

Comparison of Judgment Document Information Between China and Russia Based on Corpus

Yuanzhuo Pi^{1,a,*}

¹Guangdong University of Foreign Studies, 2 Baiyun Avenue, Baiyun District, Guangzhou, 510420, China

a. piyuanzhuo@163.com

*corresponding author

Abstract: The study of legal comparisons has become particularly important as the number of foreign-related legal disputes has increased dramatically with the growing number of cross-border trade. This study discovers the differences and similarities between Chinese and Russian judgment documents and provides a linguistic basis for the academic theoretical researches and the call for linguists to enter the courtrooms by establishing a corpus of Chinese and Russian judgement documents, quantifying and analysing the Chinese and Russian judgement documents based on the Legal Discourse Information Theory. It can be found that the macro-structure of civil ruling documents is clearer but does not provide more micro-information. By adding language elements, it can be seen that there are more information points and information components in Russian criminal judgment documents. The information points in Chinese civil judgment documents are more concentrated than in Russian civil judgment documents, while they are more similar in terms of the proportion and type of information components.

Keywords: Judgment Documents, China-Russia Comparison, Information Comparison, Corpus.

1. Introduction

Since the beginning of the 20th century, China has step on the path of modernising its legal system by learning from the West. Due to the national system, the social system and various historical reasons, the Soviet Laws had a great impact on our country. After the founding of PRC, China embarked on an all-encompassing introduction of Soviet law, which was mainly demonstrated on the legislative and judicial levels. Macroscopically speaking, the current concept of law in China carries the general definition of law from the Soviet period, which holds that law is used to protect, consolidate and develop social relations and order favourable and suitable for the ruling class. However, in socialist countries, law is used to reflect the will of the working people, through which it is used to undermine and eliminate the old system of exploitation and establish new socialist relations, and is an instrument of class domination. China's legal theory has always followed the viewpoint of instrumentalism. For a long time, scholars have believed that law is a tool in the service of politics, closely linking the state, politics and law. It was only with the change in thinking brought about by the changes in people's lives after the Reform and Opening that scholars have begun to look at this viewpoint in a dialectical way, and reformed some of the systems that had been affected by this viewpoint.

In the era of global economic integration, transnational trade is increasing particularly, which causes the dramatic rise in foreign-related economic legal disputes. Taking Fujian Province in China as an example, the number of foreign-related cases increased by 15.44% year-on-year in 2022, and the most foreign-related maritime affair disputes also exceeded more than 50%. The situation of development urgently requires a group of legal practitioners who not only have expertise in our laws, but also have bilingual or multilingual skills, and even erudition of the legal systems of two or more countries. But from the perspective of academic research, the study of legal comparison is especially important. A systematic and specialised contrastive study of the laws of the two countries can provide them with a solid foundation of academic theory. This study compares and analyses Chinese and Russian judgment documents to explore the differences and similarities by peeling off the texts according to information stratification.

2. Research Review

Chinese scholars' linguistic studies of judgment documents have explored more about the realisation of the trial process on the objectives the case. Guo Fei and Wang Zhenhua established an analytical framework for the embodiment of social processes in legal discourse [1] by combining Wang's Types of Social Processes (conforming, assimilating, co-operating, competing, conflicting) [2], and from a sociological perspective and by analysing lawyers' representations and trial discourse, they found that lawyers' representations include more negative approach, and that they conform to the social process by becoming a judicial 'unity' with the represented to comply with the social process and achieve their own ends; by adjusting competition and conflict, reducing polar oppositions and promoting conformity, assimilation and cooperation they are accessible to achieve the ultimate ends of justice. Yang Fan [3], on the other hand, analysed the trial courses by using Damaska's Theory of Judicial Types, arguing that even under the influence of the inquisitorial ideology of the dynamic national judicial system, the exercise of judges' power follows the principle of the neutrality to a certain extent, and that most of the exercise of judges' power is also intended to maintain the procedures and order of the courtroom. After analysing 6 civil judgment documents, Dong Min [4] drew a conclusion that in civil judgment, the judge gives a verdict as the sender of information, which is in a dominant position; the both parties and the public receive judgments as the recipients of information, which are in a subordinate position. And the both parties are in an equal litigation position but have the greatest social distance between them. The judgment document, as the written carrier for the formation of the jural relations of procedure to the litigation between the court and the both parties, has two main functions: to announce the conclusion of the verdict and to substantiate the conclusion of the judgment.

