

# Wikipedia Entries: Collaborative Editing of Bias under Linguistic Hegemony?

Qing Xia

*School of Economics and Management, Communication University of China, Beijing, 100024, China  
xiaqing0330@gmail.com*

**Abstract:** *Based on the content analysis of the multilingual articles of selected Wikipedia entries, this study examines the Wiki-content editing processes from the aspects of topic, framing, discourse, reference preferences and the relevance among languages, questioning with evidence that Wiki's editing processes do not implement its editorial principles of objectivity and neutrality: Western discourse frameworks have taken root in the civilian global digital community; and there exists a biased "consensus" multilingually though the consistency of Wikipedia entries in various languages is currently weak and relatively independent. The intensified trend of linguistic hegemony highlights the necessity of better utilization of the Wiki-like worldwide platforms, and the importance of China to plan up-to-date international communication strategy and execution path with wisdom.*

**Keywords:** *Wikipedia; Linguistic Hegemony; Discourse Framework; National Image; International Communication Strategy*

## 1. Introduction

Under the multiple overlapping of Sino-US trade disputes and the Covid-19 epidemic, the trend of anti-globalization has intensified, China is facing a more severe international environment with rumors and smears emerged, interfering the pace of development. Eurobarometer survey in 2017 showed that people in 21 of the 28 EU member states held a negative view of China. The Pew Research Center survey in 2020 showed that the image of China among the people in 14 developed countries was 73% negative to 24% positive <sup>[1]</sup>, and the proportion of people of a negative view of China has risen to the highest point since 2002 in half of the countries. Let alone a Gallup survey showed that Americans' disapproval of China reached the highest point (67%) since 1979.

As an important part of national soft power <sup>[2]</sup>, national image (in international communication) has great impact on the political and economic relations between countries <sup>[3]</sup>. On May 31<sup>st</sup> 2021, the Central Committee of the Communist Party of China emphasized the necessity of "strengthening the construction of international communication capabilities" to "form an international discourse power that matches China's comprehensive national strength and international status". However, the Western media has formed a hegemonic power of discourse long, by setting the international reporting framework, constructing the media image, and shaping the audience's cognition, to dominate the international public opinion. China lacks depth of global influence still though its resources and global exposure increase much <sup>[4]</sup>. In terms of influence, Chinese media hold far less than that of traditional Western media such as The New York Times, CNN News, BBC News and other multi-channel competitors such as Yahoo News, MSN News, etc. The weakness in international communication, leads to the dilemma that China being unable to correct the misunderstanding of other media and failing to guide the audience with its own voice. And the predicament has seriously affected the interests of all aspects of China, holding back its further rise, arrives to a point that urges to seek a way out.

As the Internet penetration rate in various countries increases year by year, Internet users' preferences for accessing content online change, as the 2006 Eurobarometer report stated, to online information channels (including social media), differing in the choice of frequent contacted media forms and brands; they no longer accustom to obtaining information from a single source. Following the trend and with the help of Google's search algorithm recommendation preference, Wikipedia has become the world's most traffic platform of information searching and sharing, to be not only more influential than traditional media, but also an important tool and primary channel for journalists, scholars, politicians and public to search for information and understand news, as well a habitual news background query database and a citation source for a large number of media <sup>[5]</sup>. Wikipedia has covered

314 language versions as of March 2022, with a total of more than 55 million articles, 490 million independent users, and 18 billion page views per month; ranked 13<sup>th</sup> and 9<sup>th</sup> of Alexa and SimilarWeb in the global website traffic rankings.

Wikipedia's user visits, influence, audience breadth, and importance in international communication are beyond doubt, and as an important window for global netizens to understand events and things related to China, scholars generally acknowledged its importance, necessity and suitability in the practice of China's international communication [6]. To a certain extent, the content of Wikipedia entries can reflect the global netizens' acculturation understanding of China's image after exposure to increasingly negative China-related reports in recent years. This kind of "consensus" image rather than the media image is more beneficial for China to formulate accurate and effective international communication strategies. In addition, understanding the production and flow patterns of information in different languages in this multilingual Internet community would be conducive to China's international communication against linguistic hegemony, so that measures can be tailored to specific conditions and targeted.

## 2. Content Analysis

Concerned that China's national image largely equates to the image of Chinese government and the image of the Communist Party of China, with the characteristics of "three images in one", this study intends to analyze the content of Wikipedia's multilingual articles of the Communist Party of China based on the Framing Theory and Agenda Setting Theory. By summarizing the characteristics of framework, discourse, agenda, similarity with the content and attitude of traditional media reports, and degree of linguistic hegemony, suggestions on international communication in terms of strategy and specific path are proposed then.

