peaceful rise in the sci-tech domain and thus provide support for China's countermeasures against US sanctions.

Another point worth paying attention is the vacillation between positive and negative attitudes in the reports by NYT. In the NYT corpus, there are 62 cases where positive evaluations are followed by contrastive negative appraisals. For instance, example (35) at first acknowledges China's achievements in developing supercomputers (positive capacity) before turning to its reliance on American chips to support these computers and China's lagging behind the United States in state-of-the-art technologies (negative capacity). In such cases, the force of the former positive attitude seems to be weakened by the negative attitude, thus potentially positioning readers to take the negative stance. This vacillation can be seen as a common narrative structure in journalism (Martin & Rose, 2003), which is also observed by Hu and Huang (2021) in their corpus-based analysis of British news coverage on the Belt and Road Initiative.

(35) Now, China has 167 systems on the list compared to 165 from the United States. China also leads a more obscure category — total processing power... Despite those achievements, Intel still provided the chips for 91 percent of the machines on the list. And China is still catching up with the United States in state-of-the-art technologies...

Unfortunately, the interaction between positive and negative attitudes can't be detected by annotating individual expressions horizontally, but requires the coder to look at the text as a whole vertically. This might be an area that calls for further exploration.

Drawing on the appraisal framework, this article has investigated how evaluative resources are employed to represent China's science and technology by CD and NYT. It also explores the underlying factors influencing the representations. Apart from confirming certain trends discovered in previous studies, such as positive representation by CD and negative representation by NYT, the article reveals that there are also positive evaluations in NYT,

¹ <https://www.commerce.gov/news/press-releases/2021/04/commerce-adds-seven-chinese-supercomputing-entities-entity-list-their>; <https://www.commerce.gov/news/press-releases/2021/06/commerce-department-adds-five-chinese-entities-entity-list>; <https://www.commerce.gov/news/press-releases/2021/07/commerce-department-adds-34-entities-entity-list-target-enablers-chinas>; <https://www.commerce.gov/news/press-releases/2021/11/commerce-lists-entities-involved-support-prc-military-quantum-computing>; <https://www.commerce.gov/news/press-releases/2021/12/commerce-acts-deter-misuse-biotechnology-other-us-technologies-peoples>

² https://www.mfa.gov.cn/eng/xwfw_665399/s2510_665401/2511_665403/202210/t20221011_10780974.html

although they mainly serve to establish China as a threat. The oscillation between positive and negative attitudes is also characteristic of the reports by NYT.

This study contributes to existing cross-cultural literature on media representation of China by focusing on the less-discussed dimension of science and technology. Methodologically, it reveals the analytical challenge of the appraisal framework in detecting the relations between positive and negative attitudes. However, as the data is limited to news articles from two newspapers, our findings may not be generalizable to other media outlets in China and the US. It is recommended that future research utilize corpus-based techniques to investigate a wider range of data for a more comprehensive understanding.

References

- Badenschier, Franziska & Holger Wormer. (2011). Issue selection in science journalism: Towards a special theory of news values for science news? In Simone Rödder, Martina Franzen, Peter Weingart (eds.), *The Sciences' Media Connection—Public Communication and its Repercussions*, 59-85. Dordrecht: Springer Netherlands.
- Boykoff, Maxwell T. & Jules M. Boykoff. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change*, 14(2), 125-136.
- Carvalho, Anabela. (2007). Ideological cultures and media discourses on scientific knowledge: re-reading news on climate change. *Public Understanding of Science*, 16(2), 223-243.
- Chen, Lily. (2004). Evaluation in media texts: A cross-cultural linguistic investigation. *Language in Society*, 33(5), 673-702.
- Chinn, Sedona, P. Sol Hart & Stuart Soroka. (2020). Politicization and polarization in climate change news content, 1985-2017. *Science Communication*, 42(1), 112-129.
- Dirikx, Astrid & Dave Gelders. (2010). Ideologies overruled? An explorative study of the link between ideology and climate change reporting in Dutch and French newspapers. *Environmental Communication*, 4(2), 190-205.
- Dutta, Soumitra, Bruno Lanvin, Lorena Rivera León & Sacha Wunsch-Vincen. (2022). Global Innovation Index 2022: What is the Future of Innovation-driven Growth? (Vol. 2000). WIPO.
- Elmer, Christina, Franziska Badenschier & Holger Wormer. (2008). Science for everybody? How the coverage of research issues in German newspapers has increased dramatically. *Journalism & Mass Communication Quarterly*, 85(4), 878-893.
- Fairclough, Norman. (1992). *Discourse and Social Change*. Cambridge: Polity Press.
- Fowler, Roger. (1991). *Language in the News*. London: Routledge.
- Guenther, Lars, Jenny Bischoff, Anna Löwe, Hanna Marzinkowski & Marcus Voigt. (2019). Scientific evidence and science journalism: Analysing the representation of (un)certainly in German print and online media. *Journalism Studies*, 20(1), 40-59.
- He, Juan & Helen Caple. (2020). Why the fruit picker smiles in an anti-corruption story: Analyzing evaluative clash and news value construction in online news discourse in China. *Discourse, Context & Media*, 35, 100387.
- Hu, Weihua & Xinjing Huang. (2021). Narrative Strategy of News in International Communication: Comparative Analysis of Report about "Belt and Road" Initiative between Chinese Press and British Press. *Foreign Language and Literature Research*, 7(04), 79-91.
- Huan, Changpeng. (2017). Evaluating news actors in Chinese hard news reporting: language patterns and social values. *Text & Talk*, 38(1), 23-45.
- Inkster, Nigel. (2019). The Huawei Affair and China's technology ambitions. *Survival*, 61(1), 105-111.
- Jungherr, Andreas, Oliver Posegga & Jisun An. (2019). Discursive power in contemporary media systems: A comparative framework. *The International Journal of Press/Politics*, 24(4), 404-425.
- Kania Elsa B. (2018). China's Threat to American Government and Private Sector Research and Innovation Leadership. Testimony before the House Permanent Select Committee on Intelligence. *Center for a New American Security*, 19.
- Kennedy, Andrew B. & Darren J. Lim. (2018). The innovation imperative: technology and US-China rivalry in the twenty-first

- century. *International Affairs*, 94(3), 553-572.
- Lee, Chin-Chuan. (2002). Established Pluralism: US Elite Media Discourse About China Policy. *Journalism Studies*, 3(3), 343-357.
- Li, Zheng, Xizhen Zhou, Samuel Jung & Jun Li. (2020). China's 40-year road to innovation. *Chinese Management Studies*, 14(2), 335-357.
- Li, Tao & Yifan Zhu. (2020). How does China appraise self and others? A corpus-based analysis of Chinese political discourse. *Discourse & Society*, 31(2), 153-171.
- Liss, Alexander. (2003). Images of China in the American print media: A survey from 2000 to 2002. *Journal of Contemporary China*, 12(35), 299-318.
- Litvinenko, Anna, Alexandra Borissova & Anna Smoliarova. (2022). Politicization of science journalism: How Russian journalists covered the covid-19 pandemic. *Journalism Studies*, 23(5-6), 687-702.
- Luke, Douglas A., Charlene A. Caburnay & Elisia L. Cohen. (2011). How much is enough? New recommendations for using constructed week sampling in newspaper content analysis of health stories. *Communication Methods and Measures*, 5(1), 76-91.
- Martin, James R. & David Rose. (2003). *Working with Discourse*. London/New York: Continuum.
- Martin, James R. & Peter R.R. White. (2005). *The Language of Evaluation: Appraisal in English*. Hampshire: Palgrave Macmillan.
- Mori, Satoru. (2019). US technological competition with China: The military, industrial and digital network dimensions. *Asia-Pacific Review*, 26(1), 77-120.
- Ooi, Su-Mei & Gwen D'arcangelis. (2017). Framing China: Discourses of othering in US news and political rhetoric. *Global Media and China*, 2(3-4), 269-283.
- Peng, Xuanwei. (2015). *Handbook for Analysing Chinese and English Appraisal Meanings*. Beijing, China: Peking University Press.
- Peng, Zengjun. (2004). Representation of China: An across time analysis of coverage in the New York Times and Los Angeles Times. *Asian Journal of Communication*, 14(1), 53-67.
- Song, Yunya & Tsan-Kuo Chang. (2012). Selecting daily newspapers for content analysis in China: A comparison of sampling methods and sample sizes. *Journalism Studies*, 13(3), 356-369.
- Stocking, S. Holly & Lisa W. Holstein. (2009). Manufacturing doubt: journalists' roles and the construction of ignorance in a scientific controversy. *Public Understanding of Science* 18(1), 23-42.
- Sun, Haiyong. (2019). The U.S.-China Tech War: Impacts and Prospects. *China Quarterly of International Strategic Studies*, 5(02), 197-212.
- Thompson, Geoffrey & Susan Hunston. (2000). Evaluation: an introduction. In Susan Hunston, Geoffrey Thompson (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*, 1-27. London: Oxford University Press.
- Van Dijk, Teun. (1995). Discourse semantics and ideology. *Discourse & Society*, 6(2), 243-289.
- Van Dijk, Teun. (2009). News, discourse, and ideology. *The Handbook of Journalism Studies*, 191-204.
- Vestergaard, Gunver Lystbaek & Kristian Hvidtfelt Nielsen. (2016). Science news in a closed and an open media market: A comparative content analysis of print and online science news in Denmark and the United Kingdom. *European Journal of Communication*, 31(6), 661-677.
- Wang, Mei-ling. (1991). Who is dominating whose ideology? New York Times reporting on China. *Asian Journal of Communication* 2(1), 51-69.
- Wu, Xiangning. (2020). Technology, power, and uncontrolled great power strategic competition between China and the United States. *China International Strategy Review* 2(1), 99-119.
- Xie, Wanying & Jiang Sheng. (2025). Digital technology races between China and the US: a critical media analysis of US media coverage of China's rise in technology and globalization. *Online Media and Global Communication*, 4(2), 216-264.
- Yu, Wenting & Liu Hui. (2020). Ideology in science communication: analyzing American news reports on AI. *Science and Society*, 10(03), 111-124.
- Yuan, Zhengqing & Qiang Fu. (2020). Narrative framing and the United States' threat construction of rivals. *The Chinese Journal of*

International Politics, 13(3), 419-453.

A Comparative Study of *Oliver Twist* and *Jane Eyre* with the “Structure of Feeling” – The Contradiction Writing of the Victorian Era

Jing Hou¹, Lingling Xu¹ & Tao Tao¹

¹ Huaiyin Institute of Technology, Huai'an, China

Correspondence: Jing Hou, Huaiyin Institute of Technology, Huai'an, China.

doi:10.63593/JLCS.2025.12.02

Abstract

Raymond Williams, an important British Marxist literary theorist, coined the term “Structure of Feeling” to analyze the shared personal feeling and experiences of people during specific historical periods. This framework reveals unstructured yet pervasive social mentality and emotional responses, offering a unique perspective for interpreting literature from social transition periods. Based on this theory, this study compares Charles Dickens’ *Oliver Twist* (1838) and Charlotte Brontë’s *Jane Eyre* (1847), two iconic Victorian-era (1837-1901) works, to explore the shared emotional tensions in 19th-century British industrialization and urbanization. The research demonstrates that both novels sharply criticize structural injustices like the Poorhouse system, class divisions, and gender oppression while compromising with mainstream ideology and values through magic solutions such as kinship redemption, unexpected inheritance of fortune, and marital order. These narrative modes reflect the dynamic tension unique to Victorian era’s transitional period between old and new values. This contradictory nature mirrors Williams’ “Structure of Feeling”, revealing the complex interplay between individual experiences and social structures, as well as the collusion between critical demands and ideological frameworks during social transformation.

Keywords: Structure of Feeling, *Oliver Twist*, *Jane Eyre*, contradiction

1. Introduction

1.1 The Social Background of Both Novels

After the First Industrial Revolution, Britain underwent profound transformations in politics, economy, and social-cultural spheres. As the first nation to complete the Industrial Revolution, it became the “world’s factory” and rapidly amassed vast wealth. The bourgeoisie emerged politically and consolidated their power after several Parliament reforms. By the mid-19th

century, the constitutional monarchy in Britain entered its golden age, along with a period of intense industrialization and urbanization. While the economy grew exponentially, the moral order of traditional agrarian society clashed violently with the values of emerging capitalist society, resulting in severe class divisions and widening wealth gaps. Thus, social contradictions became increasingly acute. “On one side stood aristocratic manor houses, while on the other, farmers’ shabby cottages; on

one side, factory owners enjoyed comfortable lives, while on the other, unemployed workers struggled desperately for survival" (Chengdan, 2002). What's worse, the British Parliament enacted the Poor Law Amendment Act of 1834 under pressures from large factory owners, stipulating that "all relief work must be conducted in workhouses (indoor relief), where living standards must be lower than those of the poorest outside the houses" (Chengdan, 2002). As punishment, recipients of relief lost their voting rights. This legislation embodied Adam Smith's political economy philosophy: "The rich should remain rich and the poor remain poor, and the government do not bear social relief responsibilities towards the poor" (Hongtao, 2022). Working conditions in workhouses resembled hell, making poverty a persistent threat. The enactment of the New Poor Law further aggravated the already dire living conditions of the workers. As a consequence, more workers began to actively engage in labor movement in order to survive and protect their own rights, kicking off the vigorous Charter Movement. Literary works are rooted in life and, to some extent, mirror social realities. Charles Dickens' *Oliver Twist* and Charlotte Brontë's *Jane Eyre*, both 19th-century realist works, target different social contradictions due to respective gender and life experiences. Yet, as they were written under the same social backdrop, they inevitably share similarities and both embody the same "Structure of Feeling", a cultural theory created by British Marxist cultural critic Raymond Williams to examine literary works.

1.2 Theory of "Structure of Feeling"

The term "Structure of Feeling", coined by Raymond Williams (1921-1988) in his book *Film: A Short Introduction* (1954) co-authored with William Michael, originally refers to the collective live experience of people in a given era. The connotation of this term underwent significant evolution in his successive works. In *Marxism and Literature* (1977), Williams expanded this concept by incorporating Gramsci's theory of cultural hegemony, transforming it into a critique of capitalist hegemony. He conceptualized Structure of Feeling as a mechanism for resisting dominant cultural hegemony (Yifan, 2017). Williams categorized social-cultural elements during critical social transformation period into three types: the dominant, emergent, and residual—corresponding to bourgeois,

working-class, and aristocratic cultures respectively in the 1840s. These cultural forces interact with each other, resulting in constant conflicts and clashes of values, which were reflected by the mindset and immediate experiences of people at that historical time. The Structure of Feeling, consisting of the universal experiences and perceptions of people, particularly the conflicts between people's actual experiences and the dominant moral values, is the preemergent stage of the fight against cultural hegemony.

Williams posits that the prevalent Structure of Feeling in popular culture "persist in nearly all the novels which are referred to as literary works." To elaborate his theory, Williams analyzed 19th-century British industrial novels (including Elizabeth Cleghorn Gaskell's *Mary Barton* and Charles Dickens' *The Great Expectations*) in *The Long Revolution* (1961), identifying a distinctive structure of characteristic of the period (Raymond, 2019), which is the tension between middle-class (bourgeois) ideology and novel writers' lived experiences and pursuit of life. Williams introduced the concept of "social character" to describe dominant social groups' moral values that embody cultural ideals, emphasizing how these characteristics interact. His analysis of mid-19th-century British society reveals: work as a value system where personal effort directly determines success; poverty attributed to personal shortcomings, necessitating punitive measures like the New Poor Law to motivate hard labor; hardship as a noble virtue cultivating humility and courage; and virtues like frugality, moderation, and piety being prioritized, with family as the central institution and marriage's sanctity inviolable (Raymond, 2019). To sum up, this framework, based on bourgeois work ethics and family-oriented values, promotes diligence and personal accountability while attributing economic achievement to moral superiority.

1.3 Research Framework and Significance

So far, there have been numerous studies on both *Oliver Twist* and *Jane Eyre* around the world. However, most of them focus on the class or gender issues of individual works, and there is still a lack of systematic comparisons on how they jointly reflect the contradictions in the structure of feeling in social transformation. This study adopts a method combining close textual reading and comparative analysis, and takes

Williams' theory of Structure of Feeling as the framework to interpret from two dimensions: "Institutional contradictions and personal redemption" and "magical solutions". By selecting the survival predicaments and endings of orphan characters in two novels as core samples and combining them with social backgrounds such as the Poor Law Amendment Act of the Victorian era, the reform of women status and property rights, the interactive mechanism between literary texts and social characteristics is revealed. By comparing two novels, this study aims to explore the explanatory power of the theory of emotional structure in analyzing literature during the period of social transformation, explain how the two works expose the same conflict between institutional oppression and individual moral redemption in the Victorian era through orphan narratives, and critically reflect on the masking and mediating functions of symbolic solutions in literary works on real-life contradictions. At the theoretical level, this study helps to expand the literary interpretation boundaries of Williams' theory of the Structure of Feeling. At the social and cultural level, it provides historical references for the study of cultural contradictions during the social transformation period (such as the predicament of class mobility and the reconstruction of gender identity) by decoding the collective anxiety and strategic ideological games in 19th-century Britain through text analysis.

2. Institutional Criticism and Personal Redemption

Guided by Adam Smith and Ricardo's "free economic theory" and Bentham's utilitarianism, Britain was the first to embark on the path of free capitalism and became the first industrialized country. In the 19th century, Britain witnessed tremendous changes in its social economy, politics and culture, and at the same time, intense social conflicts emerged, especially the large gap between the rich and the poor. The working class had long been exploited and oppressed, and was constantly at risk of unemployment due to the use of machines. Meanwhile, with the expansion of middle class, the significance of family was consolidated. During that period, confronted with discrimination and suppression, women were strictly confined at home by the role of "family angels" and had an obvious unfavorable social status. Their success depends on their

contributions to the family. Dickens and Charlotte lived in such an era. Their respective life experiences and genders enabled them to focus on different social contradictions. Both *Oliver Twist* and *Jane Eyre* expose the injustice of social systems from the perspective of orphans, yet they pin their redemption plans on individual morality or good luck.

