
| RESEARCH ARTICLE

An Analysis of the English Translation of *The Class Struggles in France* Based on AntConc

Qingkun Lin¹ ✉ and **Xiaohong Li²**

^{1,2}*College of Foreign Languages, North China University of Science and Technology, Tangshan, Hebei, China*

Corresponding Author: Qingkun Lin, **E-mail:** linqingkun@126.com

| ABSTRACT

Studying Marx's analysis of class struggles in France provides insight into historical socio-economic dynamics, revealing patterns of oppression, resistance, and revolution. It elucidates the role of the bourgeoisie and proletariat in shaping societal structures, informing contemporary understanding of power dynamics, inequality, and potential pathways for social change. Based on the English version of *the Class Struggles in France*, the language use and organization of it are analysed by using the retrieval and analysis functions of AntConc. Through the analysis, its language features are concise, clear structure, strong logicity, key terms centred around the main theme, and a coherent logical relationship. The vocabulary used exhibits emotional tendencies, particularly concerning communism and capitalism, highlighting the author's emotional attitude.

| KEYWORDS

Corpus, Marxism, AntConc, Marx, corpus linguistics, *The Class Struggles in France*.

| ARTICLE INFORMATION

ACCEPTED: 02 February 2024

PUBLISHED: 01 March 2024

DOI: 10.32996/ijllt.2024.7.3.3

1. Introduction

The Class Struggles in France is one of the most significant classic works written by Karl Marx in 1850. This work reveals the contradictions and conflicts between the proletariat and the bourgeoisie in French society during the mid-19th century, elucidating the theoretical framework and strategies of the proletarian revolutionary struggle. While previous discussions on *The Class Struggles in France* have delved deeply into its significance, there is a scarcity of research articles exploring its linguistic form. The emerging focus on "returning to Marx and exploring the true meaning of Marx's thought" has become a research hotspot. The essential characteristics of Marxism are dialectical materialism and historical materialism. Therefore, through the use of corpus as a research method, it is possible to unearth a plethora of language expressions and linguistic conventions related to these characteristics. This approach can reveal new issues, contradictions, and trends, allowing for a more in-depth understanding of the essential features of Marxism. It also provides new perspectives and methods for the innovation and development of Marxist theory.

2. Methodology

In the interpretation of discourse language, various methods can be employed. A common approach involves exploring the use of rhetorical devices from a stylistic perspective to unveil the artistic and aesthetic aspects of language in the text. By analyzing the application of rhetorical devices, a deeper understanding can be gained of how the author employs techniques such as imagery, metaphors, parallelism, etc., to enrich the expressive effects of the text. Another method is to study the overall linguistic tone of the discourse from a pragmatic perspective, aiming to uncover various implied meanings and the author's intentions within the text. By examining the pragmatic functions, connotation, and implicit information of language, deeper layers of meaning and rhetorical purposes in the text can be revealed.

However, this paper adopts corpus software as an analytical tool to achieve more objective statistical analysis. By utilizing corpus software, it becomes possible to extract and study the collocations and semantic relationships of certain keywords in the text of *The Class Struggles in France* from a large-scale corpus of textual data. Through the statistical analysis of the frequency of words, collocation patterns, and semantic characteristics of these keywords, the distinctive features of the text in terms of conciseness, logical coherence, and emotional expression can be revealed. This analysis further allows for inferences regarding the author's expressive intent and the conveyed information within the text.

2.1 Advantages of Corpus Research

The earliest definition of a corpus can be traced back to 1982 when Professor Francis from Brown University in the United States introduced the concept. He defined a corpus as a collection of texts for language analysis that is representative and can cover a certain language, dialect, or specific aspect of language (Francis, 1982). Scholars in corpus research firmly believe that frequently occurring forms, meanings, and functions often reveal the core and typical elements of language use. Research methods based on frequency reflect important inherent features of language systems and language use, providing a more objective and intuitive research process thereby increasing the credibility of results (Fang & Zhan, 2012).

Conducting textual analysis through corpus software makes research results more objective and authentic, helping to minimize the interference of subjective emotional factors to the greatest extent. This advantage sets corpus research apart from other methods of text interpretation.