The Wikipedia has a total of 90 language pages of the Communist Party of China entries, with more than 12 million visits (excluding redirects). Among them, the Portuguese "Partido Comunista da China" and the Norwegian "Kinas kommunistparti" were rated content quality as excellent. This study selects entries in English, Chinese, Spanish, Portuguese, Japanese, Russian, French, German, Norwegian, Italian etc. (content version of August 1<sup>st</sup> 2021) according to their ranking performance of page views, language users, neutrality disputes and the quality of articles.

The top 5 entries which contributed the most page views are of Chinese (28.4% of total), English (14.1%), Japanese (9.7%), Spanish (6.9%), and Russian (5.6%), accounting for nearly two thirds of the total traffic of all 90 languages. Over 90% users access Wikipedia's Communist Party of China entries through mobile networks and desktops. There is no obvious periodic change in the number of page views of various languages entries. Most visits increased in June and July of 2021, not only Chinese entry "中国共产党" ranked 182<sup>nd</sup> in July among all Chinese entries on Wikipedia, but also of the various languages as Chinese, English, Japanese, Russian, Spanish, French, Italian, Dutch, Arabic, Vietnamese etc., the single-day user visits reached the highest on July 1<sup>st</sup> 2021 which was the 100<sup>th</sup> anniversary of the founding of the Communist Party of China, among them Chinese one had the maximum 38,166 as a certain effect of China's enhancing international communication.

### 2.1. Abstracts of Entries

As the overview paragraph of each Wikipedia entry, the abstract part is located in the most prominent position below the theme of the article, summarizing the entire content and setting the tone.

The 1413-word abstract of Chinese entry "中国共产党" consists of three paragraphs, which includes a brief history of party development, leaders, organizational structure, number of party members, constitution and so on. It has 11 citations of sources all in simplified Chinese, and about three quarters of them are Chinese official media websites like Xinhua and People's Daily. The 545-word abstract of English entry "Chinese Communist Party" consists of four paragraphs, which includes a brief history of party development, organizational structure, leaders, guiding ideology, inter-party relations, number of party members and so on. It has 1 note and 6 citations of sources half in simplified Chinese and half in English, averagely from Chinese official media, English books published in China, and books published outside China.

The English entry quotes a larger proportion of materials originally in English language in its abstract part, but still uses Chinese official media as the main source. The contents covered in both Chinese and English entries are roughly the same, only differ in orders and details, for instance,

Chinese abstract devotes much length to describe the historical developments of the party; while the English one adds contents such as China's eight democratic parties, the party-army relationship, and relationship with Kuomintang and other parties around the world.

In terms of expression, they differ in word choice and details in length on the one hand. In most cases, English entry has more detailed expression. For instance, when expressing the party leaders, the English entry has an added word "elected"; in describing the organizational structure of the party, the Chinese entry simply states that it mainly based on democratic centralism, advocating those policies that have gained consensus should be jointly maintained, however the English one adds contents like "democratic centralism was proposed by Lenin" and "(it) requires democratic and open discussion of policies" to this description.

On the other hand, they each have implicitly biased statements, some even disregarding the editing principles (neutrality of opinion, availability of evidence, and unoriginal research) set by the Wikipedia community. For example, when expressing that the National Congress is held every five years, the English entry prefaces the sentence with the word "theoretically", which means "that could possibly exist, happen or be true, although this is unlikely" according to the Oxford Advanced Learner's English-Chinese Dictionary, implying a negative connotation about the possibility of the existence of the matter described, dissolves the sense of objectivity of the fact.

And, both entries go to great lengths to explain how Chinese current economic system develops the capitalist mode of production, even with expression as "official explanation", but without any cited references. Moreover, both lack the neutrality in words and phrases, using "(the party) has always proclaimed its ideals and goals as" "(it) sees itself as" "officially (it) is committed to" etc. to distinguish between the source of information and the editor's attitude toward the cited material; as with the choice of the word "officially", which means "according to information that has been told to the public but that may not be true" in dictionary, to imply that the cited information is published but authenticity in doubt, provoking readers to question its objectivity.

The abstracts of two rated-excellent entries are basically the same as the English one in terms of the structure and expression. The Portuguese 539-word abstract has two citations from western publications; the Norwegian 448-word abstract has none. As for the content and reference of the abstracts of other languages, the Spanish 259-word abstract has one citation from xinhuanet.com report in simplified Chinese; the French 135-word abstract has three citations in English, two from china.org.cn and one from The Times of India; the Russian 130-word abstract has three citations, two from Chinese official media and one from book published in Russia; the Japanese 249-word abstract has five citations mainly in Japanese, but one in Korean and one in Hindi. Their content and orders are varied, some with expressions violating the principle of objectivity and neutrality. For instance, the Japanese abstract begins with a statement "事実上の一党独裁" to define the party as a dictatorship with negative connotations such as oppressive control.