2.1 *Oliver Twist: Hypocrisy of the Workhouse and Bloodline Redemption*

Dickens, born into the middle class, was forced to drop out of school and work as a child laborer in a blacking factory to help support the family when he was 12 years old because his father was imprisoned for debt. For several months, Dickens worked with children from poor families and experienced the hardships of the lower class, which directly inspired his writing of "David Copperfield". The darkness of society and the injustice of the system also frequently popped up as the major theme of his novels. *Oliver Twist* was born in a poorhouse, suffered severe abuse and went hungry all day long, rendering it a miracle that he should have survived and grown up. At the age of eight, he was sent to an undertaker to work as an apprentice because he was disowned by the house for a "crime" of asking for "a little more porridge" (Charles, 2004). Later, unable to bear the bitter treatment of the undertaker's wife, he fled to London and was lured into a theft gang, but he remained true to his kind nature. Finally, with the help of kind-hearted people like Mr. Brownlow, he regained his status as the heir of a rich man. The most significant change in the New Poor Law of 1834 was the prohibition of providing assistance outside of poorhouses to those who were capable of working, and those receiving assistance must participate in the work arranged by the diocese. The aim is to make the poor understand that they must rely on themselves rather than on the government and society to rid poverty, thus shifting the root cause of the poverty problem from the imperfect economic structure to the laziness and bad habits of the poor without any concern for the living conditions of the poor (Hongtao, 2012). The children in the poorhouse had been in a state of hunger for years. Most of them fell ill due to cold and hunger, or die accidentally because of negligence. More often than not, the dead children's stomachs were found to empty after dissection. The children were starving, but if they dare complain, they would be harshly

punished (Charles, 2004). The relief for each child was 6.5 pence a week, which was limited, but Mrs. Mann appropriated the greater part of the weekly stipend to her own use. Oliver's eight birthday found him a pale thin child, diminutive in stature and decisively small in circumference. At eight years old, he started working at six every morning but earned scarce food. When the undertaker's wife saw him, she kept complaining that he was too young, "I see no saving in parish children; for they always cost more to keep" (Charles, 2004). Through the life experiences of the protagonist Oliver, Dickens sharply criticized the cruelty and hypocrisy of the poorhouse system.

In *Oliver Twist*, Dickens not only vividly depicted the living conditions of children in the poorhouse to expose the greedy and hypocritical faces of the poorhouse administrators (Mrs. Mann, Mr. Bumble), but also criticized the corruption in the judicial system (Mr. Fang the police magistrate). The novel delves into the darkness and hypocrisy of the bourgeoisie by depicting the overall survival predicament of children at the bottom of society during the process of industrialization. Oliver's salvation does comply with the middle-class values about the poor: outstanding people will eventually stand out (Yifan, 2017). Indeed, it was Oliver's integrity and kind nature that bond him with Mr. Brownlow, who in turn help him regain his middle-class status as the son of a rich man. However, it also reflects that the poor was not able to extricate himself from darkness, and the poor remains poor not because they do not strive hard enough or born with shortcomings but because of the unjust social structure. Consequently, the solution for the poor orphan arranged by Dickens deviated from the track of structural criticism by attributing Oliver's ultimate redemption to the compassion of the charity, his flawless morality and his accidental middle-class background (his blood relationship with Brownlow). This narrative strategy simplifies structural oppression into an opposition between good and evil, suggesting that the continuation of traditional bloodline order and moral conscience is the solution to social problems. At the same time, it also indicates the realistic predicament that individuals cannot break through class barriers through their own strength. This contradiction is precisely a typical manifestation of the emotional structure — the writer is aware of the

existence of social problems, but is unable to transcend the limitations of his time to propose more practical solutions, and can only resort to traditional moral resources and accidental luck. This compromising solution has, to a certain extent, weakened the critical nature of the novel, shifting the focus that should have been directed at the social system to the individual moral level.

2.2 *Jane Eyre: Hypocrisy in Charity Schools, Gender Oppression and Redemption*

Charlotte Bronte was born in a middle-class parsonage family at the end of the flamboyant Romanticism and the beginning of the tender Victorian culture. On the one hand, the factory system of the Industrial Revolution separated work from family, making the family a warm place free from external pressure and reinforcing its importance. On the other hand, after her marriage, Queen Victoria focused on family affairs and indifferent to politics, forming a pattern where women take care of the household and men work outside. Due to its alignment with the interests of the bourgeoisie and male chauvinism, the pattern was highly prized and became the unique Victorian value system that was popular among all social strata (Lanxin, 2017). Many elegant, pure and gentle "family angels" also emerged in literary works. Unlike the lower-class working women who were forced to go out to work, middle-class women were mostly confined to the family and are accessories without social functions who relied on their husbands (Chengdan, 272). Middle-class wives going out to work were likely regarded as losing dignity. Intellectual women like the Bronte sisters could only choose semi-servant work with an awkward social status as governesses like *Jane Eyre*, because the job did require public exposure. As petty bourgeoisie, the Bronte sisters had a very narrow path to personal success, nor career ladder to climb. Moreover, they did not have generous dowries and were unable to change their impoverished family situation through an upward marriage. Due to her solitary and proud nature, Charlotte found it hard to adapt to the insulting job as a governess. At first, she attempted to achieve a class leap by founding a boarding school, but failed. Later, she turned to writing, which were challenges to gender oppression and institutional injustice. In her work, *Jane Eyre* is a female brimming with romanticism. She is full of personal emotions

and passion and dreams of gender equality in love and marriage. The realistic novel criticized the class gap and gender oppression but still could not break off from the background of the time.

Jane Eyre, an orphan girl with no wealth, has no beauty nor status. However, with her independent will, spiritual strength and uncompromising resistance, she eventually becomes the mistress of Thornfield Hall. This growth trajectory is not only a reversal of an individual's fate, but also profoundly reveals how structural injustice permeated family, education, marriage and property systems in the Victorian era. Moreover, Jane's success also highlights the universal contradictions in the emotional structure of that era. The Lowood Charity School where Jane Eyre lived for eight years was strictly managed and adhered to the principle of "punishing the body to save the soul" (Charlotte, 2015). Students were short of food and clothing and lived in poor conditions. Mr. Brocklehurst, who was full of religious preaches and always giving moral talk, advocated "enduring hardship, patience and self-restraint". When students are hungry, he claimed to "endure hardship for a while to cultivate their spirit" (Charlotte, 2015). Under the pretext of saving souls, he ordered Jane Eyre to cut her hair and humiliated her in public, yet his wife and daughter were dressed in fancy clothes. His act of cutting off supplies and treating students harshly was not exposed until typhus broke out at the school and many students died. His pseudo-charitable nature and religious hypocrisy were exposed to the fullest. However, the final solution was that several wealthy and benevolent people donated money to rebuild the school. The Mr. Brocklehurst still performed his duties with a few broad-minded gentlemen assisting from the side, thus Charlotte's criticism of the institutional system was also incomplete. As an educated woman, although she had a certain degree of foresight and criticism, her solutions did not touch upon the essence of institutional flaws. She still placed her hopes on beneficence of some rich people and, to some extent, replicated the values of the middle class.

Jane posted a job advertisement in the newspaper, and was employed as a governess, the only job she could find in spite of her talent. At Thornfield Hall, Jane and Rochester fell deeply in love crossing the gap of status and

money. Later, Jane left because of the disclosure of his "mad wife". On the one hand, the author portrays Jane as a woman with a strong sense of self and an independent spirit. She declared, "We are equal", which not only challenged the values of the nobility and the middle class in the early Victorian era but also criticized the Victorian gender norms. On the other hand, Jane's resolute departure upon discovering that Rochester already had a wife was a defense of the middle class value towards marriage and family. "Marriage is absolutely sacred, and adultery and infidelity are unforgivable" (Yifan, 2017). Personal desires and traditional concepts are constantly in conflict and compromise with each other, and contradictory emotional structures are everywhere. When Jane receives unexpected property and returns to Thornfield in a Gothic plot at Rochester's call, the initially powerful male character is injured and blind. In marriage, the strength of men and women is swapped. Such a plot arrangement undoubtedly sends a certain signal challenging the patriarchal culture (Gengxin, 17). However, this romantic and radical gender and class stance relied on the concept of family in Victorian ideology to exist, while the romantic exterior was filled with a rebellious spirit against tradition. After marriage, Jane became a "family angel" and smoothly achieved a social class promotion in her life. This arrangement exposed the limitations of Charlotte as the son of her time: she could envision the spiritual independence of women, but found it difficult to completely break away from traditional economic structures and gender roles. As scholar Su Gengxin pointed out in the introduction of "British Fiction and Romanticism" when analyzing the gender roles in Victorian novels: In romantic novels, the more common reason for the dramatic change or transformation of characters' roles and status is that during a transitional period when old values are about to fade away and new values have yet to take shape, the author himself lacks stable values and ideological positions, and is indecisive in the face of two competing orders and systems. This contradictory mentality precisely constitutes the core features of the Structure of Feeling during this period described by Raymond Williams. As Williams put it, art is a living culture. On the one hand, authors challenge authority in pursuit of their ideals; on the other hand, they consciously or unconsciously uphold the values of the middle

class and strengthen the dominant culture. This is the most genuine interaction between literary texts and social reality.

2.3 Shared Contradictions in the Structure of Feeling

Both works take orphans, a marginalized group in society, as the main characters, exposing the flaws of the social system through their perspectives. However, they both coincidentally transform institutional contradictions into individual moral narratives, revealing the tension of criticism and compromise in the Structure of Feeling. This common narrative strategy is not accidental but a typical manifestation of the Structure of Feeling during the social transition period of the Victorian era when intellectuals were acutely aware of social problems, but were constrained by the mainstream ideological framework when it came to how to thoroughly solve these problems. Williams once pointed out that the characteristic Structure of Feeling lies in that it captures “the special and vivid results brought about by all factors in a general organization” (Raymond, 2019), which is a reflection of the constant conflict among the dominant culture, the emergent culture and residual culture in people’s mentality, rather than a solidified ideological stance. In this sense, both *Oliver Twist* and *Jane Eyre* have successfully captured the emotional contradictions brewing in Victorian British society — dissatisfaction with the traditional order and the yearning for new social relations, while at the same time being unconsciously constrained by the ideological limitations of the times. This contradiction precisely makes the two works appropriate for studying the Structure of Feeling of the Victorian era, as they not only record the social issues of that time, but also document people’s mental states and emotional responses when facing these problems.

3. The Ideological Function of Magical Solutions

Williams pointed out that literary works often compromise contradictions through magic schemes to meet the demands of mainstream ideology. “Magic is indeed necessary; it can be used to postpone the conflict between ethics and experience” (Raymond, 2019). When reality is too cruel and the protagonist is in despair, they often resort to the standard model to solve the urgent problem. Among the various magical solutions to the contradiction between personal

feelings and social characteristics, the most common arrangement is “an unexpected gift and the empire” (Raymond, 2019). Despite the increasingly acute social conflicts and the constant clashes between personal ideals and aspirations and popular values, middle-class values still dominate the field of cultural production, and fiction still serve the dominant culture. Both *Oliver Twist* and *Jane Eyre* offer emotional comfort in the form of moral fables, weakening the radical criticism of social issues and bridging the gap between reality and ideals through magic solutions. This narrative strategy not only reflects the writer’s personal choices, but also embodies the deep-seated anxiety over the collapse of social order within the emotional structure of The Times.

3.1 The Magical Plan of *Oliver Twist*

In *Oliver Twist*, Oliver was unfortunately drawn into a den of thieves. Facing Fagin’s threats and inducement, he remained as kind-hearted as a child from a wealthy family. It was his kindness that enabled him to meet Mr. Brownlow and receive help to escape the abyss. Dickens simplified the complex social collisions into the issue of moral purification by eliminating the villains (Fagin being hanged and Sykes falling to his death) and restoring to family order (Oliver’s return to his middle class family). This black-and-white moral view undoubtedly satisfies the psychological needs of middle-class readers for order and justice while concealing the structural roots of poverty (such as the flaws in the New Poor law). When Oliver was finally revealed to be a middle-class heir and recovered his legitimate social position, Dickens strengthened the legitimacy of the class essence theory, suggesting that a person’s social status is ultimately determined by their birth.

This narrative arrangement is no accident; it closely aligns with the ideology of the middle class in the Victorian era when social discourse often attributed poverty to individual moral flaws rather than institutional injustice. Although Dickens felt deep sympathy for the poor, he ultimately failed to completely break free from the shackles of mainstream thinking. Through the magical plan of bloodline redemption, the novel ingeniously avoids fundamental questioning of the capitalist system and transforms possible social criticism into moral exhortation. This strategy to some extent weakened the critical nature of the work, but perhaps it was precisely his compromise that

enabled *Oliver Twist* to gain widespread resonance at that time.

3.2 The Magical Plan and Imperial Imagination of *Jane Eyre*

On her path to becoming the wife of a noble estate owner, Jane Eyre's main obstacles were money, status and morality. In the 1840s, the concept of family status was deeply rooted in people's minds, which indicated that the influence of the residual noble culture was still profound. However, the mainstream middle-class concept began to embrace money determinism — tying money to success and status. Jane's financial problem was the first to be resolved. The author arranged for Mr. Eyre, who was doing business in the Madeira Islands, to pass away and leave all his property (a total of 20,000 pounds) to his niece Jane, who became rich overnight from poverty. Williams believed that the empire was a broad and circuitous route, "often an unexpected source of wealth" (Raymond, 2019). In Charlotte's description, opportunities were abound overseas, colonies were full of wealth, and the depiction of empires also appeared repeatedly in her works.

Jane's wealth still lagged behind that of Rochester to some extent, so the author arranged for a fire at Thornfield Hall. This magical plan achieved three goals at once. Not only did it reduce Rochester's wealth and cause him to lose his eyesight, but it also cleared another obstacle to their marriage — Rochester's mad wife Bertha died in the fire, resulting in a happy ending for all. Bertha was the daughter of Mason, a plantation owner and merchant in the West Indies. To enhance her family's status, she married Rochester, the second son of a nobleman who lacked property, with a dowry of 30,000, which solved Rochester's financial problem as a second son to his noble father. However, the author arranged for Bertha to dehumanize and set herself on fire to fulfill the love story of Jane and Rochester. Spivak believed that "when female authors of published works come from the ruling 'culture', they sometimes, like male authors, have a tendency to create a vague 'other' (often female)" (Spivak, 2014). As a Creole from the West Indies, Bertha embodies the colonial other, female desires and madness. Bertha's death symbolized the suppression of the colonial other and female desires, thereby achieving the self-redemption of Jane Eyre, suggesting that female independence comes at the cost of excluding "irrational" others. This

magical solution reflects the inner predicament of middle-class women in the Victorian era: they longed for equality but were afraid of the consequences of completely breaking away from traditional gender roles. By externalizing the inner conflicts with Bertha as the other and achieving psychological balance by eliminating her, Bronte actually found a narrative path that expresses women's demands without completely subverting social order.

Both the beautiful imagination of colonial wealth and opportunities and the narration of other colonial 'barbarians' in literature directly or indirectly promoted the construction of the British Empire.

3.3 The Historical Context of Magical Solutions

Both works exchange the sacrifices of marginalized groups (such as lower-class criminals and colonial women) for the redemption of the protagonists, exposing their common structure of feeling, which is the coexistence of radicalism and conservatism. This narrative mode needs to be examined within the specific historical context of the Victorian era when British society was in a crucial period of transition from tradition to modernity, and various social conflicts intensified. For instance, workers' movements such as the Charter Movement (which started in 1839) occurred one after another, causing widespread concern in British society about the situation at that time. Against this backdrop, cultural works often assume the function of a safety valve for social anxiety. From the perspective of structure of feeling, these magic solutions are not only the narrative choices of individual writers, but also projections of the collective psychology of the time. They reflect the contradictory mindset of Victorian British society in the face of the pains of modernization: both longing for change and fearing disorder, criticizing the current situation while also being attached to tradition. Williams emphasized that the value of Structure of Feeling lies in its ability to capture common experiences that have not yet been explicitly expressed but are already taking shape. In this sense, the symbolic solutions of *Oliver Twist* and *Jane Eyre* precisely document the complex emotional responses of people in the Victorian era when facing social transformation — a delicate state between hope and fear, criticism and compromise.

4. Conclusion

From the perspective of the theory of emotional structure, when reinterpreting *Oliver Twist* and *Jane Eyre*, we can find that at the level of social system criticism, both novels keenly capture the social problems of the Victorian era (class oppression, gender injustice), but both rely on individual morality and accidental opportunities to solve the structural contradictions. They jointly reflect the contradictory mentality of individuals during the social transformation period — both longing for change and fearing disorder, demonstrating the limitations of the ideology of the time. As carriers of the structure of feeling of the Victorian era, the two works reveal the oppressive nature of industrialization and patriarchy but maintain the traditional order through moralized narratives. This contradiction is precisely the embodiment of what Williams called the outcome of “fighting between old and new cultures”. The conflict between individual experience and social structure, the collusion between critical demands and ideology jointly constitute the complex aspects of literature during the transitional period. At the magic solution level, both works have reconciled real contradictions with fulfilled ideological functions by eliminating marginal others (lower-class criminals, colonial women) and maintaining the core order (family, marriage).