2.2 Advantages of AntConc in Research

AntConc, developed by scholar Laurence Anthony at Waseda University in Japan, is a powerful, user-friendly, and efficient corpus research tool. It assists researchers in conducting in-depth corpus analysis, providing valuable insights into textual data. Specifically, the main features of AntConc include: 1. Word Frequency Statistics: AntConc allows for the calculation of word frequencies, providing insights into the usage preferences and textual emphasis of translators by analyzing the frequency of different vocabulary in translated texts; 2. Keyword Extraction: The tool identifies frequently occurring keywords in the text, aiding in determining the focus and areas of attention in translation; 3. Concordance Analysis: AntConc reveals lexical combinations in the text, offering an understanding of how translators handle complex concepts and providing insights into translation quality; 4. Vocabulary Distribution: It displays the distribution of vocabulary in the text, enabling the evaluation of stylistic features, fluency, and contextual adaptability. Additionally, AntConc possesses a crucial feature for calculating the keyness of target corpus keywords relative to a reference corpus. This functionality helps determine the importance and prominence of keywords in the target corpus within the entire corpus.

The advantage of using AntConc for analyzing the English translation of *The Class Struggles in France* lies in its ability to provide in-depth and comprehensive corpus research. Through functions such as word frequency statistics, keyword extraction, concordance analysis, and vocabulary distribution, AntConc facilitates a better understanding and analysis of the text. It enables readers to delve into the characteristics, translation strategies, and language processing of the translation, offering researchers more objective facts about the translation process and outcomes. This, in turn, assists readers in conducting meticulous research on the English translation of *The Class Struggles in France*.

3. Retrieval Results and Analysis of the *Class Struggles in France*

In this study, *The Class Struggles in France* from Volume 10 of "Marx/Engels Collected Works" (known as MECW) (Marx & Engels, 1982) is selected as the corpus base. Utilizing AntConc to retrieve and analyse vocabulary, keywords, and collocations, a relevant analysis of the language usage and organization in its English translation was conducted. Before the analysis using the software, the author performed text cleaning to eliminate irrelevant characters and symbols, ensuring the obtained data's objectivity and accuracy. The cleaning process aimed to remove interference factors, enhancing the accuracy and reliability of the analysis results. It involved removing punctuation, numbers, and special characters, as well as deleting irrelevant parts or non-text content to focus the analysis on crucial language content and reduce interference.

3.1 Analysis from "File View"

According to AntConc statistics, there are a total of 5,709 file types, 3,9819 file tokens, and a type-token ratio (TTR) of 0.1433, as shown in Table 1.

Table 1. Number and Ratio of File Types and File Tokens

Analysis	Results
File Types	5709
File Tokens	39819
Type-Token Ratio (TTR)	0.143

“File types” refer to the variety of vocabulary in the corpus, while “file tokens” refer to the number of words in the corpus. Yang (2002) pointed out that corpus linguistics primarily relies on a metric called Type-Token Ratio (TTR) to analyze vocabulary variation in texts. TTR measures the ratio between unique words (types) and total words (tokens) in a text. By calculating TTR, researchers can gauge the diversity of vocabulary used. However, TTR is influenced by text length; shorter texts tend to have higher TTR due to the frequent repetition of common words, while longer texts tend to have lower TTR (Mojca & Špela, 2022). A higher TTR indicates a more varied vocabulary usage in the text.

Based on this ratio, the variety of words in *The Class Struggles in France* is not very diverse. This aligns with the characteristic of a politically oriented article that does not prioritize linguistic richness. Such articles should maintain seriousness and utilize concise vocabulary to express the author’s firm stance and consistency because excessive use of synonyms may make the expression of the article’s viewpoint unclear and lack a strong position, which is undesirable in this genre of writing.

From a corpus perspective, a relatively low Type-Token Ratio (TTR) in *The Class Struggles in France* suggests that Marx and Engels chose to articulate their viewpoints using a limited set of words. This intentional choice of vocabulary contributes to making the entire manifesto more accessible and easily understandable, facilitating the better dissemination of their ideas. The emphasis appears to be on clear expression of thoughts rather than elaborate language embellishments.

This linguistic style is aimed at effectively communicating their viewpoints and eliciting a broader resonance among a diverse readership. The focus on clarity and simplicity in language enhances the accessibility of their ideas, aligning with a purposeful strategy to reach a wider audience with their revolutionary concepts.