Foreign scholars have studied more extensively the information conveyed by the language of judgment documents. Melissa [5] used a corpus to analyse the impact of the judge's subjective attitudes on female victims in domestic violence cases, with discourses indicating that ending a violent relationship was the preferred route for female victims of domestic violence to exercise their power of agency and a few discourses suggesting that the main reason women chose to stay in a violent relationship was because they were influenced by the expert testimony given at the trial. Diana [6] discusses four ideologies that recur for different subjects in the courtroom: contradictions (inconsistencies), decontextualization, narrator identity, and the ideology of repeated questioning. This linguistic analysis of particular discourses contributes to an understanding of how social injustice is present and practised in the legal process.

In recent years, there has been a gradual increase in Russian scholars' research on litigants in judicial discourse, with a particular focus on the impact that litigants have on judicial discourse. The scholars have explored the discourse characteristics [7], semantic content [8], textual features (Никифорова), and means of verbal influence [9] of courtroom participants in different countries.

Pavlova [10] focuses on the characteristics of the discourse of a certain category of court participants, arguing that the linguistic messages conveyed by court participants of different identities differ. She argued that lawyers, as a special category of participants in litigation whose interest in the outcome of litigation is not self-interest, are usually guided by a combination of emotion and rationality in selecting arguments that may influence jurors, whereas more important explanatory information may be achieved by providing comments or explanations, with the emotional part of the expression being more intended to have an interaction or impact on the listener.

3. Collection and Labelling of Corpus for Comparison

The main content of this study is the comparative analysis of the structure and types of information in Chinese and Russian judgment documents, based on which the corpus of information on judicial discourse required for this study has been created, has been designed and set up with the aim of providing a corpus and a uniform classification specifically for this study. The corpus involved in the study is judicial discourse, i.e., spoken or written discourse produced in the course of judicial proceedings, which includes judgment documents, rulings documents, and trial transcripts. The specific quantities and classifications are listed in the table below:

Table 1: Types and Quantities of Judgment Documents Included in the Corpus

Languages	Type	species	Quantity
Chinese	Judgement	Civil judgment	9
		Criminal judgment	9
	ruling	Civil ruling	7
Russian	Judgement	Civil judgment	9
		Criminal judgment	9
	Ruling	Civil ruling	7

This table categories and labels the corpus according to the Legal Discourse Information Theory. The Legal Discourse Information Theory combined Lasswell's 5W (Interrogative words) Mass Communication Model, Labov's Information Categorization Model and Buring's Discourse Tree Model. It forms a basic framework for the analysis of the information structure of applicable legal discourse by dealing with information in an interrogative-word mode and combines the concept of hierarchical systems in the Discourse Tree Model and the specific methods of information categorization in the information categorization mode.

Information point is the most essential element of Discourse Information Structure which plays a key role in the establishment of the information structure [11]. It clarifies the relationship between two information units through the identification of information points. There are 15 types of information points in total: What True(WT), What Basis(WB), What Fact(WF), What Inference(WI), What Disposition(WP), Who(WO), When(WN), Where(WR), How(HW), What Effect(WE), What Condition(WC), What Attitude(WA), What Change(WG), What Conclusion(WJ).. According to the type of relationship they represent, they can be categorized into categories of state (WT, WF, WO, WN, WR), process (HW, WY, WE, WC, WG) and attitude (WB, WI, WP, WA, WJ). The judgment document has a strong institutional and chapter layout of the normative. In case of judgment document, for example, the facts of the case need to be ascertained, and the trial court will hear and judge the evidences submitted by the parties, and apply the law to the circumstances of the case, draw conclusions and make decisions. Throughout the process, the number of information points in the disposal category (WT, WB, WI, WP, WJ) will far exceed the representation category (the remaining

10 except for the disposal category). The labelling of information points enables the formation of a macro-structure of the adjudication document, which is a framing statement of the whole discourse.