The limitation of this study lies in that it mainly focuses on two novels and have not fully examined the depiction of the working class in other important Victorian texts such as Mrs. Gaskell’s *Mary Barton*. Future research could further explore how other Victorian texts (such as *Mary Barton* and *Mill on the Floss*) responded to the dynamics of structure of feeling in different forms, thereby deepening the understanding of the cultural transformation in 19th-century Britain. Just as Williams reminds us, the structure of feeling is always in a process of flow and formation. Only through multidimensional and multi-textual examination can we have a more comprehensive grasp of the spiritual outlook of an era.

Funds Project

This paper was supported by the project A study on the “Structure of Feeling” of Charlotte Bronte’s novels (The general Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province in

2021 commissioned by Jiangsu Provincial Department of Education and Jiangsu Social Science Funding). Project number—2021SJA1808.

References

- Charles Dickens. (2004). *Oliver Twist* (pp. 18-50). Printer Industria Grafica, Barcelona.
- Chengdan Chan, Jieming Xu. (2002) *A General History of Britain* (pp. 222-270). Shanghai Social Sciences Press.
- Gayatri Chakravorty Spivak. (2014). *A Critique of Postcolonial Reason: Towards a History of the Vanishing Present* (pp. 117). Translated by Beiwen Yan. Yilin Press, Nanjing.
- Gengxin Su. (2017). *British Fiction and Romanticism* (pp. 17). Peking University Press.
- Lanxin Wang. (2017). *An Analysis of Victorian Values in the 19th Century*. The New West (Theoretical Version), (04), 98+108.
- Raymond Williams. (2019). *The Long Revolution* (pp. 65-83). Translated by Ni Wei. Foreign Language Teaching and Research Press, Shanghai.
- Yifan Zhao, Zhongzai Zhan. (2017). *Key words of Western Literary Theories* (pp. 433-441). Foreign Language Teaching and Research Press, Shanghai.

A Conversation Analysis of Turn-Final *haobuhao* in Chinese Parent-Child Interaction

Shuyu Pan¹

¹ Ocean University of China, Shandong, China

Correspondence: Shuyu Pan, Ocean University of China, Shandong, China.

doi:10.63593/JLCS.2025.12.03

Abstract

Haobuhao (“好不好”) is a recurrent linguistic practice in Mandarin conversation and performs different interactional functions depending on where it is positioned within a turn. This study examines occurrences of turn-final *haobuhao* in parental talk in naturalistic parent-child interactions, drawing on a corpus of video-recorded data and using the analytic framework of Conversation Analysis (CA). The analysis shows that when *haobuhao* is deployed in responding position, it frequently implements other-initiated repair targeting the child's conduct. In these cases, parents use *haobuhao* to mark the child's behavior as inappropriate, to frame it as a departure from social norms, and to assert their epistemic authority. When used in initial position, *haobuhao* commonly accompanies directive formats. In directives delivered through commands or complaints, *haobuhao* strengthens an accusatory stance and highlights the problematic nature of the child's behavior. When accompanying request-based directives, however, *haobuhao* mitigates the directive force and downshifts parents' deontic authority. By revealing how *haobuhao* contributes to the management of epistemic and deontic relations in interaction, this study demonstrates the intricate ways in which interactional practices participate in the construction of social norms and parental authority. The findings contribute to a more nuanced understanding of Mandarin parent-child interaction and expand CA research on epistemics in family and institutional settings.

Keywords: conversation analysis, *haobuhao*, parent-child interaction, sequential position, epistemics

1. Introduction

Haobuhao is a recurrent interactional practice in Mandarin that appears in a range of sequential environments. While it may function as part of a clause, it frequently occurs turn-finally as an interrogative appended to a complete declarative or imperative. In such cases, the interactional role of *haobuhao* cannot be understood solely through its grammatical interrogative form. Rather, its function emerges from the social action implemented by the

surrounding turn-constructive units (TCUs) (Sacks et al., 1974), as shown in Examples 1 and 2. Here, *haobuhao* is mobilized not to solicit information, but to accomplish additional interactional work.

Extract 1: [去北京: 00:20:38-00:20:51]

- 01 妈妈: 你知道北京>有什么<?你还去玩?
02 女儿: 北京有<游乐园>,<游乐场>.
03 [还有()]

04 妈妈: [游乐园我们这]也有,好不好?

05 (1.5)去北京要(.)去看升国旗.

Extract 2:[去吃水饺了: 00:00:24-00:00:34]

01 妈妈: hehe.我们中午在(0.6)饭店去吃水饺了.

02 女儿: 哦.

03 (2.4)

04 妈妈: 你嚼嚼再咽,好不好?=-

05 女儿: =嗯:↑,太好吃了.

Parent-child interaction constitutes a form of institutional talk in which epistemic and deontic asymmetries are inherent. Through everyday interactions, parents construct their identities, enforce social norms, and socialize their children. Understanding how specific linguistic practices contribute to the negotiation of epistemic and deontic relations is therefore essential to understanding the organization of family interaction.

This study employs CA to investigate the interactional functions of turn-final *haobuhao* in naturally occurring Mandarin parent-child interaction. Drawing on concepts of epistemics (Heritage, 2012a, 2012b, 2018) and deontic rights (Stevanovic, 2012; Peräkylä & Stevanovic, 2013), the analysis examines how parents use *haobuhao* to display epistemic expectations, enact authority, and manage children's conduct. In doing so, the study contributes to broader discussions about the relationship among social norms, linguistic practices, and social actions in Mandarin family interaction.

2. Literature Review

Research on *haobuhao* in Mandarin has largely been conducted within a pragmatic framework. Prior studies typically examine *haobuhao* as a postposed negative marker or as part of the *X-bu-X* interrogative pattern, identifying its discourse- and pragmatics-related functions (Gao, 2009; Yu & Yao, 2009; Yan, 2015; Wang, 2017). Drawing largely on literary texts, film scripts, and constructed examples, these studies identify a range of discourse functions—such as initiating a turn, facilitating turn exchange, or introducing a third party—and discuss pragmatic functions related to modality, politeness, and intersubjectivity, including seeking permission or agreement, requesting advice, and establishing interactional alignment. While such work offers insights into certain uses of *haobuhao*, it is limited in capturing the full

complexity of *haobuhao* as an interactional practice. A more adequate understanding requires examining *haobuhao* within actual, naturally occurring conversational sequences, where its function is shaped by its sequential position and by the social actions implemented by the surrounding turns.

Motivated by this gap, the present study employs CA to examine turn-final *haobuhao* in video-recorded Mandarin parent-child interaction. Focusing exclusively on *haobuhao* used as a turn-final interrogative appended to a prior TCU, the study describes its sequential distribution and identifies the social actions it accomplishes across different environments. By further considering epistemic and social norms that parents orient to, the study aims to elucidate how this practice operates as a conversational routine in the management of knowledge, authority, and social behavior in parent-child interaction.

3. Research Methodology and Data Collection

This study adopts the methodological framework of CA, a branch of microsociology developed by Harvey Sacks, Emanuel Schegloff, and Gail Jefferson. CA draws on Goffman's structural conception of interaction order and Garfinkel's notion of shared common-sense reasoning, and its distinctive contribution lies in the creative use of audio and video recordings and the development of a specialized transcription system, which together allow for detailed observation and analysis of social action in talk. CA is a naturalistic, descriptive qualitative approach. Researchers aim to uncover and describe how participants accomplish social actions in interaction, focusing on what is observable in the sequential organization of turns rather than imposing subjective interpretations. Its naturalistic character is reflected both in the collection of data—recordings of spontaneous interaction without experimental manipulation—and in analysis, which examines how social actions are produced and normatively organized through participants' turn construction.

The dataset consists of video-recorded, naturally occurring parent-child interactions collected in everyday home settings. All recordings capture spontaneous activities such as mealtime conversations, play sessions, and routine household interactions. The participants are Mandarin-speaking parents and their preschool-

or school-aged children, all capable of producing complete and coherent utterances. The data are transcribed using CA transcription conventions (Sacks, Schegloff & Jefferson 1974), with attention to turn-taking, sequence organization, and action formation. The total duration of recordings is approximately 190 minutes, within which 24 instances of turn-final *haobuhao* functioning as appended interrogatives are identified.

4. Data Analysis

The interactional functions of *haobuhao* depend on both its sequential position and the social action implemented by other TCUs within the same turn. A close, data-driven examination of the corpus shows that when *haobuhao* occurs in a responding position, parents typically use it while performing repair. In contrast, when *haobuhao* appears in a sequence-initial position, it is commonly embedded within turns implementing directive actions, such as issuing commands, complaints, or requests. The following sections examine how turn-final *haobuhao* displays epistemic norms and deontic orientations in these different sequential environments, and what interactional functions it accomplishes.

4.1 *Haobuhao* in Responding Position

The data show that turn-final *haobuhao* in responding position is frequently used when parents initiate or implement repair. In these cases, the appended interrogative *haobuhao* serves as a conversational practice through which parents highlight the inappropriateness of the child's prior behavior and maintain their epistemic authority.

In everyday social life, and especially in talk-in-interaction, participants routinely encounter troubles of hearing, speaking, or understanding. When mutual understanding is not achieved, participants deploy conversational repair to address and resolve the trouble (Schegloff et al., 1977). Because repair interrupts the progressivity of talk, it takes priority over other ongoing social actions. Only after the trouble is resolved and intersubjectivity restored can interaction continue and other social actions be carried out. Thus, repair holds a structurally privileged status within the organization of conversation.

In the context of parent-child interaction, turn-final *haobuhao* in repair sequences is not simply a yes/no interrogative. Rather, it

functions as a parental practice for re-establishing normative expectations, framing the child's prior conduct as problematic, and reinstating the parent's epistemic authority. The detailed sequential analysis below demonstrates how *haobuhao* participates in the organization of repair and in the management of knowledge and responsibility between parents and children.

Extract 3. [亲子饭间对话: 00:00:45-00:01:08]

- 23 女儿: 好凉啊.
24 (0.4)
25 妈妈: 凉?
26 女儿: 嗯.
27 妈妈: 没在冰箱里放着,在外面放着呢.
28 (0.7)
29 女儿: [()]
30 妈妈: [>那应该<]是热的,好不好?
31 (0.4)
32 女儿: 那不是热的,关键是.
33 (2.2) ((妈妈品尝薯条))
34 妈妈: >一点儿也不凉.< (0.6).h ¥哈哈哈哈哈
¥[.h
35 女 儿 :
[>蘸上<番茄酱就凉了].
36 妈妈: 哦,番茄酱凉.
37 (.)
38 女儿: °he°=
39 妈妈: =>番茄酱在冰箱里呢,<番茄-番茄
酱凉,(0.8)薯条不凉.
40 女儿: 对.

Excerpt 3 occurs after the daughter requests ketchup, the mother retrieves it from the refrigerator, and the daughter eats a French fry with ketchup. In line 23, the daughter produces an assessment of the food's temperature ("It's so cold"). Following a 0.4-second silence, the mother initiates repair in line 25 by partially repeating the trouble source ("Cold?"). This partial-repeat repair initiation (Schegloff et al., 1977) indicates that the mother's trouble is not one of hearing but of understanding: she displays difficulty comprehending the grounds on which the daughter has produced this assessment, or treats the assessment as misaligned with the actual state of affairs.

After the daughter confirms the trouble source in line 26, the mother continues in line 27 by

offering an explanation: “It wasn’t in the fridge; it was left outside.” The design of this turn shows that the mother treats the daughter’s assessment as inconsistent with what she knows to be the case and provides information about storage conditions to recalibrate the daughter’s understanding. Here, the mother displays greater epistemic access to the temperature and storage status of the food, positioning herself as the more knowledgeable participant in this domain.

From an epistemic perspective, the mother’s turns in lines 27 and 30 reflect orientations to epistemic access and epistemic primacy. According to epistemic norms, speakers should avoid making claims in domains where they lack sufficient access (Heritage & Raymond, 2005). Epistemic primacy further holds that participants with superior knowledge have greater rights to make evaluations concerning that domain (Stivers et al., 2011). In this sequence, the mother displays her knowledgeability regarding the fries’ storage and temperature; thus, when the daughter’s assessment conflicts with her epistemic territory, she assumes epistemic responsibility for resolving the discrepancy. Her turn in line 30 implements repair to restore intersubjectivity.

Importantly, the appended interrogative *haobuhao* in line 30 follows a syntactically complete TCU and is not necessary for the implementation of repair. Thus, *haobuhao* here does not function as a request for confirmation. Prosodically, *haobuhao* is latched to the preceding TCU without a gap and is produced with stress, suggesting its role as a discourse marker rather than an information-seeking interrogative. In this position, *haobuhao* serves to draw the child’s attention and to strengthen the mother’s stance. It highlights the epistemic asymmetry between the participants and reinforces the mother’s epistemic authority. More specifically, turn-final *haobuhao* here works as a conversational practice for marking the child’s assessment (the repairable) as inappropriate and for foregrounding the parent’s entitlement to correct it.

Extract 4. [一起剪窗花: 00:00:22-00:01:53]

- 20 妈妈: 你想剪个什么样的?
21 (0.6)
22 女儿: 嗯:;(0.7)随便选什么样的都行:.
23 (.)>但是<(0.8)°先对折,这样对折.°

- 24 妈妈: 对折:;,
25 (2.6) ((折东西的声音))
26 再对折:;.
27 (1.4) ((折东西的声音))
28 女儿: 对:;(0.2)°很好.°
29 (0.5) ((折东西的声音))
30 妈妈: 很好?是我<教你>,好不好?
31 hehe[°he°
32 女儿: [h 这是<老师>教我的.
33 (0.2)
34 妈妈: 再对折:;=
35 女儿: =哎,不不不.
36 妈妈: 我们>再对折<,我再对折,
37 (0.4)你不对折了吗?
38 女儿: 对呀.

Example 4 occurs during a mother-daughter activity in which they are folding paper in preparation for making window cutouts. In line 20, the mother initiates an open-ended question (“What kind do you want to make?”). After a 0.6-second silence, the daughter, in lines 22-23, not only provides a type-selecting response but also proceeds to deliver step-by-step instructions (“Fold it first, fold it like this”). In this sequential position, the daughter presents detailed procedural directives, thereby positioning herself as the party with greater epistemic access and epistemic authority, and consequently proposing a reconfiguration of the ongoing activity framework.

In lines 24-27, the mother performs the actions in accordance with the daughter’s prior instructions, apparently ratifying the epistemic stance the daughter has established. Then, in line 28, the daughter produces a positive assessment of the mother’s folding performance (“Yes, very good”). Such an assessment, produced in second position, is generally contingent upon the speaker’s entitlement that derives from their epistemic authority in the relevant domain (Pomerantz, 1984; Heritage, 2002). By evaluating the mother’s action with “very good,” the daughter further reinforces her epistemic claim regarding the paper-folding activity.

However, in line 30, the mother initiates a repair through a partial repeat (“Very good?”), followed by “I taught you, *haobuhao*?” (“I taught you, okay?”), thereby making explicit her

understanding and stance toward the current epistemic relationship. The partial repeat here does not display a hearing problem but instead indexes the mother's suspicion regarding the action the daughter's prior assessment implements—namely, whether the daughter is entitled to evaluate her performance. Furthermore, through “I taught you,” the mother re-establishes the epistemic hierarchy for this activity, asserting that she—rather than the daughter—should occupy the instructional role.

The appended *haobuhao* in this position occurs after a syntactically and pragmatically complete TCU, and its removal would not affect the execution of the repair. Therefore, it does not function as an information-seeking question. Instead, it serves to reinforce the speaker's stance display and to emphasize the propositional claim just made. In other words, the mother uses *haobuhao* to strengthen her epistemic position and to secure alignment with her understanding of “who is entitled to teach whom” within this activity framework.

In the subsequent line 32, the daughter counters with “The teacher taught me this,” signaling her non-acceptance of the mother's epistemic reconfiguration. By invoking the authority of “the teacher,” she supports her prior epistemic access. This resistance provides further evidence that the *haobuhao* in line 30 does not solicit information but rather constitutes an interactional strategy through which the mother defends her epistemic primacy when confronted with a challenge.

Overall, in Example 4 the mother's use of the appended *haobuhao* in repair sequences consistently occurs after an assertive or evaluative TCU, functioning to reinforce, highlight, and protect her epistemic authority rather than request information. Its deployment is occasioned by the recipient's production of actions that violate the speaker's understanding of epistemic norms—for instance, “who is entitled to instruct” or “who may assess another's actions.” Thus, in this position, *haobuhao* constitutes an interactional practice used to point out the inappropriateness of the recipient's behavior and to restore the expected epistemic order.

4.2 *Haobuhao* in Initial Position

When parents use *haobuhao* turn-finally in a initial position, it most commonly occurs in the implementation of directive actions. Directives

are pervasive in everyday interaction; as long as a speaker's utterance is designed to “get another person to do something,” it may be treated as a directive (Goodwin, 2006). A directive is not a single speech act but rather a complex speech genre (Goodwin, 1990) that encompasses a range of linguistic formats (e.g., imperatives, descriptions) and a range of actions (e.g., commands, requests). In parent-child interaction, parents typically possess higher epistemic authority and deontic authority, which grants them the right and responsibility to prompt the child to adjust behaviors that violate social norms. This is particularly salient with younger children, for whom parents routinely undertake the work of socialization through directive actions.

Example 5 occurs during mealtime talk. Structurally, prior to line 153 the mother is explaining how she knows what the child ate in kindergarten; lines 154-158 constitute an inserted directive sequence, and in line 159 the talk returns to the prior topic. This shows that the directive is not projected by the preceding turns but is launched in response to the mother's real-time monitoring of the child's behavior. Such side-sequence directives demonstrate that, given the parental role and the authority attached to it, directive actions can legitimately disrupt the ongoing progressivity of the main sequence.