3.2 Analysis from “Keyword”

Based on the analysis results in Table 1, this section aims to interpret the content by analyzing the outcomes of the “keywords”. Overseen by Oxford University’s Computing Services, the database used is British National Corpus (BNC), which is a vast collection of British English language samples totaling 100 million words, covering the period from the 1980s to 1993. It includes written texts from newspapers, journals, books, letters, and more, as well as transcribed spoken language from a variety of contexts and demographics. Construction of the corpus began in 1991, resulting in the release of the third edition, BNC XML, in 2007 (Reuben, 2021). By using BNC, one of the most representative contemporary English corpora, as a reference, the author conducted a “keyness search” (sorted by likelihood) to interpret the keywords. The results were organized, and Table 2 presents the refined outcomes.

Table 2. Top 30 Content Words

Keyword	Rank	likelihood	Frequency	Keyword	Rank	likelihood	Frequency
bourgeois	2	2350.688	228	barrot	17	531.412	34
assembly	3	1712.474	234	legitimists	18	515.781	33
bourgeoisie	4	1497.918	142	order	19	487.385	160
republic	5	1469.326	197	ledru	20	468.89	30
bonaparte	6	1239.089	95	june	21	466.81	116
proletariat	7	1214.602	113	cavaignac	22	460.055	30
revolution	9	842.283	130	provisional	23	447.56	56
national	11	728.652	214	paris	24	442.644	86
constituent	12	711.521	82	orleanists	25	421.999	27
montagne	13	622.102	53	ministry	27	397.426	75
revolutionary	14	619.995	90	party	28	379.149	147
petty	15	613.236	71	rollin	29	364.027	30
constitution	16	549.93	90	changarnier	30	359.478	23

Based on the keywords from the top 30, as presented in Table 2 (content words only), they can be classified into four categories based on their semantics and functional components. The first category encompasses function words devoid of substantive meaning, such as “the”, “of”, and “its”. These words contribute to the coherence and cohesion of the text. The second category includes proper nouns like personal names and political party names that play specific roles or represent particular organizations within the text. The third category consists of vocabulary closely associated with the proletariat, such as “proletariat”, “revolution”, and “republic”. These terms reflect the status of the proletariat and underscore the significance of their involvement in the revolution, serving as crucial terminology for describing the struggles of the proletariat and societal transformation. The fourth category comprises vocabulary related to the bourgeoisie, including terms like “bourgeois” and “bourgeoisie”. These words pertain

to the class status and characteristics of the bourgeoisie, used to depict their economic, social, and political influence, as well as the antagonistic relationship with the proletariat.

Considering the lower significance of the first category of words, they may be temporarily excluded from content interpretation. The second category comprises proper nouns related to historical figures and political parties. The evaluation of these nouns may vary among individuals and is also contingent on different historical periods; therefore, a detailed discussion is omitted here. In contrast, the third and fourth categories of words exhibit distinct class characteristics. These terms can offer crucial information and reveal the reflections and perspectives on the bourgeoisie, indicating a focused consideration of the capitalist class within the discourse.

By interpreting the content words among the top 30 keywords, an overall understanding of the content of *The Class Struggles in France* can be gained and helpful to recognize its evident political stance. Reviewing the previously mentioned word frequency list, is it possible to extract subtle features of the textual language from those low-frequency words that only appear a few times? Gui suggests that studying rare vocabulary in the process of textual interpretation, especially words that occur only once (hapax legomenon), also holds significance (Gui, 2009). This kind of research can indirectly illustrate the frequency of word usage and the characteristics of the text. Through the Invert order function of AntConc, after organization and analysis, it is noted that there is no hapax legomenon, and the minimum word frequency is twice. Nevertheless, even with these relatively low-frequency words, one can still explore the subtle features and characteristics of the text. Some low-frequency words are presented in Table 3, as shown below:

Table 3. Partial Low-Frequency Words

Words	Frequency	Words	Frequency	Words	Frequency	Words	Frequency
râteau	2	rateau	2	betook	2	hitte	2
baroche	2	girardin	2	braggadocio	2	cabet	2
barbes	2	berryer	2	gouvernement	2	carlier	2
boutiquiers	2	proroguing	2	civitavecchia	2	citoyens	2
crémieux	2	marchands	2	gloire	2	réforme	2
ledrurollin	2	handicraftsmen	2	napoléon	2	petto	2
coquettish	2	encumbering	2	laffitte	2	vacillated	2

“râteau”, “hitte”, and “berryer” are names of historical figures, while “citoyens”, “marchands”, “gouvernement” are French vocabulary. Considering the historical period, namely the mid-19th century in France, the country was undergoing industrialization, social transformation, and political upheaval. The names of historical figures are relevant to the various factions and political figures on the political stage of France at that time. The presence of French words was influenced by the linguistic environment of France during the 19th century, for French was the official language, and many terms in social and political domains were expressed in French. Therefore, the use of French vocabulary in the text is closely tied to the societal and political background of France during that period.