Micro-structure is the interpretation and deconstruction of the macro-structure of the discourse, i.e., the analysis of propositions and inter-propositional connections. The propositions at the micro level can be understood as information components, which form a chain of information and connect with each other to form an information tree. According to the Legal Discourse Information Structure, information components are mainly divided into three categories: process, individual and environment, as shown in the following table:

Table 2: Types and Symbols of Information Components

Names	Symbols	Name	Symbols	Name	Symbols
Process	P	entity	e	condition	c
State	S	agent	1	instrument	i
Quality	Q	dative	2	location	l
Appear	A	patient	3	source	s
Relation	R	facilitate	4	goal	g
Behaviour	B	attribute	5	comitative	c
Cause	C			time	t
Turn	T			affected	a
Negation	N			with	w
				situation	o
				basis	b
				manner	m
				elaboration	e

This table provides a systematic data analysis of the annotated and organized corpus. It presents the key data in the discourse of each judgment document to form a comprehensive data quantification table containing the total number of information points, the total number of words/phrases in the discourse, and the total number of information components. Meanwhile, it performs a statistical analysis of key data (number of words/phrases in the discourse, number of information points, number of information components), captures and illustrates extreme data, and performs preliminary and macro comparative analysis of Russian-Chinese judicial discourse based on tabular data.

4. Data Comparison and Analysis

1) Relationship Between the Total Number of Words/Phrases in the Discourse and the Total Number of Information Points

In the selected corpus data, there is an overall positive correlation between the total number of words/phrases of the discourse and the total number of information points. Although there are some peculiarities that not all words and information points of each discourse are in equal proportions, but the overall trend is consistent, as shown in the figure below:

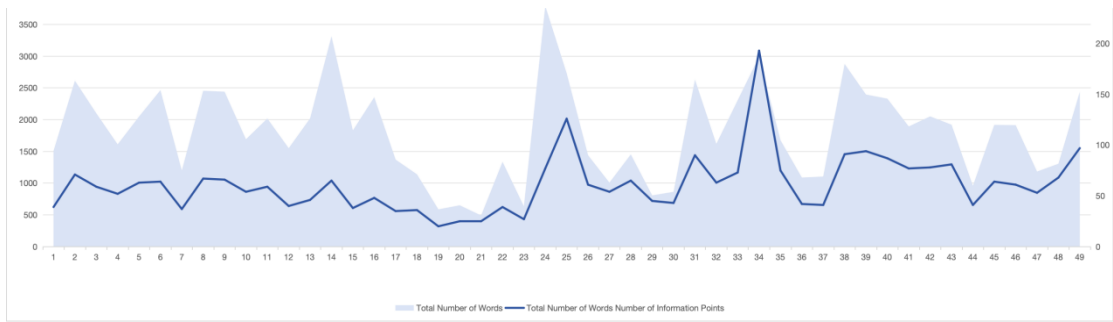


Figure 1: Correlation Between the Total Number of Words and the Number of Information Points in Selected Discourse in the Corpus

The rigorousness and accuracy of legal documents require that judicial discourse should be concise and to the point, in particular, as the outcome of the case and the basis for accountability, the judgment document should provide a brief description of the case, investigate and draw conclusions with regard to the main points of contention between the two parties or the main points of the case. Repetitive language and redundancy of information should be avoided in each part of it. Therefore, the more content there is in the text, the greater the number of information points.