Extract 5. [亲子饭间对话: 00:03:47-00:04:02]

150 妈妈: 我星期五的时候看见你王老师贴在小学-幼儿园门口了。

151 (0.2)

152 女儿: 哦。

153 妈妈: 这一星期的饭(0.4)都贴在>门口了<。

154 你把腿放好了,>好不好?<

155 女儿: 好吧。((女儿把腿拿了下去))

156 (1.2)

157 妈妈: 老拿起来干什么。

158 (1.7)

159 女儿: 你>刚才<说啥?

In this case, the mother notices that the daughter has lifted her feet onto the chair—a breach of basic table manners and social norms—and therefore issues a directive in line 154 requiring her to put her leg down. In this directive, *haobuhao* does not function as a genuine inquiry

into the child's preference, nor does it present a real choice between "yes" and "no." Rather, it constrains the projected response to a single acceptable option: the child should immediately comply with the mother's instruction. The daughter's behavior in line 155 confirms this interpretation demonstrating shared understanding of the deontic rights at play. Accordingly, the removal of *haobuhao* would not affect the implementation of the directive itself.

When *haobuhao* occurs in directives that require the recipient to correct an inappropriate behavior, it adds a layer of reproach to the action. Although the directive succeeds, the daughter's response *hao ba*—rather than a straightforward, unmarked *hao*—constitutes a dispreferred format (Wu & Yang, 2020; Wu, 2022). From a facework perspective, this mitigated acceptance may reflect her experience of the mother's turn as face-threatening due to its embedded criticism. Since a single turn may accomplish multiple actions, the *haobuhao*-marked directive here not only enforces behavioral compliance but simultaneously indexes the mother's negative assessment of the child's conduct.

Overall, in initial position, turn-final *haobuhao* within directive actions functions not as a request for information but as an interactional resource for highlighting the inappropriateness of the child's behavior and asserting parental deontic authority. It thus operates as a conversational practice through which parents enact norm enforcement and maintain the expected moral and social order in parent-child interaction.

Extract 6. [亲子饭间对话: 00:08:44-00:08:51]

- 355 妈妈: >你那筷子还用吗?<
 356 (0.4)
 357 女儿: 不用啊.
 358 (0.4)
 359 妈妈: 不用,你举着它-耽误我 chao 菜菜,好不好?
 ((方言 chao 菜是夹菜的意思))
 360 女儿: >好吧.<
 不用,你举着它耽误我夹菜

Excerpt 6 occurs during a parent-child mealtime interaction. In line 355, the mother initiates a pre-directive with "Are you still using those chopsticks?" After a 0.4-second silence, the daughter responds in line 357 with "No,"

thereby providing the conditional relevance and legitimacy for the mother to proceed with a directive—that is, confirming that the mother has both the right and an adequate reason to ask her to put down the chopsticks that are not currently in use.

In line 359, the mother issues the directive in the format of a declarative plus *hao bu hao*: "Since you're not using them, you holding them up is getting in the way of my picking up food, *hao bu hao*?" Although the grammatical format differs from Excerpt 5, where the mother used an imperative plus *hao bu hao*, the identification of a directive action is not determined by syntax. The daughter's immediate compliant response "Fine" in line 360 shows that she recognizes line 359 as a directive. Her subsequent action of putting the chopsticks down demonstrates that the directive has been successfully implemented, and then the sequence is closed.

From a sequential perspective, since the daughter has already confirmed in the pre-sequence that she is not using the chopsticks, the mother has obtained the requisite entitlement and grounds to issue the directive. Thus, even without *hao bu hao*, the directive in line 359 would still be fully performable. However, the TCU preceding *hao bu hao* does more than issue a directive: it also frames the daughter's current behavior as causing inconvenience to the mother, thereby containing a complaint component. The addition of *hao bu hao* does not solicit information or offer a genuine choice; rather, it upgrades the stance of complaint, strengthening the reproachful tone embedded in the mother's directive.

As in Excerpt 5 and 6, when *hao bu hao* occurs within a directive that is implemented through command or complaint, its function is augmentative: it adds a layer of reproach to the directive without altering its basic action. It thus allows the parent not only to exercise behavioral control but also to morally evaluate and discipline the child's conduct. The daughter's reluctant-sounding response across both excerpts further supports the claim that *hao bu hao* in this position carries a reproachful tone with recognizable interactional consequences.

In sum, when appended to directives delivered through commands or complaints, *hao bu hao* serves to intensify the parent's negative stance toward the child's behavior. Rather than seeking information, it contributes a reproach

component, enabling parents to both enforce behavioral compliance and mark the child's conduct as inappropriate within the ongoing activity.

Example 7 differs from the preceding two cases in that the mother does not seek to modify the child's behavior due to any impropriety; instead, she enacts a directive through a request.

Extract 7. [折纸比赛: 00:01:45-00:01:52]

- 01 女儿: 咱们来比赛,好不好?
02 妈妈: 我还没找到(.)折纸王子教我呢,
03 妈妈: 我找到折纸王子那一页 he¹,好不好?
04 女儿: 好啊.

Example 7 takes place in a playful context in which the mother and daughter are preparing to compete in an origami contest. Because the daughter has learned origami at school while the mother has not, the mother needs to locate the "Paper-folding Prince" tutorial on a tablet before she can begin. After the daughter proposes starting the competition, the mother in line 03 requests that the daughter wait until she locates the tutorial. Following Couper-Kuhlen's (2014) analysis of the relationship between benefactor/beneficiary roles and action types, the mother's move in line 03 constitutes a request.

In the activity framework of a "competition," social norms stress fairness: participants should enjoy equivalent rights and obligations at the starting point. Thus, the epistemic and deontic authority ordinarily associated with the mother role is not expected to be strongly exercised here. As co-participants in the competition, the mother's request is legitimately open to acceptance or rejection. The relative low entitlement of the mother in this local context renders the turn-final *haobuhao* genuinely interrogative: it offers a real choice in the response space rather than projecting a single preferred outcome.

If the *haobuhao* were omitted from line 03, the remaining TCU would still implement a directive, but the action type would shift from a request to a command, violating the locally relevant norm of participant equality. Thus, in low-entitlement positions, speakers may append

haobuhao to directives to downgrade the directive force, enabling them to implement a directive in the form of a request. This adjustment aligns the action with expectations of politeness and interactional order, while simultaneously indicating the mother's recognition of the child's behavioral autonomy in this context.

To summarize, turn-final *haobuhao* in parent-child directives is both positionally and actionally sensitive. First, when *haobuhao* is attached to directives enacted through commands or complaints, it intensifies a blaming or admonishing tone, highlighting the child's behavioral impropriety and allowing the parent to perform a blaming action alongside the directive. Second, when *haobuhao* is attached to directives enacted through requests, it softens the directive force, temporarily downshifting parental deontic authority and enabling the directive to be realized in a socially compatible manner. These contrasting uses illustrate how parent-child power relations are maintained, negotiated, and recalibrated in situated interaction, demonstrating the multifunctionality of *haobuhao* in directive environments.

5. Discussion

Drawing on a conversation-analytic approach, this study has examined in detail the turn-final appended interrogative *haobuhao* ("OK?/alright?") as it appears in Mandarin parent-child interaction. The analysis has shown how, across different sequential environments, parents' deployment of *haobuhao* invokes distinct epistemic and deontic norms, and how it functions as an interactional practice within the institutional context of parenting.

The findings indicate that turn-final *haobuhao* in responding positions is typically associated with repair. In these cases, *haobuhao* does not carry an interrogative meaning; prosodically, it is tightly attached to the preceding TCU, forming a continuous and often emphasized unit. This usage serves as a conventionalized practice through which parents highlight the inappropriateness of the child's prior action and assert their epistemic authority. Notably, all instances of turn-final *haobuhao* in repair sequences are parent-initiated and parent-carried-out repairs which belongs to the other-initiated other-repair type. This diverges from the widely observed preference for

¹ he: In the dialect, "he" means "wait until I finish...". Here, the mother's meaning is "Let's start the competition after I find...".

self-initiated self-repair in ordinary conversation (Schegloff et al., 1977). It can be argued that this divergence is attributable to the institutional nature of parent-child interaction: given parents' role as caregivers and the asymmetrical distribution of power, parents bear epistemic responsibility and deontic rights to initiate repair when children display behavior that deviates from social norms. Through such repair initiations, parents not only maintain their epistemic authority but also guide the child's socialization into appropriate conduct.

Turn-final *haobuhao* in initial positions is predominantly associated with directive actions. However, its interactional function varies depending on the specific type of directive being performed. When parents issue directives through commanding or complaining formats, *haobuhao* adds a layer of moral accountability and transforms the directive into a directive-plus-reproach action. In such cases, the interrogative morphology does not implement a genuine question but strengthens the deontic force of the directive and foregrounds the child's breach of normative expectations. By contrast, when *haobuhao* is used in a request-based directive, it serves a different interactional purpose. In these contexts—often those in which the parent does not hold clear deontic superiority, such as collaborative activities or games—turn-final *haobuhao* mitigates the directive's impositive force and renders it more affiliative and negotiable. Here, parents modulate their deontic authority in order to align with the moral and structural requirements of the activity (e.g., fairness in a game) and to acknowledge the child's behavioral autonomy. Thus, *haobuhao* functions as a resource for recalibrating authority relations within the interactional moment.

In sum, the use of turn-final *haobuhao* is closely tied to the negotiation of knowledge and rights in parent-child interaction. When responding to children's norm-deviant actions, parents—who hold greater epistemic and deontic authority institutionally—tend to use *haobuhao* to strengthen repair or directive-plus-reproach actions, thereby asserting or protecting their epistemic standing and marking the child's conduct as inappropriate. In these contexts, *haobuhao* embodies adherence to institutional norms associated with the parental role. However, parents do not always exercise authority in a straightforward or unilateral

manner. In request-based directives, parents employ *haobuhao* to soften their deontic stance and to display sensitivity to the child's autonomy. Here, the meaning and function of *haobuhao* shift in accordance with the social action being implemented, and the practice contributes to the local reconstruction of epistemic and deontic relations. These findings demonstrate that language use in specific sequential environments plays a constitutive role in shaping norms of knowledge and action rights within the parenting institution.

6. Conclusion

Drawing on naturally occurring Mandarin parent-child interactions, this study has examined the sequential distribution and social action functions of turn-final *haobuhao* as an appended interrogative, demonstrating how this practice is mobilized in local negotiations of knowledge, authority, and behavioral regulation. Rather than a mere modal particle, *haobuhao* emerges as a finely tuned interactional resource through which parents modulate entitlement, enact or mitigate directive force, and display adherence to—or recalibration of—norms of epistemic and deontic authority. These findings highlight the central role of micro-interactional practices in the socialization process and contribute empirical evidence to understanding how epistemic and normative orders are accomplished in Mandarin parent-child talk. Nonetheless, the present study is limited by the size and situational scope of the dataset, and has not systematically addressed variation across families or participant characteristics. Future research could draw on larger and more diverse corpora to further test the generalizability of these patterns, and explore how turn-final *haobuhao* intersects with broader socio-cultural factors—such as family dynamics, social class, and trajectories of language socialization—to build a more comprehensive account of its sequential and social functions.

References

- Atkinson, J. M., & Heritage, J. (1984). *Structures of social action: Studies in emotion and social interaction*. Cambridge University Press.
- Couper-Kuhlen, E. (2014). What does grammar tell us about action? *Pragmatics*, 24(3), 623-647.
- Drew, P. (2013). Conversation analysis and social action. *Journal of Foreign Languages*, 3, 2-19.

- Gao, H. (2009). On the discourse modality of the “hao-bu-hao” tag question. *Journal of Shenzhen University (Humanities & Social Sciences)*, 26(4), 98-102.
- Goodwin, M. H. (1990). *He-said-she-said: Talk as social organization among Black children*. Indiana University Press.
- Goodwin, M. H. (2006). Participation, affect, and trajectory in family directive/response sequences. *Text & Talk*, 26(4-5), 515-543.
- Heritage, J. (2002). Oh-prefaced responses to assessments: A method of modifying agreement/disagreement. In C. Ford, B. Fox, & S. Thompson (Eds.), *The language of turn and sequence* (pp. 196-224). Oxford University Press.
- Heritage, J. (2012a). Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction*, 45(1), 1-29.
- Heritage, J. (2012b). The epistemic engine: Sequence organization and territories of knowledge. *Research on Language and Social Interaction*, 45(1), 30-52.
- Heritage, J. (2018). The ubiquity of epistemics: A rebuttal to the “epistemics of epistemics” group. *Discourse Studies*, 20(1), 14-56.
- Heritage, J., & Raymond, G. (2005). The terms of agreement: Indexing epistemic authority and subordination in assessment sequences. *Social Psychology Quarterly*, 68(1), 15-38.
- Margutti, P., Tainio, L., Drew, P., & Traverso, V. (2018). Invitations and responses across different languages: Observations on feasibility and relevance of a cross-linguistic comparative perspective on the study of actions. *Journal of Pragmatics*, 125, 52-61.
- Pomerantz, A. M. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge University Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289-327.
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53(2), 361-382.
- Stevanovic, M. (2013). Deontic rights in interaction: A conversation analytic study on authority and cooperation (Doctoral dissertation). University of Helsinki.
- Stevanovic, M., & Peräkylä, A. (2012). Deontic authority in interaction. *Research on Language and Social Interaction*, 45(3), 297-321.
- Stivers, T., Mondada, L., & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 3-24). Cambridge University Press.
- Wang, S. (2017). The discourse functions of the “X bu X” construction: Based on stance taking. *Chinese Language Learning*, 5, 94-103.
- Wu, Y. (2022). The conversation analytic approach to particles in Mandarin Chinese. *Journal of Foreign Languages*, 45(6), 21-33.
- Wu, Y., & Yang, Y. (2020). Haoba as the sequence-closing third in Chinese mundane conversation: A conversation analytic approach. *Linguistic Research*, 27, 19-31.
- Wu, Y., & Zhou, X. (2015). A comparison between pragmatics and conversation analysis. *Journal of Shanxi University (Philosophy and Social Science Edition)*, 38(4), 46-50.
- Yan, Y. (2015). Syntax and pragmatic functions of Modern Chinese tag questions. *Linguistic Research*, 3, 44-50.
- Yu, G. (2022). *What is conversation analysis* [In Chinese]. Shanghai Foreign Language Education Press.
- Yu, G., & Wu, Y. (2018). Building a scientific system for Chinese conversation analysis research. *Journal of Foreign Languages*, 4, 7-9.
- Yu, G., & Yao, Y. (2009). A pragmatic research on the expressional function of haobuhao and its process of formation. *Linguistic Sciences*, 8(6), 625-632.

Voice-Over as a Substitute for Visual Explanation in Short Videos

M. D. Reynolds¹

¹ University of Guelph, Canada

Correspondence: M. D. Reynolds, University of Guelph, Canada.

doi:10.63593/JLCS.2025.12.04

Abstract

Short-form video has become a dominant mode of audiovisual communication on contemporary digital platforms, where limited duration, rapid editing, and fragmented viewing practices increasingly undermine the explanatory capacity of visual continuity. This paper examines the growing role of voice-over narration as a substitute for visual explanation in short videos and argues that voice-over should be understood not as a stylistic or technical supplement but as a core explanatory mechanism. Drawing on audiovisual theory, narration studies, and platform media research, the analysis shows that visual explanation traditionally depends on temporal development, spatial coherence, and sustained attention, conditions that are structurally weakened in short video formats. In response, voice-over assumes primary explanatory authority by guiding interpretation, condensing processes, and stabilizing meaning across fragmented visuals, offering a level of clarity and abstraction that images alone struggle to achieve under accelerated and distracted consumption. The paper situates this shift within broader media conditions such as everyday multitasking, the perceptual stability of sound, platform norms favoring rapid comprehension, and cultural preferences for explicit guidance, and it discusses the narrative and cultural implications of this transformation, including a movement from showing to telling, reduced interpretive openness, and the normalization of guided meaning. By reframing voice-over as a substitute for visual explanation, the study challenges assumptions of visual dominance in audiovisual media and highlights a rebalancing of sound and image in platform-based communication.

Keywords: voice-over narration, short-form video, visual explanation, audiovisual theory, platform media, guided meaning

1. Introduction

1.1 Background

Short-form video platforms have become a central infrastructure of contemporary media consumption and everyday communication. They are no longer marginal or supplementary to traditional media forms but operate as a

primary interface through which information, knowledge, and cultural meanings circulate. The defining characteristics of short videos include extreme temporal compression, algorithmically driven distribution, and an emphasis on immediate perceptual impact. These conditions shape not only what kind of content is produced but also how meaning is constructed and

received.

Within this environment, visual material is subject to intense pressure. Images are required to attract attention instantly, maintain engagement within seconds, and remain legible on small mobile screens. Editing practices privilege speed and rhythm over continuity, while narrative structures are often reduced to highlights, summaries, or isolated moments. As a result, visuals are frequently fragmented and decontextualized. They present effects without causes, outcomes without processes, and moments without extended temporal development. The traditional expectation that images can explain through duration, spatial coherence, and sequential logic becomes increasingly difficult to sustain.

At the same time, short videos circulate within conditions of everyday distraction. Viewing often occurs alongside other activities, with intermittent visual attention and fluctuating levels of cognitive engagement. Under such circumstances, visual meaning alone becomes unstable. Images may be seen but not fully processed, or they may be skipped altogether while audio continues to play. These viewing practices further undermine the explanatory capacity of visuals and create demand for alternative channels of meaning transmission.

It is within this media ecology that voice-over narration has become pervasive. Across diverse genres, spoken language is used to explain, summarize, interpret, and frame visual content. Voice-over no longer appears as an optional stylistic layer added to otherwise self-sufficient images. Instead, it increasingly functions as a structural component that stabilizes meaning in an environment where visual explanation is weakened by fragmentation, speed, and distraction.