3.3 Analysis from “collocate”

Based on the retrieval results from AntConc, the author identified key terms including two opposite classes, “bourgeoisie” and “proletariat” (refer to Table 2 for details). Given this scenario, the author analyzed the co-occurrence and collocation of these two terms. By examining their collocational patterns and frequencies, the study further explores the semantic collocational characteristics of “bourgeoisie” and “proletariat”, as depicted in Tables 4 to 7.

Table 4. High-Frequency Words to the Left of “Proletariat”

Collocation	Rank	Frequency
paris	2	18
revolutionary	3	12
industrial	4	6

Conducting a collocational search for the term “proletariat” with a span of 1 word to the left, the results were organized and presented in Table 4. According to the statistical analysis, the expression “Paris proletariat” appeared 18 times, “revolutionary proletariat” appeared 12 times, and “industrial proletariat” appeared 6 times in the text. These concepts reflect his recognition of the proletariat as a primary force in societal transformation. *The Class Struggles in France* is a political document written by Karl Marx between 1848 and 1850, a period marked by industrialization and social transformation in France. During this era, Paris

served as the political and industrial hub of the country, witnessing the gradual emergence of the industrial proletariat as a significant social class. Concurrently, the Marxists emphasized class struggle and attached importance to the proletariat in Marx's ideological framework.

Table 5. Words to the right of "proletariat"

Collocation	Active/passive	frequency
compelled, forced	passive	2
decimated, pushed, driven	passive	1
allow	active	2
gain, revel, oppose	active	1

The author also conducted a search for verbs occurring within the immediate context of the term "proletariat" with a span of 5 words to the right, and the results are presented in Table 5. During this period, the proletariat was subjected to passive oppression and exploitation by the bourgeoisie; hence, the passive forms of verbs reflect the social status and predicament of the proletariat. The passive voice of verbs implies the role of the proletariat as an oppressed class under the dominion and control of the bourgeoisie. This aligns with Marx's theoretical perspective on the bourgeoisie exploiting the proletariat. Considering the influence of social structure and class relations, in capitalist societies, the bourgeoisie holds control over the means of production and social resources, while the proletariat is compelled to sell their labor power, assuming a passive role as laborers. This social structure determines the passive status and controlled situation of the proletariat. Analyzing the aforementioned verbs expressing agency, it is evident that the proletariat engages in struggles against oppression, striving for rights (suffrage) in the process.

Table 6. High-Frequency Words to the Left of "bourgeoisie"

Collocation	Rank	Frequency
petty	2	31
big	3	8
industrial	3	8
French	5	6

Similarly, a collocational search for the term "bourgeoisie" was conducted with a span of one word to the left, and the results were organized as presented in Table 6. During this period, the bourgeoisie emerged as the dominant force in French society. The term "petty bourgeoisie" in the text refers to relatively smaller factions of the bourgeoisie, such as small business owners and independent professionals, reflecting the presence of small-scale business owners and independent craftsmen in French society at the time. The "petty bourgeoisie" often faced economic competition and oppression from the larger bourgeoisie while also having certain conflicting interests with the proletariat.

On the other hand, the "big bourgeoisie" and "industrial bourgeoisie" denote the affluent class possessing substantial capital and industrial enterprises, exercising economic and political power. The contradictions and struggles between them and the proletariat were focal points of Marx's attention. The term "French bourgeoisie" encompasses the entire bourgeois class in France, reflecting the status and influence of large-scale business owners and industrial capitalists in the country. Describing and analyzing different types of bourgeoisie reflects Marx's contemplation on the roles and relationships of various classes in capitalist society. It unveils Marx's attitudes and viewpoints towards these classes in the text.

Table 7. Words to the right of "bourgeoisie."