2) Relationship Between the Number of Information Points and the Number of Information Components

In the selected corpus, the number of information points shows an absolute positive correlation with the number of information components, and it is particularly obvious in the Chinese judgment documents (discourse 1-18). As mentioned previously, although the total number of words in a discourse is not absolutely correlated with the number of information points, a higher number of information points in a discourse implies more information in the macro-structure. And information points are made up of many information components. Therefore, more information components are necessary to generate more information points.

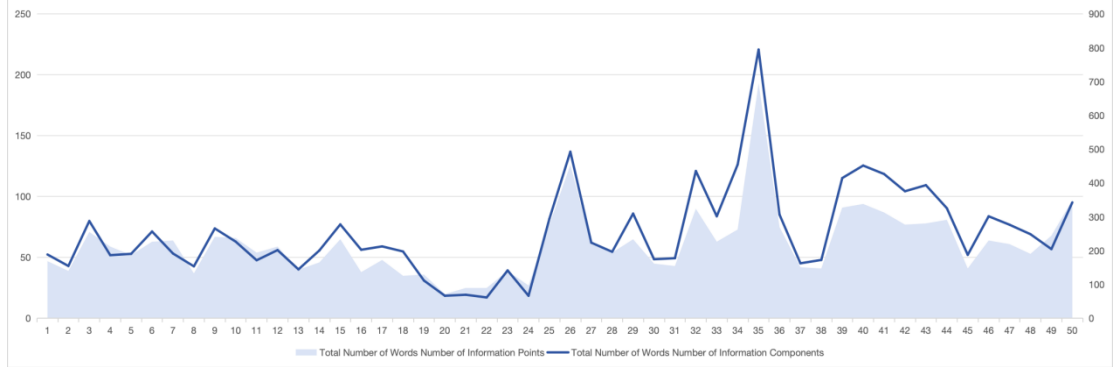


Figure 2: Correlation Between the Number of Information Points and the Number of Information Components in Selected Parts of Speech in the Corpus

3) Comparison of Information Content in Various Types of Judgments Documents between China and Russia

To highlight the interlanguage comparison and to explore the scale of the information content of each type of Russian and Chinese discourse respectively, the average of discourse data for 6 sections from 3 types of documents from Russia and China have been taken to calculate the ratio of the total number of information points to the total word count, as well as the ratio of the total number of information components to the total word count, as shown in the table below:

Table 3: Multi Level Information Values and Ratios of Different Language Document Types

Type	Language	Total Number of Words (average)	Number of Information Points (average)	Rate	Number of Information Components (average)	Rate
Civil judgment	China	1961	55	2.83%	209	10.65%
	Russia	1651	69	4.17%	308	18.64%
Criminal judgment	China	2066	50	2.43%	204	9.89%
	Russia	2046	86	4.23%	389	19.02%
Ruling	China	1233	36	2.89%	116	9.44%
	Russia	1663	66	4.00%	270	16.24%

The total word count of Chinese civil judgment documents is higher than that of Russian civil judgment documents, and the word number of Russian ruling judgment documents is much larger than that of Chinese ruling judgment documents. Additionally, the word count of criminal judgment documents between China and Russia is comparable. Nevertheless, the information point ratio of Russian documents of various types is much higher than that of Chinese documents, especially civil and criminal judgment documents, which can provide more information points in the overall discourse volume compared to China. This is caused by the fact that Russian judgment documents usually contain direct quotations from the original text of legal provisions in factual findings and legal support, while it is included approximately 10 articles in a case, each of which has a small number of words but is more scattered and trivial in the perspective of information points. Moreover, Russian criminal judgment documents usually contain multiple repeated information points, such as WN (when), WR (where), etc.. Thus, the information content of criminal judgment documents is larger than that of Chinese ruling documents because some physical issues, such as the parties' claims in the lawsuit, are included in Russian ruling documents, which generate more information points for discourse of ruling documents.