1.2 Research Focus

This paper approaches voice-over not as a technical feature or a narrative embellishment but as a functional response to the changing conditions of audiovisual meaning-making. The focus lies on why voice-over increasingly operates as a substitute for visual explanation rather than a supplement to it. The analysis centers on the relationship between fragmented visuals and spoken narration, asking how voice-over compensates for what images can no longer reliably convey on their own.

Rather than examining production techniques or

conducting empirical audience studies, the paper adopts a theoretical and conceptual perspective. It seeks to understand how explanatory labor is redistributed between sound and image in short video formats. Attention is directed toward the ways spoken language guides interpretation, condenses information, and resolves ambiguity in contexts where visuals are insufficiently continuous or complete.

The core question is not whether voice-over improves clarity but why clarity itself has become increasingly dependent on speech. By situating voice-over within broader platform conditions such as acceleration, attention scarcity, and algorithmic optimization, the paper frames narration as an adaptive mechanism shaped by structural constraints rather than individual creative choice.

1.3 Research Significance

Conceptualizing voice-over as an explanatory mechanism has implications for how short videos are understood within media theory. Audiovisual studies have often privileged the image as the dominant carrier of meaning, with sound positioned as supportive or secondary. The prominence of voice-over in short videos challenges this hierarchy and calls for a reconsideration of how meaning is organized across sensory channels in contemporary media.

This perspective contributes to discussions on narration, perception, and platform culture by highlighting a shift in explanatory authority from visual demonstration to verbal articulation. It suggests that short video platforms foster forms of communication in which meaning is increasingly guided, condensed, and stabilized through speech. Such a shift has consequences for audience interpretation, narrative openness, and visual literacy.

By focusing on voice-over as a substitute rather than an addition, this study provides a framework for understanding broader transformations in audiovisual communication under platform conditions. It opens space for further inquiry into how sound, voice, and language reassert their centrality in media environments often assumed to be dominated by images.

2. Visual Explanation and Its Limits in Short Videos

2.1 Visual Explanation in Audiovisual Media

Visual explanation has long been regarded as a foundational capacity of audiovisual media. In cinema, television, and documentary traditions, images are expected to carry explanatory weight through their organization in time and space. Visual meaning emerges from the ability of images to show relations rather than simply present isolated objects or moments. Processes become intelligible when actions unfold gradually, when changes can be observed across duration, and when spatial continuity allows viewers to orient themselves within a scene. Explanation is achieved not by stating facts directly but by arranging visual elements so that understanding arises through observation.

This mode of explanation presupposes a certain stability in viewing conditions. Images are given time to develop, and viewers are assumed to maintain sustained attention. Continuity editing, establishing shots, and sequential framing all contribute to a visual logic in which causes precede effects and actions lead to consequences. Even when audiovisual works rely on montage or symbolic imagery, explanation still depends on the accumulation of visual cues that can be connected through interpretive effort.

Visual explanation also relies on the contextual richness of images. Background details, environmental cues, and bodily gestures provide information that exceeds explicit narration. A scene can communicate social relations, emotional states, or power dynamics without verbal clarification because viewers are able to scan the image, notice details, and integrate them into a coherent understanding. In this sense, visual explanation depends on openness and density rather than speed. The image offers more information than is immediately consumed, allowing meaning to emerge gradually.

Another key feature of visual explanation is its reliance on inferential participation. Viewers are not simply recipients of information but active interpreters who construct meaning by linking visual elements together. Explanation is distributed across shots and sequences rather than concentrated in a single moment. This model assumes that viewers are willing and able to invest cognitive effort in interpretation. Visual explanation therefore aligns with media forms that value immersion, narrative depth, and interpretive engagement.

2.2 Structural Constraints of Short-Form Video

Short-form video platforms introduce a radically different set of structural conditions that challenge the assumptions underlying traditional visual explanation. Duration is the most obvious constraint. When a video lasts only a few seconds or minutes, there is little room for gradual development. Processes must be compressed or omitted, and visual sequences are often reduced to highlights rather than complete representations. The image no longer unfolds but appears in fragments that are immediately replaced by the next visual stimulus.

Editing practices in short videos prioritize rhythm and immediacy. Rapid cuts, jump edits, and abrupt transitions are common, not as stylistic experimentation but as normative conventions. Spatial coherence is frequently disrupted as scenes shift locations without establishing context. Temporal continuity is often abandoned, with past, present, and future moments collapsed into a single sequence. Under these conditions, images struggle to establish causal relations. Actions may be shown without sufficient buildup, and outcomes may appear without clear explanation of how they were achieved.

The visual economy of short videos is also shaped by platform competition. Countless videos vie for attention within algorithmically curated feeds, encouraging creators to maximize visual impact in the shortest possible time. Images are designed to capture attention instantly rather than sustain understanding. Visual shock, novelty, and emotional intensity are rewarded, while slow exposition is penalized by the risk of being skipped. This incentive structure discourages visual explanation that requires patience or sustained observation.

Screen size and viewing environment further complicate visual explanation. Short videos are primarily consumed on mobile devices, often in vertical formats that limit visual scope. Fine details may be difficult to perceive, and complex spatial arrangements may lose clarity. Viewing conditions are rarely controlled, with ambient distractions competing for attention. Visual explanation that depends on subtle cues or careful observation becomes fragile under such circumstances.

Short-form video narratives are also shaped by

repetition and modularity. Many videos follow similar templates, formats, or trends, leading to visual standardization. When images become predictable, they lose explanatory specificity. A gesture, setting, or visual trope may signal a genre or emotion without providing concrete information about a particular situation. Visual meaning becomes shorthand rather than explanation, relying on prior familiarity rather than visual demonstration.

2.3 *The Weakening of Visual Self-Sufficiency*

These structural constraints contribute to a weakening of visual self-sufficiency in short videos. Visual self-sufficiency refers to the capacity of images to explain themselves through their internal organization and contextual richness. In short video formats, this capacity is undermined by fragmentation, speed, and decontextualization. Images are no longer expected to carry meaning independently but function as partial signals that require external support.

One consequence of this shift is the separation of attention from comprehension. Viewers may register visual stimuli without fully understanding them. An image can be seen, recognized, or emotionally felt without being cognitively processed in depth. Visual attention becomes fleeting, while understanding requires additional cues that images alone cannot reliably provide. The gap between perception and comprehension widens as images circulate faster than they can be interpreted.

Another consequence is the reduction of visual causality. Short videos often present results rather than processes. A transformation, achievement, or outcome is shown without visual evidence of how it occurred. The explanatory chain is broken, leaving images as endpoints rather than narratives. In such cases, visuals attract curiosity but fail to satisfy it. They prompt questions that cannot be answered visually within the available time and structure.

Visual ambiguity also takes on a different role. In longer audiovisual forms, ambiguity can invite interpretation and deepen engagement. In short videos, ambiguity is often experienced as confusion rather than openness. When viewers lack time or context to resolve ambiguity, unclear images risk disengagement. The tolerance for interpretive uncertainty decreases, and the demand for immediate clarity increases. Visual explanation struggles under these

conditions because it traditionally relies on interpretive openness and gradual meaning formation.

The weakening of visual self-sufficiency also affects the authority of images. When visuals cannot fully explain, they lose their status as primary sources of meaning. They become dependent on external framing to anchor interpretation. This dependence does not necessarily diminish the aesthetic appeal of images, but it changes their function. Images shift from being explanatory structures to being illustrative elements that support meaning articulated elsewhere.

This transformation is closely linked to changes in audience expectations. Viewers accustomed to narrated short videos may no longer approach images with the assumption that they should be interpreted independently. Instead, they may wait for verbal cues to clarify meaning. Visual literacy practices adapt to platform norms, and the skill of extracting explanation from images alone becomes less central. Visual self-sufficiency declines not only because of structural limitations but also because of changing habits of interpretation.

The limits of visual explanation in short videos therefore cannot be attributed solely to technical constraints. They reflect a broader reconfiguration of audiovisual communication. Images remain powerful in capturing attention and evoking emotion, but their capacity to explain is increasingly constrained by the conditions under which they are produced, distributed, and consumed. Visual explanation, once a defining strength of audiovisual media, becomes unreliable in environments characterized by acceleration, fragmentation, and partial attention.

This does not imply the disappearance of visual meaning. Instead, it suggests a redistribution of explanatory labor. Visuals continue to play an important role, but they no longer operate as autonomous explanatory systems. Their meaning is increasingly stabilized through other channels, particularly spoken language. Understanding the limits of visual explanation in short videos is therefore a necessary step toward understanding why voice-over emerges as a substitute rather than a supplement in contemporary platform media.

3. **Voice-Over as an Explanatory Mechanism**

3.1 *Voice-Over as Semantic Guidance*

Voice-over operates as a direct mechanism for organizing meaning in short videos by providing explicit semantic guidance. Spoken language has the capacity to name, classify, evaluate, and explain in ways that images alone often cannot under conditions of fragmentation and speed. Through voice-over, creators are able to establish what a scene is about, which elements are relevant, and how the viewer should interpret what is shown. This guidance does not emerge gradually but is delivered immediately, often within the first seconds of a video, shaping interpretation before the visuals have had time to develop their own logic.

Semantic guidance through voice-over functions by reducing uncertainty. Images in short videos frequently present incomplete information. A scene may show an action without context, a result without process, or an object without explanation. Voice-over fills these gaps by supplying missing links. It can state causal relations that are visually absent, define abstract concepts that cannot be shown directly, or situate an image within a broader narrative or argument. In doing so, voice-over transforms visual fragments into legible units of meaning.

This guiding function also operates at the level of attention. Voice-over directs viewers toward specific aspects of the image, telling them what to notice and what to ignore. In a visually dense or rapidly changing scene, not all elements can be processed equally. Spoken narration prioritizes certain details by naming them or assigning significance to them. The viewer's gaze is guided not by visual exploration but by verbal instruction. Interpretation becomes aligned with the logic of speech rather than the openness of the image.

Voice-over guidance also structures emotional interpretation. Images can evoke multiple emotional responses depending on context and personal experience. Voice-over narrows this range by framing how a scene should be felt. A visual moment that might otherwise appear neutral or ambiguous can be defined as inspiring, shocking, tragic, or humorous through narration. Emotional meaning becomes less dependent on visual cues and more dependent on verbal framing.

In this sense, voice-over does not merely add information to visuals. It actively organizes perception. It establishes a hierarchy of meaning in which spoken language occupies a directive

role, shaping how images are seen, understood, and evaluated. Semantic guidance through voice-over reflects a shift in explanatory authority away from visual inference toward verbal instruction.

3.2 Substitution Rather Than Supplementation

In traditional audiovisual theory, voice-over is often described as supplementary. It is seen as an additional layer that supports or enhances visual meaning without replacing it. In short video practices, this relationship is increasingly reversed. Voice-over becomes the primary explanatory channel, while visuals function as secondary elements that illustrate or reinforce what is already explained verbally.

This substitution occurs because visuals in short videos often lack the conditions required for autonomous explanation. Fragmented shots, compressed timelines, and minimal context limit the ability of images to construct meaning independently. Rather than attempting to restore visual continuity, creators rely on voice-over to carry the explanatory burden. Spoken narration provides the structure that images can no longer sustain.

In many short videos, the narrative logic unfolds entirely through speech. The voice-over introduces the topic, explains its significance, and delivers conclusions. Visuals appear as loosely connected scenes that accompany the narration but do not determine it. If the audio were removed, the video would often become difficult or impossible to understand. If the visuals were removed, the core message might still remain intelligible. This asymmetry indicates a clear substitution rather than mutual dependence.

The substitution of voice-over for visual explanation also reshapes the function of images. Images become illustrative rather than demonstrative. They serve to maintain viewer interest, provide sensory stimulation, or evoke emotion, while the task of explanation is handled by language. Visuals no longer need to show how something works or why it happens. They only need to align loosely with what is being said. This reduces the demand placed on visual coherence and allows for greater flexibility in image selection.

This shift also affects narrative hierarchy. In classical audiovisual storytelling, images often lead and sound follows. In short videos with dominant voice-over, sound leads and images

follow. Visual sequences are edited to match the rhythm and content of speech rather than the other way around. The narrative spine is verbal, and the image track adapts to it. This inversion reflects a broader transformation in how audiovisual meaning is assembled under platform constraints.

Substitution rather than supplementation also alters the viewer's interpretive role. When explanation is carried by voice-over, viewers are less required to infer meaning from visual relationships. Interpretation becomes a process of listening and accepting rather than observing and connecting. Visual engagement remains present, but it is guided and constrained by verbal explanation. The image loses some of its autonomy as a site of meaning production.

3.3 Efficiency of Verbal Explanation

The efficiency of verbal explanation is a key reason for the prominence of voice-over in short videos. Spoken language is capable of compressing complex information into concise statements. Causal chains, abstract ideas, and evaluative judgments can be expressed in seconds. Visual explanation of the same content would require extended sequences, multiple shots, and careful pacing, which short video formats rarely allow.

Language excels at summarization. A voice-over can condense an entire process into a brief verbal description, bypassing the need to show intermediate steps. This is particularly valuable in educational, informational, or commentary-based short videos, where the goal is to transmit knowledge quickly. Visual demonstration of a process may be incomplete or misleading if shown too briefly. Verbal explanation avoids this risk by stating conclusions directly.

Verbal explanation is also efficient in handling abstraction. Many concepts addressed in short videos are not easily visualized. Ideas related to motivation, social trends, ethical judgments, or personal reflection often lack concrete visual referents. Voice-over allows creators to address such topics without relying on symbolic or metaphorical imagery, which may be difficult to interpret quickly. Speech provides clarity where visuals might introduce confusion.

The temporal flexibility of voice-over contributes to its efficiency. Speech can overlap with visuals, allowing explanation and imagery to coexist within the same time frame. Visual

explanation often requires sequential time. A process must be shown step by step. Voice-over can explain while something else is being shown, maximizing the use of limited duration. This parallelization of information channels makes voice-over particularly suited to short formats.

Efficiency also relates to cognitive load. Visual explanation often requires viewers to actively process relationships between images. This demands attention and interpretive effort. Voice-over reduces this burden by delivering meaning directly. Viewers can understand the message without reconstructing it themselves. In environments characterized by distraction and rapid consumption, this reduction in cognitive effort is advantageous.

The efficiency of verbal explanation aligns with platform incentives. Algorithms favor content that is quickly understood and easily consumed. Voice-over supports these goals by ensuring that meaning is accessible even during brief viewing moments. Speech can convey the main point early, increasing the likelihood that viewers will continue watching or engage with the content.

Efficiency, however, is not a neutral quality. While it enables communication under constraints, it also shapes the form of meaning that is communicated. Efficient verbal explanation tends to favor clarity over complexity and conclusion over exploration. It supports content that delivers answers rather than raises questions. The dominance of voice-over as an explanatory mechanism reflects a broader orientation toward speed and legibility in platform media.

Taken together, semantic guidance, substitution of visual explanation, and the efficiency of verbal language explain why voice-over has become central to meaning-making in short videos. Voice-over is not merely a response to technical limitations but a structural adaptation to the conditions of contemporary media consumption. It reorganizes the relationship between sound and image, redefining how explanation is achieved in compressed audiovisual forms.

4. Media Conditions Favoring Voice-Over Explanation

4.1 Partial Attention in Everyday Media Use

Short-form video consumption is deeply embedded in everyday routines rather than

separated as a dedicated viewing activity. Unlike cinema or television, which traditionally require a fixed viewing position and a relatively stable attentional state, short videos are consumed in transit, during breaks, and alongside other tasks. Viewing often takes place while commuting, waiting, eating, or switching between applications. In these situations, attention is divided and unstable. The screen may be glanced at briefly, then ignored, then returned to, creating a fragmented pattern of visual engagement.

Under such conditions, visual explanation becomes unreliable. Visual meaning presupposes sustained attention, even when images are simple or familiar. Understanding visual relations requires looking, comparing, and integrating what is seen across time. Partial attention interrupts this process. When viewers miss key visual moments or fail to track visual continuity, images lose their explanatory power. They remain perceptible as stimuli but fail to cohere into intelligible sequences.

Audio operates differently in distracted contexts. Sound can be received without direct orientation toward the screen. A viewer may hear narration while looking elsewhere, performing another task, or momentarily disengaging from the video interface. Voice-over remains accessible even when visual attention drops to a minimum. This asymmetry between visual and auditory attention makes sound a more reliable channel for meaning transmission in everyday media use.

Voice-over aligns with this attentional reality by ensuring that the core message does not depend on uninterrupted viewing. Spoken narration allows meaning to persist across moments of visual disengagement. A viewer who looks away from the screen may still follow the argument, explanation, or story through audio alone. When attention returns to the screen, visuals can be reintegrated without having been solely responsible for explanation.

Partial attention also shapes cognitive expectations. Viewers accustomed to distracted viewing may no longer expect to derive meaning primarily from images. Instead, they rely on narration to anchor understanding. Visuals become supplementary cues that enhance or illustrate what is already understood verbally. The explanatory burden shifts toward voice because it accommodates the realities of

divided attention more effectively than visual sequences.

This pattern does not indicate a decline in visual interest or appreciation. Images remain central to attracting attention and maintaining engagement. Their role, however, changes from explaining to accompanying. In a media environment structured around partial attention, voice-over becomes the channel through which coherence is maintained despite interruptions, distractions, and fluctuating focus.

4.2 Sound as a Stable Perceptual Channel

Sound possesses perceptual characteristics that make it particularly suited to explanatory functions in short video environments. Auditory perception unfolds over time in a continuous manner. Speech remains intelligible as long as it is heard, regardless of whether the listener maintains visual contact with the source. Images, by contrast, depend on spatial fixation and visual orientation. When the screen is not actively watched, visual information ceases to be available.