Collocation	Active/passive	frequency
allow	active	2
vote, preserve, rule, discover, wish, refute, regulate	active	1
condemned, abolished, forced	passive	1

Similarly, the author conducted a search for verbs occurring within the immediate context of the term "bourgeoisie" with a span of 5 words to the right. The results revealed a higher proportion of verbs indicating agency, while passive verbs were less frequent. This suggests that the bourgeoisie rose to prominence during this period, attaining significant economic and political status. Consequently, the presence of active verbs in the text reflects the bourgeoisie as the dominant class in society at that time, playing crucial roles in societal transformation and social relations. Marxism emphasizes class struggle and the liberation of the proletariat. Marx posited that the bourgeoisie constitutes the exploiting class, while the proletariat is the exploited class. The active verbs

associated with the bourgeoisie in the text imply their proactive role in driving social change for their interests, concurrently imposing oppression and exploitation upon the proletariat.

Therefore, from the analysis of the four tables above, it is evident that in *The Class Struggles in France*, Marx's depiction and portrayal of the bourgeoisie and proletariat are filled with criticism and conflict. Through corpus analysis, it is observed that in collocations related to the bourgeoisie, there is a prevalence of verbs indicating active actions, reflecting the dominant status of the bourgeoisie in society at that time and their active role in societal transformation. In contrast, collocations associated with the proletariat may lean more towards passive or constrained expressions. This difference in the use of verbs reflects Marx's observations of the contradictions and opposition between the bourgeoisie and the proletariat. As the founder of Marxism, Marx considered the bourgeoisie to be the exploiting class, obtaining benefits through the exploitation of the proletariat. Marx's ideological inclination is aligned with the perspective of standing with the proletariat, focusing on their liberation and societal transformation.

4. Results and Discussion

In *The Class Struggles in France*, Marx often employs language with strong emotional undertones. Through criticizing and exposing the actions of the bourgeoisie, as well as expressing support for the proletariat, he conveys his anger and dissatisfaction with social inequality and exploitative systems. This emotion is reflected in his portrayal of the bourgeoisie, where the use of verbs emphasizes the bourgeoisie's agency and pursuit of selfish interests.

5. Conclusion

By applying the corpus-based method to analyze *The Class Struggles in France*, this paper aims to uncover language expressions and habits related to the essential features of Marxism, thus deepening the understanding of Marxism and providing new ideas and methods for its theoretical innovation and development. The analysis reveals that *The Class Struggles in France* strives for concise language selection, avoiding redundant usage of words, resulting in clear and understandable language expression with precise meanings. Key terms are prominently highlighted in the text, indicating various interests of the exploited classes and emphasizing the position and goals of the proletariat, which suggests that *The Class Struggles in France* exhibits overall logical coherence, clear viewpoints, and vivid language imagery.

However, it should be noted that since the original version of *The Class Struggles in France* is in German, while this paper refers to its English translation, there may be differences in word choice and expression between the translated version and the original German version, which may lead to some inconsistencies in text analysis. Thus, it is important to note that any interpretation of the text should be considered from multiple dimensions and should take into account the author's intent, historical context, and limitations of the research methods. Therefore, when studying the text, researchers should maintain critical thinking and integrate various sources of information to understand the text's connotations more comprehensively and accurately.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Fang, X., & Zhan, H. W. (2012). Interpretation of the English translation of the Manifesto of the Communist Party: A corpus-based analysis. *English Teacher*, 12(10): 48-53.
- [2] Francis, N. W. (1982). Problems of assembling and computerizing large corpora. In S. Johansson (ed.), *Computer corpora in English language research* (pp. 7-24). Norwegian Computing Centre for the Humanities.
- [3] Gui, S. C. (2009) *A Corpus-Based Analysis of the Register of English Linguistics*. Beijing: Foreign Language Teaching and Research Press.
- [4] Marx, K., & Engels, F. (2010). *Engels. Marx/Engels Collected Works (10)*. Lawrence & Wishart, 47-145.
- [5] Mojca B., & Špela, V. (2022). Lexical Diversity in Statistical and Neural Machine Translation. *Information*, 13(2): 93-93.
- [6] Reuben, N. (2021). Societal Age Stereotypes in the U.S. and U.K. from a Media Database of 1.1 Billion Words. *International Journal of Environmental Research and Public Health*, 18(16), 8822-8822.
- [7] Yang, H. Z. (2002). *An Introduction to Corpus Linguistics*. Shanghai: Shanghai Foreign Language Education Press.