In summary, the multi-language abstracts of the Communist Party of China entries are in one same zone of topics and contents, only differing in the selection of details and logic order, which indicates that the editing process of later entries should have referred to some extent to some earlier entry information; however, their contents laid in various framework with different quotes and attitudes of wording, suggests that they have written independently, not formed of translation.

## 2.2. Content of Entries

The table of contents is located beside the abstract, and framed by a prominent text box. It has links of each heading, with the convenience of jump directly to the corresponding section, which reflects the editing thoughts of content compilation structure.

Apart from the usual modules set by Wikipedia such as "See also" "Notes" "References" and "External links", the Chinese entry has seven chapters, with six of the contents of history, symbols, ideology, governance, organization, and party relations, corresponding to the chapters of the English entry (except that the last chapter of the English one lists the votes and ranking data of party leaders' electoral history), and one 2000-word last chapter to place the related comments. And the three parts of comments involve politics and economy, corruption and speech control: they use the perspective of the Kuomintang, criticizing that China's rapid development after the reform and opening up costs much on the environment, culture and demographics; question the disciplinary actions inside the party, expressing that it reflects a systemic problem of the party; and depreciate the anti-corruption measures as well as the efforts and the facts of improvement on freedom of speech and democracy, for the reason

of lacking the effective recognition of some international organizations and foreign journals, with the only focus on criticism of publications and offshore websites censorship.

The comments chapter seems to comply with the principle of objectivity which Western journalism and communication appreciates, texting roughly balanced quantity of Chinese and foreign media sources (covers international organizations, governments, foreign published books, commercial media, Chinese official media etc., with most frequently cited sources as CPC website, BBC News, the New York Times and People's Daily Online), with positive and negative information in alternating structure. However, this chapter of Chinese entry indeed uses the agenda setting methods to influence the readers' perceptions of the importance and the logic of the events and opinions mentioned in the text as to "determine how the public relates different info"<sup>[7]</sup>, specifically as: positive and neutral statements are mostly placed unimportantly, used to introduce the weighted and evident criticisms; paragraphs generally emphasize negative tendencies carried in quoted texts from Taiwan media and foreign media, mainly locating focuses on the unfavorable judgments, to associate the party's image with dictatorship and standing oppositely to the people. The party/government image formed from the unobjective descriptions has been received by global Internet users, though the original website pages of three of the negative quotes are no longer accessible.

Another example of the agenda setting is that the English entry has listed a separate section on "Funding" in the chapter "Organization", claiming that "the funding of all CCP organizations mainly comes from state fiscal revenue" with a citation of the data summed up by editors themselves from a concise fiscal revenue and expenditure report released in the local government website of Mengmao in Yunnan Province in October 2016, then comes to a conclusion that the party costed more than 1.66 million yuan that year, which amounts to 5.63% of fiscal revenue for its own operation. Although there is no further explanation of the figure, the source of funds has been set as a topic related, which would cause misunderstanding and doubts the reasonableness of the party operation when the message read by netizens. And, as a comparison to note, Russian entry mentions the funding subject as well, yet chooses to be more objective: only quoted from a Russian media report to state that the CPC members monthly pay some percentage of their salary for the party fee as "юаней (около 456 долларов) в месяц уплачивается 0.5%, а с заработка более 3 тыс. юаней в месяц — 2%".

The content of the entries also spreads their favorable power relations and ideological contexts through framing, highlighting the subjective interpretation of meaning, attributive inferences, moral assessments and biases of the editors themselves in producing<sup>[8]</sup>. The "human rights" framework and the "democracy and freedom" framework are found throughout the articles of the entry, representing to some extent the dominant consciousness and the way of interpreting meaning in Western society. The English entry has a subsection named "Artificial Intelligence", citing a PBS television programme Frontline which states that artificial intelligence "has been aggressively developed...by its one-party state-enabled consolidation of power" to monitor citizens extensively; the subsection "Controversias" of Portuguese entry, the "Controverisie" of Italian entry, the "Contrôle idéologique", "Religion" and "Féminisation" of French entry, are full of cases forming negative image of China by using "human rights" framework; the Norwegian and German entries even create a separate section on "human rights violations" to locate the hundreds of words of criticisms.