The stability of sound lies in its temporal persistence. Voice-over flows continuously, creating a stable thread of meaning that runs through rapidly changing images. Even when visuals shift abruptly, narration provides continuity. This continuity allows viewers to maintain a sense of progression and coherence despite visual fragmentation. Spoken language connects moments that might otherwise appear disjointed.

Sound is also less susceptible to perceptual overload. Visual environments in short video platforms are highly saturated. Bright colors, rapid motion, text overlays, and transitions compete for attention. This density can overwhelm visual processing, making it difficult to extract explanatory information. Voice-over cuts through visual noise by occupying a different sensory channel. It simplifies the interpretive task by delivering meaning directly, reducing the need to decode complex visual arrangements.

Auditory explanation also benefits from familiarity. Human speech is one of the most deeply learned and practiced forms of communication. Listeners are skilled at processing spoken language quickly, even under suboptimal conditions. Accents, background noise, and variable audio quality can be

accommodated with relative ease. This robustness enhances the reliability of voice-over as an explanatory medium in uncontrolled viewing environments.

The stability of sound also extends to emotional and narrative continuity. A consistent voice can create a sense of presence and authority that anchors the viewer's experience. Even when visuals change style, location, or pace, the voice remains constant. This constancy supports narrative cohesion and reinforces the explanatory role of narration. The voice becomes a guide through visual variation rather than a supplement to it.

In short video contexts, where images are often selected for impact rather than coherence, sound assumes the task of stabilizing meaning. Voice-over does not merely accompany visuals but provides a continuous interpretive framework that compensates for visual volatility. This perceptual stability makes sound an ideal channel for explanation under conditions of rapid visual change.

4.3 Platform Norms and Audience Expectations

Platform environments play a crucial role in normalizing voice-over as an explanatory strategy. Short video platforms are not neutral spaces but structured ecosystems shaped by algorithms, content trends, and user behavior. Over time, certain formats and conventions become dominant because they perform well within these systems. Narrated videos are among the most prominent of these conventions.

As voice-over becomes widespread, audiences develop expectations for verbal explanation. Viewers learn to anticipate narration as a guide to meaning. When a video lacks voice-over, it may be perceived as incomplete, confusing, or less accessible. This expectation is not necessarily conscious but emerges through repeated exposure to narrated content. Platform cultures gradually define what a legible video looks and sounds like.

Creators respond to these expectations by incorporating voice-over as a default element. Even when visuals could potentially explain on their own, narration is added to ensure clarity and alignment with audience habits. Over time, this reinforces the association between explanation and speech. Voice-over becomes institutionalized as a standard feature rather than a creative choice.

Platform metrics further reinforce this trend. Videos that are quickly understood and easily consumed tend to perform better in terms of engagement, retention, and sharing. Voice-over supports these metrics by reducing interpretive effort and accelerating comprehension. Algorithms indirectly favor narrated content by rewarding clarity, which encourages creators to rely on verbal explanation.

Audience expectations also intersect with accessibility. Voice-over can make content more inclusive for viewers with limited visual attention, small screens, or situational constraints. While captions address accessibility for hearing-impaired viewers, voice-over addresses the inverse situation of limited visual access. Platforms that prioritize broad reach implicitly favor content that can be understood through multiple sensory channels, with sound playing a central role.

As narration becomes normalized, visual explanation is no longer the default standard against which videos are judged. Instead, explanation through voice is taken for granted, and images are evaluated based on their ability to complement narration rather than replace it. Platform norms thus reshape the criteria of successful communication, privileging guided meaning over visual inference.

4.4 Algorithmic Time Pressure and Accelerated Comprehension

Short video platforms operate under conditions of extreme temporal competition. Users scroll rapidly through feeds, often spending only seconds on each piece of content. Videos must establish relevance and meaning almost immediately to prevent being skipped. This creates intense time pressure at the level of perception and interpretation.

Visual explanation struggles under such pressure. Visual meaning often requires time to unfold. Establishing context, showing relationships, and allowing viewers to observe changes across duration cannot be easily compressed into the first few seconds. Voice-over, by contrast, can state the main point immediately. A single sentence can define the topic, summarize the outcome, or frame the narrative direction.

Algorithmic systems reward this immediacy. Content that communicates its purpose quickly is more likely to retain viewers and be promoted further. Voice-over allows creators to front-load

explanation, ensuring that viewers understand what the video is about before deciding whether to continue watching. Visual explanation alone cannot always achieve this level of instant clarity.

Accelerated comprehension also affects how content is structured. Videos are increasingly designed around clear verbal hooks, summaries, and conclusions. Voice-over provides these elements efficiently. The image track adapts to this structure by offering illustrative scenes rather than explanatory sequences. The explanatory core becomes verbal, optimized for speed and legibility.

This acceleration reshapes the experience of meaning itself. Understanding is no longer something that unfolds gradually through observation but something that is delivered quickly through narration. The role of the viewer shifts from interpreter to receiver. Voice-over fits seamlessly into this accelerated mode of comprehension, while visual explanation appears slow and uncertain by comparison.

4.5 Multitasking and the Audio Background Function

Short videos often function as background media rather than foreground experiences. Users may play videos while doing other activities, treating audio as a continuous stream that accompanies daily routines. In such cases, visuals may be only intermittently consulted or ignored entirely.

Voice-over enables this mode of consumption by transforming short videos into quasi-audio content. Explanation delivered through speech allows the video to remain meaningful even when not actively watched. This flexibility expands the contexts in which short videos can be consumed and increases their integration into everyday life.

The background function of audio also influences content design. Creators may prioritize narration that remains intelligible without visuals. Explanations are phrased clearly, references to specific visual details are minimized, and meaning is made explicit through speech. Visuals become optional enhancements rather than essential components of understanding.

This shift has implications for how media presence is experienced. Voice-over creates a

sense of companionship or guidance that persists even when attention is divided. The voice accompanies the viewer through other activities, reinforcing its explanatory authority. Visual explanation, which requires focused attention, is less compatible with this mode of engagement.

4.6 Cultural Preferences for Explicitness and Guidance

Media conditions also reflect broader cultural shifts toward explicit communication and guided interpretation. In fast-paced information environments, ambiguity is often experienced as inefficiency rather than openness. Viewers seek content that delivers clear messages, actionable insights, or straightforward narratives.

Voice-over satisfies this preference by making meaning explicit. It reduces the need for inference and minimizes interpretive uncertainty. Visual explanation, which often relies on implication and gradual understanding, may appear vague or demanding under such expectations.

This cultural orientation toward explicitness reinforces the role of narration as an explanatory tool. Voice-over aligns with a desire for clarity and immediacy, while visual explanation may be perceived as indirect or incomplete. Media conditions thus favor modes of communication that prioritize telling over showing.

Taken together, these media conditions explain why voice-over has become central to explanation in short videos. Partial attention, perceptual stability of sound, platform norms, algorithmic time pressure, multitasking practices, and cultural preferences converge to privilege spoken narration. Voice-over is not simply added to short videos but emerges as a structural solution to the conditions under which they are produced and consumed.

5. Narrative and Cultural Implications

5.1 From Showing to Telling

The increasing reliance on voice-over in short videos signals a profound narrative shift from showing to telling. In classical visual storytelling, meaning is expected to emerge through observation. Images unfold in time, allowing viewers to grasp relationships, motivations, and consequences by watching actions take place. The ideal of visual narration emphasizes demonstration over explanation, trusting the audience to infer meaning from

what is seen rather than what is said. Short video practices increasingly depart from this model.

Voice-over reorients narrative structure toward verbal articulation. Instead of allowing images to build meaning through sequence and duration, narration often states meaning directly. What happened, why it happened, and what it signifies are explained explicitly through speech. Visuals become evidence or illustration rather than the primary site of narrative development. This transformation alters the fundamental logic of storytelling. Narrative progression is no longer anchored in visual causality but in verbal sequencing.

This shift aligns short video storytelling with oral narrative traditions. Oral narration privileges immediacy, clarity, and direct address. The storyteller guides listeners through events, highlights relevant points, and interprets significance on their behalf. Meaning is transmitted through voice rather than discovered through observation. Short videos increasingly adopt this mode, positioning the narrator as a guiding presence who leads the audience through a compressed narrative experience.

The move from showing to telling also reflects changing assumptions about audience patience and interpretive labor. Visual storytelling requires time and attention. It assumes that viewers are willing to observe, compare, and infer. Telling assumes that viewers prefer direct explanation and quick comprehension. Voice-over satisfies this preference by delivering narrative meaning efficiently. In doing so, it redefines what counts as effective storytelling under platform conditions.

This transformation challenges long-standing ideals within media theory that value visual subtlety, ambiguity, and experiential immersion. When narration takes precedence, images lose their narrative autonomy. They no longer need to demonstrate processes or embody temporal development. Their role shifts toward supporting verbal claims. The narrative authority moves from the image track to the voice track, reshaping the hierarchy between sound and image.

The implications extend beyond individual videos. As telling becomes normalized, audiences adapt their expectations. Viewers may come to see visual storytelling without narration

as incomplete or unclear. The cultural standard of what a coherent story looks like changes. Storytelling becomes something that is explained rather than shown, spoken rather than visualized.

5.2 Reduced Interpretive Openness

One consequence of narration-centered storytelling is a reduction in interpretive openness. Visual explanation traditionally allows for multiple readings. Images can be ambiguous, inviting viewers to project their own meanings, emotions, or experiences onto what they see. This openness is not a flaw but a defining strength of visual media. It enables richness, complexity, and personal engagement.

Voice-over narrows this openness by defining meaning in advance. When narration explains what an image represents or how it should be understood, alternative interpretations are discouraged. The viewer is guided toward a preferred reading, leaving less space for personal inference. Ambiguity is resolved quickly through verbal clarification. Meaning becomes fixed rather than negotiated.

This narrowing of interpretation affects the role of the audience. Instead of actively constructing meaning from visual cues, viewers are positioned as recipients of explanation. The interpretive work is performed by the narrator, not the viewer. Visual engagement becomes confirmatory rather than exploratory. Images are used to illustrate what has already been said rather than to provoke reflection or questioning.

Reduced interpretive openness also affects emotional experience. Images often evoke complex or mixed emotions that resist easy categorization. Voice-over tends to label emotional meaning, framing a scene as inspiring, tragic, ironic, or humorous. Emotional interpretation becomes guided by language. The viewer's affective response aligns with the narrator's framing rather than emerging from personal resonance with the image.

This process does not eliminate interpretation entirely. Viewers still respond to images, sounds, and narratives in individual ways. The range of acceptable interpretations, however, becomes narrower. Deviations from the intended meaning may feel less legitimate. The space for uncertainty diminishes, replaced by clarity and instruction.

The reduction of interpretive openness aligns

with the temporal constraints of short videos. Ambiguity requires time to resolve or reflect upon. In accelerated media environments, unresolved meaning risks disengagement. Voice-over resolves ambiguity quickly, ensuring that meaning is accessible within limited attention spans. Interpretive openness becomes a liability rather than a virtue.

Over time, this shift may influence broader habits of interpretation. Audiences repeatedly exposed to guided meaning may become less accustomed to extracting meaning independently from images. The skill of visual inference may weaken as narration assumes responsibility for explanation. Visual literacy practices adapt to a media environment where meaning is spoken rather than inferred.

5.3 Normalization of Guided Meaning

The widespread use of voice-over contributes to the normalization of guided meaning in short video culture. Guided meaning refers to a mode of communication in which interpretation is structured, directed, and stabilized through explicit cues. Voice-over is a central instrument of this guidance. It tells viewers how to understand what they see and why it matters.

This normalization is closely tied to platform dynamics. Short video platforms reward content that is easily understood and quickly consumed. Guided meaning reduces the risk of misunderstanding and increases the likelihood of engagement. Voice-over ensures that viewers grasp the intended message even during brief or distracted viewing. As a result, guided meaning becomes a dominant communicative strategy.

Standardization follows from this dominance. When many creators rely on similar narration styles, tones, and structures, meaning presentation becomes uniform. Videos across different topics and genres may share a similar explanatory rhythm. Introduction, explanation, conclusion are delivered through voice-over in predictable patterns. Visual diversity persists, but interpretive structure becomes standardized.

This standardization influences cultural expectations. Viewers learn to expect clear explanations and explicit framing. Content that requires interpretive effort may feel demanding or inefficient. Guided meaning becomes the norm against which other forms of communication are judged. Visual ambiguity or silence may be perceived as absence rather than invitation.

The normalization of guided meaning also intersects with authority. Voice-over often adopts an explanatory or instructive tone. The narrator positions themselves as knowledgeable, credible, or experienced. This establishes a hierarchy between speaker and viewer. Meaning flows from narrator to audience rather than emerging through shared exploration. Cultural authority is vested in the speaking voice.

This has implications for how knowledge and experience are mediated. Short videos increasingly resemble micro-lectures, commentaries, or explanations rather than visual narratives. Information is delivered rather than discovered. The culture of short video becomes one of instruction and guidance, even in entertainment-oriented content.

Guided meaning also affects the pace of cultural consumption. When interpretation is streamlined, content can be consumed more rapidly. Viewers move quickly from one video to the next without lingering uncertainty. This supports the endless scrolling logic of platforms. Voice-over plays a key role in sustaining this flow by minimizing friction in comprehension.

At the same time, guided meaning may limit the depth of engagement. When interpretation is predetermined, there is less incentive to reflect, revisit, or reinterpret. Content is consumed and moved past rather than contemplated. The cultural value of images as sites of exploration diminishes.

Taken together, the narrative and cultural implications of voice-over as an explanatory mechanism are substantial. The shift from showing to telling redefines storytelling norms. Reduced interpretive openness reshapes audience participation. The normalization of guided meaning aligns audiovisual communication with platform demands for speed and clarity. These changes do not signal the disappearance of visual culture but mark a transformation in how images function within contemporary media environments. Voice-over becomes not only a technical feature but a cultural force that reshapes narrative authority, interpretive practices, and the balance between sound and image.

6. Discussion

6.1 Voice-Over as an Adaptation to Visual Overload

The increasing centrality of voice-over in short videos should be understood as an adaptive

response to visual overload rather than a decline in creative standards or visual ambition. Contemporary media environments are saturated with images. Viewers are exposed to an uninterrupted stream of visual stimuli across platforms, applications, and interfaces. This saturation alters how images are perceived and processed. Individual images struggle to retain explanatory force when they appear within an endless sequence of competing visuals.

In such conditions, the capacity of images to sustain attention and generate meaning through observation alone is diminished. Visual explanation depends on the viewer's ability to slow down, compare, and integrate what is seen. Visual overload undermines this process by accelerating consumption and fragmenting attention. Voice-over emerges as a compensatory mechanism that restores coherence under these pressures. By providing a stable explanatory thread, narration allows meaning to persist even as images compete for attention.

This adaptation does not imply that visuals are unimportant. On the contrary, images remain essential for attracting attention and generating affective engagement. What changes is the division of labor between sensory channels. Visuals handle immediacy and emotional impact, while voice-over handles explanation and coherence. This division reflects a pragmatic response to conditions where images alone can no longer reliably perform both functions.

Voice-over also adapts to the temporal compression characteristic of short videos. Visual explanation often requires time to unfold, while narration can deliver meaning instantaneously. Under time pressure, voice-over becomes a rational solution. It allows creators to bypass the limitations of visual duration by stating what cannot be shown fully. The prevalence of voice-over should therefore be interpreted as a structural adjustment to accelerated media rhythms rather than a stylistic preference detached from context.

This perspective challenges critiques that frame narrated short videos as inherently simplistic or lazy. Such critiques often assume that visual explanation remains fully viable and that narration merely replaces effort. In reality, creators operate within constraints that make traditional visual explanation difficult to achieve. Voice-over represents an efficient and adaptive strategy within these constraints, not a

rejection of visual storytelling but a reconfiguration of its possibilities.

6.2 Rethinking Visual Dominance in Media Theory

The rise of voice-over as an explanatory substitute invites a reconsideration of visual dominance in media theory. Audiovisual studies have historically emphasized the primacy of images, often treating sound as secondary or supportive. This hierarchy reflects the influence of cinema studies, where visual composition, montage, and framing have been central analytical categories. Short video practices complicate this framework by foregrounding the explanatory power of voice.

In short videos, sound is not merely atmospheric or emotive. It is structural. Voice-over organizes narrative flow, establishes meaning, and guides interpretation. Without narration, many short videos would be difficult to understand despite their visual richness. This suggests that sound has assumed functions traditionally attributed to images. The explanatory center of gravity shifts from the visual track to the audio track.

This shift does not negate the importance of visual analysis but calls for a more balanced approach. Media theory must account for how sound and image interact under platform conditions that privilege speed, clarity, and accessibility. Voice-over demonstrates that meaning is not inherently visual in audiovisual media. Meaning emerges from the coordination of sensory channels, and the relative importance of each channel varies with context.

Rethinking visual dominance also requires reconsidering assumptions about spectatorship. Visual theories often presume attentive viewing and interpretive engagement. Short video consumption disrupts these assumptions. When attention is partial and fleeting, sound becomes a more reliable carrier of meaning. Media theory must therefore address not only textual properties but also viewing practices and technological environments.

The prominence of voice-over also highlights the role of language in shaping perception. Spoken narration does not simply describe visuals but frames how they are perceived. It influences what is noticed, how events are categorized, and which interpretations are privileged. This linguistic mediation of visual experience challenges the notion of images as transparent or self-evident carriers of meaning.

A revised theoretical framework would treat voice not as an accessory but as a central organizing principle in contemporary audiovisual media. Such a framework would examine how narration structures time, authority, and interpretation in platform-based environments. It would also explore how the balance between sound and image reflects broader cultural shifts toward guidance, efficiency, and explicitness.