From a microscopic aspect, Russian criminal judgment documents have the highest rate of information content, which is influenced by the ratio of information points in Russian criminal judgment documents due to the positive correlation demonstrated by information points and information contents. On the contrary, the information content ratio of various documents in China is relatively low. Obviously, redundant information is avoided as much as possible in China's judicial documents, and fewer information components are provided at larger lengths, with a small order of magnitude of information. Instead, those information components are more concentrated on the trial process. In addition, there is the largest disparity in information content ratio between Chinese and Russian criminal judgment documents: Chinese judgment documents are longer in length but have smaller information content, while Russian criminal judgment documents are the opposite.

5. Conclusion

The comparison shows that the total value of information points in civil ruling documents is the highest, while that in criminal judgment documents is the lowest; The total information content of criminal judgment documents is the highest, while that of civil rulings is the lowest; The civil

judgment is relatively balanced in terms of information points and information components. It can be found that the macro-structure of civil ruling documents is clearer but does not provide more micro-information. By adding language elements, it can be seen that there are more information points and information components in Russian criminal judgment documents. Russian criminal judgments documents are not only more clearly delineated in the macro-structure, with the possibility of more long sentences or compound propositions, but also present more micro-information through information points.

The information points in Chinese civil judgment documents are more concentrated than in Russian civil judgment documents, while they are more similar in terms of the proportion and type of information components. In criminal judgment documents, there are significant differences and similar types of information points between Russian and Chinese, with Russian information points being more concentrated and the quantity and types of information components generally converging. In civil ruling documents, the proportion of information points is much higher than that in judgment documents in both Russian and Chinese, and there is also a significant difference between the two. The first three types of information points in Chinese civil rulings documents have a high proportion, which may be reflected as a single content in the discourse. The two are also similar in information point types, but the information components show large discrepancies, with a difference of nearly 10% in proportion, and there are also two inconsistent types of information components.

This study only deconstructed the judgment documents of China and Russia at the data level. The specific characteristics of various judgment documents and the differences in legal systems between the two countries reflected in them can be demonstrated by data. However, detailed comparative analysis and in-depth analysis of the reasons will be carried out in subsequent research.

References

- [1] Guo Fei and Wang Zhenhua. *The Impact of Interpersonal Semantic Distribution on Social Processes in Legal Discourse* [J]. *Journal of Northeastern University (Social Science Edition)*, 2017 (1) : 97-103.
- [2] Wang Zhenhua. *Legal Discourse as a Social Process and its scheming Semantics* [J]. *Foreign Language Teaching*, 2015 (1) : 1-6.
- [3] Yang Fan, *The function of Discourse Analysis in Judicial Research* [J]. *Journal of East China University of Political Science and Law*, 2018 (4) : 159-169.
- [4] Dong Min. *On the Linguistic Structure Potential of Chinese Civil Judgment Documents* [J]. *Rhetoric Learning*, 2006 (4) : 33-36.
- [5] Melissa, Hamilton (2010). *Judicial discourses on women's agency in violent relationships: Cases from California*. *Women's Studies International Forum*, 570–578.
- [6] Diana Eades (2012), *The social consequences of language ideologies in courtroom cross-examination*. *Language in society*, 41, 471-497.
- [7] Dubrovckaya T.V. *Judicial discourse: speech behavior of a judge (on the material of Russian and English languages)*. – M.: Izd-vo Akad. MNAPU, 2010.
- [8] Avakova M. L., Budnyaya O. N. *Metadiscursive markers as means of realizing the category of persuasion in English-language judicial discourse* // *The World of Science, Culture, and Education*. 2020. – № 4(83). – C. 465–467.
- [9] Litvishko O.M. *Syntactic regularities of English-language legal discourse as a means of realizing the normative-regulative function of law* // *Professional communication: topical issues of linguistics and methodology*. -2020. - № 13. - C. 124-132.
- [10] Pavlova Y.A., Eltsova V.I. *Speech of a lawyer as a special genre in English-language judicial discourse* // *Postulat*. - 2019. - №1-1 (39).
- [11] Du Jinbang, *Discourse Information Flow* [J]. *Foreign Languages*, 2009 (3) : 36-43.