In summary, the content themes are broadly similar across the language articles of the entry, with the following strategies being used more often in the compilation process: Firstly, direct agenda setting to influence public perception. Setting up sections with critical headings, to internalize human rights violations, corruption, speech control and so on as parts of the party's presentation, with negative content from foreign media reports even cases of unquoted sources and imaginary criticism. Secondly, the use of Western frameworks to guide the public. The so-called "human rights" framework and "democracy and freedom" framework are everywhere. Not only "Shuanggui" was introduced as "to extract confessions by cigarette burns, beatings and simulated drowning" with no regard for human rights, but also "artificial intelligence" equates with "monitoring the people", let alone to question the party system with its population policy, ethnic policy, information censorship, historical events, illegal organizations, etc., fictionalizing the antagonistic relationship between the government, political parties and the people with some non-sourced critical messages. Thirdly, the definition such as "dictatorship", wording modifying the truth to imply it may fake, and ordering of the content are tactfully used to construct a negative image of the party. The compilation of these articles is not in line with Wikipedia's stated principle of objectivity and neutrality, nor is the information without a source in the texting process in line with Wikipedia's established editing principles. Among them, the French entry is the most disregarded in terms of objectivity and neutrality, with 24.8% of the entry devoted to corruption cases addressing they were imprisoned with "luxe (luxury)" under a different and "confidentialité

(confidential)” system and process, even citing three French media reports praise “un rejet (rejection) total” attitude towards the party even the “le renversement (overthrow) du parti”. Although some editors have questioned similarly in the “discussion” pages of some entries, due to the very complex structure and levels of the Wikipedia community and editorial activity, the discourses are constantly being fought over, negotiated, and compromised <sup>[9]</sup>, remaining a great deal of bias then in the updated version of at least the research texts of this study.

### 2.3. Sources of Information

As for the information sources, the Chinese entry has 251 notes mostly in English and simplified Chinese, references of books published in the United States and Britain mostly since the year 2000, and 2 external links of Chinese websites. The English entry has 234 notes mostly in English, references 86.9% of books and 13.1% of academic articles, and two external links of China's official media websites. Among the other language entries, some such as Portuguese and Norwegian entries cite mostly (90% or so) in English; some cite mostly in English and their own language, such as French (100%) and Spanish (70% of notes, all references and external links) entries; some such as Japanese and Russian entries cite mostly in their own language (90% or so). The English entry has two quotations, which both appear in the Chinese entry: one from General Secretary Xi in a 2014 New York Times article, the other from a book named Historical Dictionary of the Chinese Communist Party by Sullivan Lawrence R; accurate and consistent when this study compares the Chinese translation with the English representation. The sources of China's official media are relatively lack of use in these entries, accounting for nearly half of the Chinese entry, about a quarter of the English and Spanish entries, and only one tenth of the entries in other languages. The entries of different languages cite inconsistent information, but all chose mainly English books as the references and quotations in English. The links to external websites prefer in their own languages, and some links to China's official media websites for it indeed is a China-related entry.

### 3. Conclusion and Discussion

On the whole, the content and themes of the Communist Party of China entries are basically the same in all languages; however, they differ in text logic, focus, notes and references, to show a certain independence. In general, the consistency between languages is weak, though some languages refer parts of English entry content in their compilation process, showing a certain degree of convergence; but there is no evidence found to show the editing as a simple translation. For China's international communication strategy and path in the new era, the findings have implications as follows:

The image of China, Chinese government, and the Communist Party of China - the “three images in one” characteristics in international communication have already existed in public opinion ecology and the perceptions of international audiences. A survey of 13 European countries in 2020 by Central European Institute of Asian Studies showed European unfavored public opinion on China, with impressions of Covid-19, large population, communism, dictatorship, authoritarian, totalitarian, absence of human rights, oppression, control, surveillance, exploitative, limited freedom, pollution, corruption, animal cruelty, forgery, low-quality products, poverty, child labor and so on. Most of the keywords above refer to the “three images”, and are highly consistent with the negative labels given by foreign media reports over the years and now even appearing in the texts of civilian digital community such as Wikipedia. This shows that the frameworks and issues set by the foreign media have become the prejudiced “consensus” of the global netizens, constantly circulating and deepening in the production and high-speed dissemination of information in the Internet era, further exacerbating the serious situation of misunderstanding and misinterpretation of China in international communication; while the academic consensus “three images in one” of Chinese researchers is not clearly reflected yet in the top-level design of China's international communication strategy. How to make an integrated consideration on this issue, in order to resist the Western-centric influence that has been disseminated continuously for a long time to take roots in the global netizens' consensus, is important and strategically urgent to be addressed.