6.3 *Limits of Voice-Over Substitution*

While voice-over effectively compensates for the limits of visual explanation, its dominance introduces new constraints and risks. One significant limitation concerns the expressive potential of visuals. When explanation is delegated primarily to speech, images may be relieved of explanatory responsibility but also deprived of narrative agency. Visuals become illustrative rather than exploratory. Their capacity to suggest, imply, or provoke reflection may be underutilized.

Excessive reliance on voice-over can lead to visual redundancy. Images may simply mirror what is being said rather than contribute independent meaning. This redundancy reduces the richness of audiovisual interaction. Instead of multiple channels offering complementary perspectives, meaning becomes centralized in speech. The audiovisual form risks collapsing into a spoken text accompanied by moving images.

Another limitation involves audience passivity. When narration delivers meaning directly, viewers are less required to interpret or infer. The cognitive work of meaning-making is shifted from the audience to the narrator. Over time, this may encourage habits of passive consumption. Viewers listen rather than observe, accept rather than explore. The interpretive skills associated with visual literacy may weaken as narration assumes explanatory authority.

Voice-over substitution also shapes power relations within media texts. The speaking voice often carries authority. It positions itself as knowledgeable and instructive. This can marginalize alternative interpretations and silence visual ambiguity. In contexts where narration frames images in a particular ideological or evaluative manner, viewers may have limited opportunity to question or reinterpret what they see. The voice becomes a

gatekeeper of meaning.

There are also aesthetic consequences. Silence, ambiguity, and visual rhythm have long been important expressive tools in audiovisual media. When voice-over becomes mandatory, these tools are constrained. Moments without narration may feel uncomfortable or incomplete within platform norms. The space for visual contemplation shrinks as continuous explanation fills every temporal gap.

These limitations do not imply that voice-over should be avoided. They suggest that substitution has costs as well as benefits. Voice-over resolves problems created by platform conditions but introduces new forms of constraint. Recognizing these trade-offs is essential for a nuanced understanding of contemporary media practices.

6.4 *Tensions Between Efficiency and Experience*

The discussion of voice-over as an explanatory mechanism reveals a broader tension between efficiency and experience in short video culture. Voice-over enhances efficiency by accelerating comprehension and reducing interpretive effort. It aligns with platform incentives that prioritize retention and engagement. At the same time, this efficiency may come at the expense of experiential depth.

Visual explanation often requires time and openness. It allows viewers to dwell on images, notice details, and construct meaning gradually. Voice-over compresses this process. Meaning is delivered rather than discovered. The experience of watching becomes more informational and less exploratory. This transformation reflects a cultural preference for speed and clarity but raises questions about what is lost in the process.

The tension between efficiency and experience is not easily resolved. Short video platforms are designed for rapid consumption. Expecting them to support slow visual explanation may be unrealistic. Voice-over responds effectively to these conditions. Yet acknowledging its limits opens space for critical reflection on platform design and cultural values.

Some creators experiment with hybrid approaches that balance narration and visual autonomy. They use voice-over sparingly, allowing images moments of silence or ambiguity. Such practices suggest that substitution is not inevitable but contingent. The

dominance of voice-over reflects prevailing conditions rather than fixed necessities.

Understanding this tension encourages a more reflective engagement with short video culture. Instead of celebrating or condemning voice-over outright, it becomes possible to analyze how it shapes experience, authority, and interpretation. Voice-over is neither inherently emancipatory nor inherently restrictive. Its effects depend on how it is used and the conditions under which it operates.

6.5 Implications for Future Media Practices

The discussion of voice-over substitution has implications for future media practices and research. As platforms continue to evolve, the balance between sound and image may shift again. New interfaces, attention patterns, or cultural preferences could alter the role of narration. Understanding current practices provides a foundation for anticipating such changes.

For creators, awareness of the limits of voice-over opens possibilities for experimentation. Narration can be used strategically rather than reflexively. Visuals can be given space to regain explanatory or expressive functions. Such choices may challenge platform norms but also expand the expressive range of short video.

For scholars, the prominence of voice-over invites renewed attention to sound, language, and voice in media analysis. Research can explore how narration shapes perception across different genres, cultures, and audiences. It can examine how voice-over interacts with issues of authority, identity, and power. The study of short videos thus becomes a site for rethinking fundamental assumptions about audiovisual communication.

The discussion underscores that voice-over as a substitute for visual explanation is both a solution and a transformation. It adapts audiovisual meaning-making to conditions of overload, acceleration, and partial attention. At the same time, it reshapes narrative authority, interpretive practices, and aesthetic possibilities. Understanding these dynamics requires moving beyond normative judgments and toward a contextual analysis of how media conditions shape the forms through which meaning is produced and received.

7. Conclusion

This paper has examined the growing prominence of voice-over as an explanatory structure in short-form video and has argued that this development is inseparable from the conditions under which short videos are produced, distributed, and consumed. The central claim is that voice-over has not simply been added to short videos as a stylistic enhancement but has increasingly taken over the explanatory role once associated with visual continuity and visual demonstration. Fragmented images, accelerated editing, limited duration, and unstable attention environments have reduced the capacity of visuals to sustain meaning on their own. Spoken narration compensates for these limitations by organizing interpretation, summarizing complex relations, and maintaining coherence across discontinuous visual sequences.

The analysis has shown that visual explanation relies on temporal development, spatial coherence, and sustained attention, all of which are undermined by the structural logic of short video platforms. Under these conditions, images remain powerful as attention-grabbing and affective elements but lose reliability as explanatory devices. Voice-over emerges as a solution to this problem because language can condense information, clarify causality, and guide interpretation within severe time constraints. Meaning that once unfolded through visual observation is now delivered through speech.

The paper has also demonstrated that this shift is reinforced by broader media conditions. Everyday viewing practices are characterized by partial attention, multitasking, and mobile consumption. Sound remains perceptible in such contexts, while visuals are easily missed. Platform norms and algorithmic incentives reward content that communicates clearly and immediately, favoring narration as a stable and efficient channel. Over time, audiences have come to expect guided explanation, and creators have adapted their practices accordingly. Voice-over thus becomes normalized as an integral component of short video communication.

Taken together, these dynamics explain why voice-over functions not merely as support for images but as a substitute for visual explanation. Short videos increasingly rely on telling rather than showing, on verbal articulation rather than visual inference. This transformation reshapes

narrative structure, audience engagement, and the cultural role of images within platform-based media.

The primary theoretical contribution of this study lies in reframing voice-over as an explanatory mechanism rather than a secondary or decorative feature. Audiovisual media theory has often assumed the dominance of the image, treating sound and voice as supplementary layers that enhance visual meaning. The analysis presented here challenges this assumption by demonstrating that, in short video environments, explanatory authority frequently shifts from image to voice.

By conceptualizing voice-over as a substitute for visual explanation, the paper highlights a redistribution of explanatory labor across sensory channels. Meaning is no longer primarily produced through visual continuity or montage but through spoken narration that organizes perception and interpretation. This perspective invites a reconsideration of how audiovisual meaning is structured under platform conditions marked by acceleration, saturation, and attention scarcity.

The study also contributes to discussions of narration and perception by emphasizing the role of voice in guiding meaning. Voice-over does not simply convey information but shapes how images are seen and understood. It directs attention, frames emotional responses, and narrows interpretive possibilities. In doing so, it alters the relationship between text and audience, positioning viewers as recipients of guided meaning rather than co-constructors of interpretation.

This theoretical framing helps bridge media theory and platform studies. It situates aesthetic and narrative changes within the material and cultural conditions of digital platforms. Voice-over emerges not as an isolated stylistic trend but as an adaptive response to structural constraints. This approach avoids normative judgments that either celebrate or condemn narration and instead emphasizes contextual analysis.

More broadly, the paper contributes to a growing body of work that questions visual dominance in contemporary media. It suggests that sound, voice, and language play increasingly central roles in shaping audiovisual experience. Recognizing this shift opens new avenues for analyzing power, authority, and

meaning in platform-based communication.

The arguments presented in this paper point toward several directions for future research. One important area concerns genre differences. Short videos encompass a wide range of content types, including education, entertainment, advertising, personal storytelling, and political commentary. The function and prominence of voice-over may vary across these genres. Comparative studies could examine how explanatory strategies differ between instructional videos and narrative content or between commercial and user-generated media.

Cultural context represents another significant dimension. Platform practices are globally distributed but locally adapted. Expectations surrounding narration, visual clarity, and interpretive guidance may differ across linguistic and cultural settings. Cross-cultural research could explore whether the substitution of voice-over for visual explanation manifests similarly in different media cultures or whether alternative strategies emerge under different norms of communication.

Audience reception also warrants closer examination. While this paper has focused on structural and theoretical analysis, empirical research could investigate how viewers perceive and respond to narrated versus visually driven short videos. Studies could examine how voice-over affects comprehension, memory, emotional engagement, and trust. Such research would deepen understanding of how guided meaning shapes audience experience over time.

Another important direction concerns visual literacy. As narration increasingly carries explanatory responsibility, questions arise about the long-term effects on viewers' ability to interpret images independently. Research could explore whether habitual exposure to narrated media influences interpretive skills, tolerance for ambiguity, or engagement with non-narrated visual forms. This line of inquiry would connect short video studies to broader debates about media education and literacy. Future work could also address creative resistance and experimentation. While voice-over is dominant, some creators deliberately minimize or avoid narration, relying on visual storytelling despite platform constraints. Examining these practices could illuminate alternative possibilities within short video culture and reveal the limits of narration-centered explanation.

Voice-over as a substitute for visual explanation is not a marginal phenomenon but a defining feature of short video media. It reflects deep transformations in how meaning is produced, distributed, and consumed under platform conditions. Understanding this shift requires attention to narrative structure, perceptual habits, and cultural norms. By foregrounding voice-over as an explanatory mechanism, this study offers a framework for analyzing contemporary audiovisual communication and sets the stage for further inquiry into the evolving relationship between sound, image, and meaning.

References

- Bordwell, D., & Thompson, K. (2019). *Film art: An introduction* (12th ed.). McGraw-Hill Education.
- Chion, M. (1994). *Audio-vision: Sound on screen* (C. Gorbman, Trans.). Columbia University Press.
- Chion, M. (1999). *The voice in cinema* (C. Gorbman, Trans.). Columbia University Press.
- Chovanec, J. (2020). Voice-over and presenter narration in TV documentaries. In J. Wildfeuer & J. Pflaeging (Eds.), *Media narrative theory* (pp. 155–170). Routledge.
- Kozloff, S. (1989). *Invisible storytellers: Voice-over narration in American fiction film*. University of California Press.
- Kruger, J.-L. (2010). Audio narration: Re-narrativising film. *Perspectives: Studies in Translatology*, 18(3), 231–249.
- Le Fèvre-Berthelot, A. (2013). Audio-visual: Disembodied voices in theory. *InMedia: The French Journal of Media Studies*, (3).
- Matamala, A. (2018). Voice-over: Practice, research and future prospects. In L. Pérez-González (Ed.), *The Routledge handbook of audiovisual translation* (pp. 465–480). Routledge.
- Pillai, S. E. (2015). The texture of interiority: Voiceover and visuals. *SVAC International Journal of Arts*, 2(1), 1–12.
- Smoodin, E. (1983). The image and the voice in the film with spoken narration. *Quarterly Review of Film Studies*, 8(1), 1–14.
- Vosmeer, M., Roth, C., & Koenitz, H. (2017). Who are you? Voice-over perspective in surround video. In G. Mark et al. (Eds.), *Interactive storytelling* (pp. 235–247). Springer.

Composite Social Actions: Question-Intoned Repeat-Formatted Repair Initiation in Mandarin Conversation

Jian Guan¹

¹ Ocean University of China, Shandong, China

Correspondence: Jian Guan, Ocean University of China, Shandong, China.

doi:10.63593/JLCS.2025.12.05

Abstract

Repair refers to the process of detecting and responding to problems with speaking, hearing or understanding in talk-in-interaction. In Mandarin conversation, repetition can function as a linguistic device for implementing other-initiated repair. This is primarily realized through three specific practices: question-intoned repetitions, repetitions suffixed with the final particle “a” and repetitions suffixed with the final particle “ma”. Focusing on the first type, this study conducts a fine-grained analysis of naturally occurring Mandarin conversations to demonstrate that such repetitions function not merely as a device for repair initiation, but as inherently composite social actions. Within a single turn, they can simultaneously accomplish additional interactional work, such as seeking clarification, highlighting unexpectedness, or negotiating responsibility. By examining their sequential environment and co-occurrence patterns, this study reveals how interactants utilize this repair form to manage both understanding and social relations within the micro-dynamics of conversation.

Keywords: conversation analysis, repeat-formatted repair initiation, clarification-seeking, unexpectedness, composite actions

1. Introduction

Repair is a central organizational feature of talk-in-interaction that ensures mutual understanding and sustains the progressivity of social interaction. It's an interactional mechanism for addressing problems in speaking, hearing, or understanding to maintain intersubjectivity and enable the progress of talk (Schegloff, 2007; Hayashi, Raymond & Sidnell, 2013). Crucially, repair is typically prompted by “inappropriateness” within the sequential context rather than by linguistic “incorrectness”. Other-repetition can serve as a form of repair

initiation. Such repetition acts as a “restricted” form of initiation, as it functions to delimit the scope of the trouble source rather than leaving it “open” (Drew, 1997). In Mandarin conversation, question-intoned repeat-formatted other repair-initiations, such as responding to “我昨天去了书店” (*I went to the bookstore yesterday*) with a partial repeat like “昨天?” (*Yesterday?*), constitute a pervasive and systematic practice.

Building on Rossi's (2018) concept of composite actions in talk, this study explores the multifaceted functions of question-intoned repeat-formatted other repair-initiations in

Chinese. While prior study has framed them primarily as understanding checks, expressions of disbelief, or resistance to medical advice, this study shows that their interactional functions are more varied and subtle than simply initiating repair.

Conversation Analysis (CA) is a rigorous, empirical methodology for studying the structures of social interaction (Heritage, 1984). Drawing on this methodology, this study examines how participants treat other-repetitions both as repair initiators and as vehicles for other social actions. Consequently, the study refines our understanding of this specific repair format in Mandarin and provides insights that can inform and extend cross-linguistic research on the interplay between repair organization and social action.

2. Previous Studies on Repair

Repair ensures mutual understanding among interlocutors at the expense of suspending conversational progress and sacrificing communicative continuity (Schegloff, 2007). Of particular relevance to this study are other-initiated repairs of the repetition type, delivered with questioning intonation and positioned in the responsive turn.

The execution of conversational repair is closely intertwined with the performance of other social actions. Drew (1997) observed that open-class repair initiators primarily serve to indicate the initiator's difficulty in understanding the trouble-source turn, thereby revealing that the topic shift in that turn was too abrupt or the action performed was inappropriate. Similarly, Kitzinger (2013) argues that repair is intricately linked to other interactional practices. Repair can be employed to redesign a granting action to meet the contextual demands of a specific sequence; to enhance the credibility of an information source; to fulfill the requirements of storytelling; or to manage epistemic authority, among other functions. Wu (2006) categorizes repetition-based repair initiators into those with interrogative intonation and those ending with the particle "a", positing that both types can perform the communicative function of conveying negative assessment. Yu and Wu (2022) found that patients' repeat-formatted repair initiators simultaneously accomplish three social actions: questioning, initiating repair, and covert resistance. They suggest that training doctors to effectively identify such resistance

actions embedded within repair initiators during medical consultations could enable intervention, thereby promoting smoother and more harmonious doctor-patient communication.

3. Research Methodology and Data Collection

Conversation Analysis (CA) is a rigorous, empirical methodology for studying the structures of social interaction (Heritage, 1984). Originating in the late 1960s through the pioneering work of sociologists Harvey Sacks, Emanuel Schegloff, and Gail Jefferson, it is built upon the foundational insight that everyday conversation is a highly structured, orderly domain of social life (Sacks, 1984).

The theoretical foundations of CA are deeply rooted in the sociological insights of Erving Goffman and Harold Garfinkel. Goffman's (1983) concept of the "interaction order" established social interaction as a legitimate institutional domain with its own moral structures, thereby providing CA with its core object of study. Garfinkel's (1967) ethnomethodology shifted the analytical focus onto the common-sense reasoning, or "ethnomethods", that members of society rely upon to create and maintain a shared sense of mutual understanding and social order. CA synthesizes these traditions, examining the interaction order not through external theories but from within, by analyzing the very procedures participants employ to render their conduct intelligible and organized.

At the heart of CA is the study of social action (Sacks, 1992). The central question is not what language means, but what it does. CA investigates how participants design their turns-at-talk to be recognizable as specific actions (action formation) and how recipients ascribe meaning to those actions through their subsequent responses (action ascription) (Schegloff, 2007). This process is inherently collaborative and sequential; actions are "context-shaped" and "context-renewing" (Heritage, 1984: 242), with each turn displaying its understanding of the prior one and shaping the next (Raymond & Robinson, 2024).

Methodologically, CA is distinguished by its commitment to naturalistic data (Pomerantz & Fehr, 1997). Analysts examine recordings of natural interactions, which are transcribed in minute detail using the system developed by Gail Jefferson (2004). This ensures that analysis is grounded in the observable orientations of the

participants themselves. The key analytical tools for this endeavor are turn-taking and sequence organization. The turn-taking system (Sacks, Schegloff, & Jefferson, 1974) explains how participants coordinate speaker change through Turn-Constructional Units and Transition-Relevance Places. Turn design examines the linguistic and non-linguistic resources selected to perform an action for a specific recipient (Drew, 2013). Ultimately, social action is understood through the interplay of turn design and its sequential position, a relationship encapsulated as “turn design + sequential position = social action”. In essence, CA provides a powerful framework for analyzing the infrastructure of human sociality, revealing the sophisticated order inherent in the flow of everyday conversation.