Although the Wikipedia entries of different languages have relatively independent and no direct translation-used content currently, the tendency towards increased linguistic hegemony in the production and dissemination of Internet information may be inevitable in the long run. Some German, French, Spanish, Portuguese, and Italian entries have shown an over-reliance on the content of English Entries in recent years; entries on global issues have adopted a similar discursive framework <sup>[10]</sup>.

English hegemony affects global knowledge production, especially that of disadvantaged culture in the world; languages of small countries and the minority would become hundreds of copies of English content, carrying its ideology and values to each corner worldwide, and the bias in English entries would spread more widely, making the negative effects of biased information more difficult to purge. And this has happened, even accelerated since the collaboration between Google Translate and Wikipedia in Jan 2019. As Wikipedia gradually becomes an important channel of knowledge for people in almost all countries, it is only a matter of time before the global digital community follows the traditional media infiltrated by cultural hegemony, turning into another “lost territory” [11]. The international communication ecology will further lose its diversity and inclusiveness: reporting a narrower range of issues with a unified world view; leaving little space but serious pressure of communication for non-mainstream languages. Therefore, effective measures should be taken soon to counteract the effects of linguistic hegemony, including: clarifying the flow, pathways and mechanisms of international communication; monitoring, evaluating and predicting the extent of linguistic hegemony in real time; determining whether China’s foreign communication strategy should be targeted precisely at the core, or extensively cover each break to analyze them specifically and wisely.

As for the specific measures, firstly, the low percentage of Chinese official media sources in the texts of the global Internet community reflects the weakness of materials quantity and linguistic richness in China producing and accumulating international reports, in particular the textual materials which have multiple dissemination advantages in terms of preservation and citation. Providing non-machine translated multilingual information to foreign countries is also crucial. Not only Chinese news reports and official websites, but also Chinese books and periodicals need to be multilingualized, providing multilingual titles, keywords and abstracts for the convenience of international search. Secondly, it is not always wise to avoid ideological controversies and focus more on softer issues, for the country image is a comprehensive and holistic perception, with its core benchmark always lies in political factors, which must be tackled as the key to strengthen foreign audiences’ positive perceptions of China along the international communication path. Last but not least, to overcome the long-lived habitual Western framework, changing a battlefield by constructing a separate framework of own possibly be a way, confronting to a certain extent in case of misunderstanding, hostility and rejection while promoting cooperation may be another. In sum, recognizing the importance of global digital platforms such as Wikipedia is urgent for China’s international communication, however, not being under any illusions about the neutrality and objectivity of these platforms while learning to use their huge volume and influence to directly reach a large international audience, then consider the steps to improve global impressions and perceptions of China in a reality full of relatively hidden biases.

## References

- [1] Silver L, Devlin K, Huang C. (2020) *Unfavorable Views of China Reach Historic Highs in Many Countries*. Pew Research Center, 6.
- [2] Boulding K E. (1959) *National Images and International Systems*. *Journal of Conflict Resolution*, 3(2), 120-131.
- [3] Zhi Y, Bao D, Luo C. (2017) *Economic Value of Country Image: Evidence from International Trade and Implications for China*. *China & World Economy*, 25(3), 87-111.
- [4] Shambaugh D L. (2013) *China Goes Global: The Partial Power*. Oxford: Oxford University Press, 16-18.
- [5] Lih A. (2004) *Wikipedia as Participatory Journalism: Reliable Sources? Metrics for Evaluating Collaborative Media as a News Resource*. *Nature*, 3(1), 1-31.
- [6] Hannah Greving, Ulrike Cress, Joachim Kimmerle. (2019) *Anger in Wikipedia: Perceived Intentionality and Threat Appraisal as Mediators of Anger about Man-make Attacks*. *Journal of applied social psychology*, 49(2), 99-116.
- [7] Guo L. (2015) *A Theoretical Explication of the Network Agenda Setting Model*. *The Power of Information Networks: New Directions for Agenda Setting*, 1.
- [8] Entman R M, Matthes J, Pellicano L. (2009) *Nature, Sources, and Effects of News Framing*. *The Handbook of Journalism Studies*, 175-190.
- [9] Bryant S L, Forte A, Bruckman A. (2005) *Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia*. *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, 1-10.
- [10] Rainer R, Gisela G. (2019) *Wikipedia as a Space for Discursive Constructions of Globalization*. *International Communication Gazette*, 81(1), 3-19.
- [11] Maier S R. (2020) *The World View(ed) through the English-speaking Media Lens: Foreign News Coverage Steadfastly Narrow and Uniform*. *The Journal of International Communication*, 26(2), 155-170.