The data for this study were drawn from two primary sources: a portion was collected firsthand by the author, while the remainder was sourced from the DIG Mandarin Conversation corpus (DMC) (Yu et al., 2024). Following the methodological principles of Conversation Analysis, all audio recordings were transcribed using the Jefferson’s transcription system (Jefferson, 2004).

4. Analysis

4.1 Seeking Clarification or Specification

In response to lexical or referential trouble sources, Chinese speakers may employ question-intoned repeat-formatted other-initiated repair as a mechanism to restore and maintain intersubjectivity. The elicited response may take various forms, including: reformulating the original statement to enhance precision, disambiguating a vague or underspecified referent, or providing additional explanatory detail to resolve the perceived opacity of the prior talk.

Thus, such repetition functions not merely as an echo but as a targeted mechanism for achieving mutual understanding and discursive precision within the sequential flow of conversation. They simultaneously perform two tasks: first, they request confirmation that the repetition is accurate, and second, they implicitly and urgently demand further clarification or specification. The following excerpt is a conversation between friends Min (敏) and Yu (玉) about what they had for lunch.

[LWZ: 03:36-04:01]

01 敏: .hhh 那你中午吃的啥?

02 (0.2)

03 玉: 我点的那个(0.2)熔岩堡.

04 (0.4)

05 敏: 熔岩堡?

06 玉: .hh 对, 就汉堡王新出的那个, 双层牛肉,

07 中间夹了特辣岩浆芝士酱的.

08 敏: 哦:::, 那个, 那个 he>我看广告了<, 辣吧?

In line 1, Min initiates a topic by asking Yu about her lunch, positioning herself as unknowing (K-) regarding this specific personal experience. Yu, in line 3, provides an information-giving turn, naming the item she ate as “*Lava Burger*” (熔岩堡). The referent “*Lava Burger*” is produced without elaboration, treating it as a recognizable or self-explanatory item for the recipient. The 0.4-second silence (line 4), projects potential interactional trouble. In line 5, Min initiates repair with a partial repeat (*Lava Burger?*) delivered with a questioning intonation. This turn treats the prior reference as problematic or insufficiently specified, displaying Min’s momentary lack of recognition or epistemic access (K-status) regarding the named object, and solicits clarification or specification from Yu as the epistemic authority (K+) on the matter. Yu’s response in lines 6-7 elaborates and provides additional explanatory detail to resolve the perceived opacity of the prior talk. The detail provided upgrades Yu’s epistemic authority (K+) on this specific personal consumption experience.

In line 8, Min’s response exhibits a change-of-state token (“o:::”), marking the receipt of new information and a shift from K- to K+ status. The repetition (“that”) demonstrate online processing and recognition, linking the now-specified item to a prior encountered advertisement. The subsequent tag question seeks confirmation on an inferred property (spiciness), which, while building on the newly established common ground, also displays Min’s continued K-status regarding the subjective experience of taste, subtly inviting further elaboration from Yu.

The sequence thus demonstrates a trajectory from problematic reference to resolution and recognition, followed by a topic extension grounded in the newly clarified referent. It further demonstrates how question-intoned repeat-formatted repair initiations serve to

reformulate the original utterance with enhanced precision and provide added explanatory details to address perceived opacity in the prior talk. This point is also demonstrated in the following excerpt.

[GXJ: 07:28-07:44]

01 嘉：我接下来两周(.)要当查操员了。

02 (0.2)

03 琴：查操员？

04 (0.3)

05 嘉：呃::对，就是-就是查大家做眼保健操的。

06 琴：hehehe 我还以为什么操呢。

In line 1, Jia produces an informing turn, introducing a new referent “*inspector*” (查操员) into the conversation. By announcing his upcoming role, Jia assumes the epistemic authority (K+) regarding this personal appointment, while treating the term as potentially recognizable or at least interpretable within the ongoing interactional context. Qin initiates repair by repeating the noun phrase with a questioning intonation. This turn treats the prior reference as problematic and solicits clarification from Jia. In line 5, following the utterance “*呃::*” (uh::), Jia offers confirmation (yes “对”) and then proceeds to provide a definitional elaboration, specifying the scope of the role (*checking everyone doing eye exercises*). This turn orients to Qin’s repair initiation as stemming from a referential gap and systematically addresses it, thereby reinforcing his K+ status on the matter.

4.2 Highlighting Unexpectedness

Question-intoned repeat-formatted repair initiations can also function as a marker of unexpectedness, signaling the speaker’s affective, epistemic, or stance toward prior talk. There are two subtly distinct categories in the classification presented below. First, speaker may repeat the interlocutor’s prior utterance to highlight its deviation from his/her expectation, thereby enacting a display of surprise. This action makes the interlocutor accountable for explaining or reconciling the discrepancy between expectation and reality. Second, the repetition can serve to implicitly or explicitly challenge the truthfulness, appropriateness of the prior talk.

4.2.1 Surprise-Oriented Repetition

Surprise in interaction serves as a social display, often conveyed through repetition and emphasis on unexpected elements to show affective

involvement. This action is oriented toward building a shared stance among participants, and can occur regardless of the positive or negative nature of the unexpected event. It thus functions as a positively valenced move that promotes alignment and affiliation.

[LYC: 01:32-01:48]

01 张：我跟你讲，我昨天做那道三不粘，

02 搅了整整四十分钟才成功。

03 (0.5)

04 王：↑四十分钟？

05 (.)

06 天哪，[¥你这手劲也太吓人¥了。

08 张：[hehehe¥不然根本成型不了¥。

In line 1, Zhang begins with the preface “我跟你讲” (*I tell you*), which frames the upcoming utterance as an informing of an A-event, something the speaker has primary epistemic authority over. Then Zhang uses the demonstrative “那” (that), implying the dish is mutually known or inferable from prior context (perhaps they had discussed cooking before). The key detail “四十分钟” (forty minutes) is prosodically stressed, marking it as a noteworthy and potentially surprising element. The 0.5-second pause allows the recipient to process the information, particularly the emphasized duration “forty minutes”.

In line 4, Wang produces a turn consisting solely of a repeat of the highlighted element: “↑四十分钟？” (↑*Forty minutes?*). The upward arrow indicates a sharp rise in pitch on the first syllable “四” (sì), making the repeat prosodically marked. This format is a classic other-initiated repair via repeat, but here it does not primarily target hearing or understanding (Zhang has clearly articulated the number). Instead, the repeat serves as a surprise or unexpectedness marker. Wang then explicitly expresses unexpectedness with “天哪” (*Oh my god*), an exclamation that upgrades the affective display from the implicit surprise in line 4. This is followed by an assessment: “你这手劲也太吓人¥了” (*Your arm strength is really frightening*). The exaggerated assessment provides a justification for the surprise—stirring for 40 minutes is physically extraordinary.

In line 8, Zhang’s laughter begins in overlap with the end of Wang’s turn, showing affective alignment. Zhang then adds an account: “不然根本成型不了¥” (*Otherwise it wouldn’t form at all*).

This justifies the extreme effort and treats it as necessary, not just voluntary, thereby modestly downplaying the remarkable achievement while still accepting the positive assessment.

4.2.2 Challenge-Oriented Repetition

A problem of acceptability in interaction often relates to perceived issues of truthfulness or accuracy in prior talk. The speaker can employ this type of repair to question the acceptability of the prior turn and express explicit dissatisfaction, thereby enacting a “negatively valenced” action that signals a belief that the previous statement was incorrect and requires correction.

[JY: 02:31-02:46]

- 01 风: .hs-,这个模块的开发,我估计大概(.)需要 5 万左右.
02 (0.4)
03 婷: 五万?
04 (0.2)
05 上次类似功能不是只花了两万.
06 风: 和上次不一样(.)结构完全不一样.

In line 1, Feng provides a cost estimate “5 万左右” (*around 50,000*) for a module development project. The 0.4-second silence allows Ting to process the figure and potentially compare it with prior knowledge. In line 3, Ting initiates a question-intoned repeat-formatted repair with “五万?”, which questions the accuracy or appropriateness of the figure and mark the unexpectedness of the speaker. In line 5, Ting explicitly contrasts this with past experience (“上次...只花了两万” *Last time... it was just twenty thousand*). The negative-interrogative construction (“不是...吗”) presupposes shared knowledge and invites confirmation of an apparent discrepancy. This formulation highlights unexpectedness by juxtaposing the past (“两万”) and present (“五万”) figures, thereby implying an inconsistency that requires explanation. She then offers a conceptual summary (“结构完全不一样” *The structures differ completely*), explicitly framing the discrepancy as structural rather than merely numerical.

4.3 Negotiating Responsibility Implicitly

Everyday conversation frequently involves narratives about personal and social problems. This study finds that this type of repair is frequently observed within problem-reporting sequences. Specifically, when a preceding turn involves problem reporting that implies “attributing responsibility to the other party”,

recipients frequently employ this type of repair as a form of implicit negotiation to responsibility attribution. Subsequently, recipients may utilize some practice to challenge the legitimacy of the reported problem, thereby attempting to negate the premises underlying its presentation.

[QYC: 06:32-07:21]

- 01 小白: 我之前给您说那个::空调漏水的情况,
02 >你啥时候能过来看看,现在还在漏着呢
04 师傅: 还漏?
05 (0.3)
06 小白: 还漏;,
07 师傅: [你那个加.你你.你那个是加到 15 米管子吧?
08 小白: [我这个是不是()
09 (0.7)
10 小白: 啊对
11 (1.0)
12 师傅: 嗯:;<搁 15 米管子的话>,按说应-
13 (0.2)
14 那-那段时间不应该漏.

In this excerpt, the temporal term “之前” (*before*) in Xiao Bai’s turn in line 1 establishes a past reference point. By contrast, the use of “现在” (*now*) in line 2—“现在还漏着呢” (*it’s still leaking now*)—frames the air conditioner leakage as a persistent issue that has continued from the past into the present without being resolved. The aspectual ending “着呢,” which marks an ongoing state, further underscores the current and unresolved nature of the problem. Xiao Bai’s request in line 2, “啥时候能过来看看” (*when can you come and take a look*), is designed with the precondition that a visit is possible (“能来看” – *can come*). By specifically inquiring about timing (“啥时候” – *when*), the turn displays Xiao Bai’s relatively high entitlement to make this request.

The way Xiao Bai presents the problem, emphasizing both its duration and persistence, serves to implicitly assign a degree of responsibility to the technician for addressing the leakage. This problem reporting also projects an expectation that the technician should take charge of arranging a repair or other relevant actions. In line 4, rather than accepting or rejecting this projected responsibility, the technician redirects the interaction by initiating a clarification: “还漏?” (*still leaking?*). This repair initiation shifts the conversational focus from “when” the repair will occur—a matter directly tied to responsibility—to “whether” the problem indeed exists, thereby reframing the issue in terms of epistemic alignment and

problem definition.

Xiao Bai confirms in line 6 with “还漏” (*still leaking*). The technician then pursues further background in line 7 by checking factual details: “你那个是加到 15 米管子吧” (*yours had the pipe extended to 15 meters, right?*). After Xiao Bai’s affirmative response in line 10 (“啊对” – *ah yes*), the technician acknowledges with “嗯” (*mm*) in line 12, and then offers a counter-assessment: “那段时间不应该漏” (*it shouldn’t be leaking during that period*). By contesting the very premise of the problem, he challenges the factual basis of Xiao Bai’s earlier complaint, thereby undoing the presupposition that necessitates a repair visit.

The following excerpt illustrates how a participant (Li) designs a problem presentation to establish the interlocutor’s (Chen’s) accountability for a problem, and how the interlocutor manages this implicit attribution through repair initiation and reframing.

[ZNX 00:05-00:39]

- 01 李: 呃,我-昨天在您这就是买了盆儿,
02 橡-橡皮树。好像,
03 .hh[>然后<您跟我说先慢慢浇水,
04 陈: [嗯。
05 李: 然后等[那个-花盆儿,就是-
06 陈: [啊对。
07 李: 它往下渗了,就不用浇了,
08 >但是<我今天给它浇水的时候,
09 它就是(0.2)不往下渗。
10 那个土就是:(.)它一直不渗水。
11 (0.5)
12 陈: <不渗水?>
13 (0.2)
14 呃::(.)[你>是不-<你是不没浇透啊,
15 李: [对。
16 陈: 你-你得>浇好几遍儿<[(0.2)好几圈儿。
17 李: [我-我
18 陈: ang:(.)就是。你。那个圈儿,
19 你得转着-转着<慢点>浇啊,
20 >就是<差不多每次都得浇四五圈儿,
21 (.)而且-你得浇两遍,

In lines 1 and 3, Li’s turn design employs the formulations “在您这” (*from you*) and “您跟我说” (*you told me*), explicitly grounding the source of both the rubber tree and the watering instructions in Chen. By framing Chen as the epistemic authority for this information, Li constructs a telling in which Chen is interactionally made accountable for the

subsequent problem reporting (Heritage, 2012). Chen’s minimal response “嗯” (*mm*) in line 4 acts as a continuer, passing the floor back to Li and encouraging further telling. In line 6, Liu’s “啊对” (*ah right*) provides a confirmatory receipt of Li’s prior description. Li continues detailing the watering practice in line 7. The turn-initial “但是” (*but*) in line 8 projects an upcoming contrast. This is realized in line 9, where Li produces the highlighted trouble description “不往下渗” (*not draining down*), which is then reiterated, constituting a clear problem presentation.

The 0.5-second gap in line 11 may foreshadow a dispreferred or non-aligning response. In line 12, Chen initiates repair with a questioning repeat, “不渗水?” (*not draining?*). This other-initiated repair does not address the accountability projected by Li but rather targets the acceptability and factual grounding of the problem description itself. It opens a verification sequence, temporarily shifting the interactional agenda from accountability to epistemic alignment. Following a brief 0.2-second gap (line 13), Chen does not await confirmation but instead produces a diagnostic reformulation in line 14: “你是不没浇透啊” (*could it be that you didn’t water it thoroughly enough?*). It undoes the presupposition underlying Li’s problem (that “non-draining” is an objective fault) and reframes the issue from one of advice failure to one of inadequate implementation. Chen then elaborates on the correct method across lines 16–21. These turns constitute a re-specification and reinforcement of the original advice. Their interactional function is to recalibrate accountability: responsibility for the problem is systematically shifted from the adequacy of the advice to the adequacy of Li’s actions, accomplishing a subtle but significant frame shift from a “problem-with-guidance” to a “problem-with-execution”.

5. Conclusion

This study demonstrates that question-intoned repeat-formatted other-initiated repair in Mandarin conversation constitutes a composite social action. It functions not merely as a mechanism for resolving problems of speaking, hearing, or understanding, but simultaneously accomplishes additional interactional work such as seeking clarification, marking unexpectedness (through displays of surprise or challenge), or implicitly negotiating responsibility attribution. The analysis reveals how participants deploy this specific turn-design within its sequential

environment to manage both intersubjectivity and social relations. These findings refine our understanding of repair organization in Mandarin by foregrounding its intrinsic multifunctionality and embodied social logic, thereby contributing to cross-linguistic research on the interplay between sequential structure and social action. Future research could productively examine the variation and constraints of this practice across different institutional settings.

References

- Drew, P. (1997). 'Open' class repair initiation in response to sequential sources of troubles in conversation. *Journal of Pragmatics*, 28(1), 69–101.
- Drew, P. (2013). Conversation analysis and social action. *Journal of Foreign Languages*, 37(3), 2–19.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Goffman, E. (1983). The interaction order. *American Sociological Review*, 48(1), 1–17.
- Hayashi, M., G. Raymond & J. Sidnell. (2013). *Conversational Repair and Human Understanding*. Cambridge: Cambridge University Press.
- Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.
- Jefferson, G. (1984). Transcription notation. In *Structure of Social Action: Studies in Conversation Analysis*, ed. by J. M. Atkinson & J. Heritage, ix. Cambridge: Cambridge University Press.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In *Conversation Analysis: Studies from the First Generation*, ed. by G. H. Lerner, 13–31. Amsterdam: John Benjamins.
- Kitzinger, C. (2013). Repair. In *The Handbook of Conversation Analysis*, ed. by Jack Sidnell & Tanya Stivers, 229–256. Chichester: Wiley-Blackwell.
- Pomerantz, A., & B. J. Fehr. (1997). Conversation analysis: an approach to the study of social action as sense making practices. In *Discourse as Social Interaction*, ed. by T. A. Van Dijk, 64–91. London: Sage.
- Raymond, C. W., R. Clift, K. H. Kendrick, & J. D. Robinson. (2024). Methods in conversation analysis. In *The Cambridge Handbook of Methods in Conversation Analysis*, ed. by J. D. Robinson, R. Clift, K. H. Kendrick, & C. W. Raymond, 3–48. Cambridge: Cambridge University Press.
- Rossi, G. (2018). Composite social actions: The use of factual declaratives in everyday interaction. *Research on Language and Social Interaction*, 51(4), 379–397.
- Sacks, H. (1984). Notes on methodology. In *Structures of Social Action*, ed. by J. M. Atkinson & J. Heritage, 21–27. Cambridge: Cambridge University Press.
- Sacks, H. (1992). *Lectures on Conversation*. Oxford: Blackwell.
- Sacks, H., E. A. Schegloff & G. Jefferson. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Schegloff, E. A. (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Wu, R. J. R. (2006). Initiating repair and beyond: The use of two repeat formatted repair initiation in Mandarin conversation. *Discourse Processes*, 41(1), 67–109.
- Yu, G., & Y. Wu. (2022). A conversational analysis of resistance to medical recommendations. *Modern Foreign Languages*, 45(1), 17–28.
- Yu, G., Y. Wu, P. Drew & C. W. Raymond. (2024). The DIG Mandarin Conversations (DMC) Corpus: Mundane phone calls in Mandarin Chinese as resources for research and teaching. *Chinese Language and Discourse*, 15(1), 105–141.