



ARTICLE



<https://doi.org/10.1057/s41599-022-01491-z>

OPEN

A corpus-based interpretation of the discourse-cognitive-society triangle on Chinese court judgments

Jingjing Wu¹ , Le Cheng²  & Yi Yang³

A court judgment is a common legal discourse and the final carrier of court trial activities. From the perspective of socio-cognitive discourse analysis, this study aims to describe and explore the Chinese court judgments in a corpus-based method, concentrating on the interactions among discourse, cognitive, and social dimensions. We have three key findings through an empirical qualitative analysis of the court judgments in China. First, the discourse dimension of Chinese court judgments is both society-oriented and cognition-oriented. The discourse components could mark the cognition sources in court judgments. Second, the cognitive source of faith is a part of social cognition on law and regulation, and induction and paraphrase provide the personal cognition to testimony, documentary, or hearsay evidence. Besides, the cognitive source of inference could change personal cognition into a social consensus through a reasoning process. Third, the social function of court judgments corresponds to the cognitive source and builds the surface structure with various discourse components. Moreover, a probe into the multi-dimensional relationship in court judgments can offer practical insights into the interpretation of legal texts in Chinese judicial decision-making.

¹College of Foreign Studies, Guilin University of Technology, Guilin, China. ²Guanghua Law School, Zhejiang University, Hangzhou, China. ³School of Economics and Management, Guangxi University of Science and Technology, Liuzhou, China. ✉email: chengle163@hotmail.com

Introduction

Influenced by ontology and post-modern trends, critical discourse analysis (hereafter CDA) concerns social issues, focusing not only on language or language use but also on the linguistic features in cultural structures and social processes (Titscher et al., 2000; Sun and Cheng, 2017). Cognitive discourse analysis has been one of the most important research directions of CDA during the past 30 years, and this approach analyzes discourse from the perspective of cognitive psychology, attaching importance to the media role of cognition (Fairclough and Wodak, 1997; O'Hallora, 2003; Hart, 2010; van Dijk, 2014a). Among the studies, van Dijk (1998) suggests a socialized method to study cognitive psychology in discourse processing, attempting to discover discourse participation in social life from the linguistic perspective. Different from the discussion of language and social relations in sociolinguistics or linguistic anthropology, van Dijk's socio-cognitive discourse analysis (hereafter SCA), also called "Discourse-cognitive-society Triangle", studies psychological representation, discourse processing, shared knowledge and ideology (van Dijk, 1998, 2009a). In SCA, discourse is redefined as the form of social interaction and the reproduction of social cognition; thus, the social attribute of discourse is explained from the social and cognitive perspective. Moreover, the role of personal and social knowledge is essential in text interpretation (van Dijk, 2014a). Thus, this study explores Chinese court judgments from social, cognitive, and discursive perspectives, and analyzes a large amount of empirical data statistically to show the deep relationship among discourse components, cognitive sources, and corresponding social functions in the judicial discourse.

A court judgment is a common legal discourse and the final carrier of court trial activities. A court judgment is "a decision made by a court in respect of the matter before it" (Martin, 2009, p. 271), which may be interlocutory in deciding a particular issue prior to the trial of the case or disposing of the case finally. A court judgment may impose personal liability on a party or determine some issues of right, status, or property (Martin, 2009). The discursive presentation of different power agents can find its root in social construction because the discursive structure is constrained by and representative of social structure (Cheng, 2012). Regarding social relations, a court judgment is a corresponding rule on a specific social conflict. Due to the different types of social conflicts, there are civil, criminal, and administrative judgments in China. However, Chinese legal texts have much more informational and slightly abstract styles (Sun and Cheng, 2017). A court judgment is a model type of logical reasoning with tangible evidence; thus, it is significant to explore the cognitive source and social function according to the discourse components. This study builds a corpus with the Chinese court judgments (hereafter CCJ) of 1,740,000 words from the Supreme People's Court (SPC) and the local people's courts in China. There are 54 civil, 57 criminal, and five administrative judgments, including cases with murder, robbery, property disputes, and other sinful plots. Many data statistics and manual labeling are carried out in the discourse components, cognitive sources, and social functions to reveal the socio-cognitive discourse construction of Chinese court judgments.

Literature review

As a social value system, the law is composed of language. Even the rules, regulations, court trials, and various judicial procedures are all realized through language (John, 2003). In the sociology of law, many studies explore the forms of specific inequalities built into the law, legal processes, and decision-making (Seron and Munger, 1996; Sandefur, 2008). Although CDA shows interest in the role of language and communication in social functions, particularly in the exercise of power and control, it is notable that

much less attention is paid to the analysis of legal language (Cheng and Machin, 2022). A court judgment is an essential type of legal discourse, while previous studies often focus on the wording, syntactic structure, and discourse style of legal provisions. Many comparative studies analyze the discourse styles of court judgments in different countries, concluding that the different roles of judges in the judicial system produce completely different discourse styles (Wetter, 1960; Kurzon, 2001; Cheng, 2007a; Cheng et al., 2008). The textual features of court judgments in different countries are constrained by and representative of their respective social structures from a multi-dimensional perspective (Bhatia, 2004; Cheng, 2010). Solan (2010) believes that in writing a court judgment, judges can only accept the views of evidence admitted by the judicial system and must decisively explain the trial. Besides the role of the judge, the value system of court judgments can be divided into legal and social value systems. Thus, judgments will reflect multiple values, which could be compatible or conflicting (Cheng, 2007b). In SCA, social cognition is located both in the social structure at the macro level and in the specific interactions or events at the micro level, which helps to reveal the deep rules of court judgments from the discourse analysis.

In CDA, the early researchers analyze the types of court judgments to explore their intertextuality, interpersonal relationship, and communicative purposes (Bhatia, 1993; Maley, 1994; John, 1994). Intertextuality permeates the judicial discourse in many different aspects, and all judgment precedents are a matter of intertextuality. Court judgments are the written forms of interaction among previous judges, litigants, lawyers, and expert witnesses, which embodies heavy intertextuality (Yu, 2021). In this sense, a court judgment involves the consultation among different levels of texts, and it should use the different texts effectively and simultaneously, taking into account the political and ideological functions and social customs (Bell and Pether, 1998). Meanwhile, legal texts exhibit distinctive discourse structures and message distribution depending on communicative functions and genres (Cheng et al., 2008). As the typical legal discourse, a court judgment is the discursive representation of judicial thinking, representing how judges apply the principles and methods in judicial proceedings through adjudication, including case hearings, trials, and decision-making (Cheng, 2010). In this way, a court judgment carries ideas and values and shapes social practices (Cheng and Machin, 2022). In the past 20 years, the research perspective of court judgments has been more diversified, especially of the Chinese judgments. Cheng et al. (2008) examine the linguistic characteristics, moves, and rhetoric of Chinese and American court judgments to specify the rhetorical preferences of "standard" judgments. Solan (2010) analyzes the reflexive pronouns in English court judgments, believing the ambiguity of reference can cause potentially disastrous consequences. Cheng (2012) deals with the attribution and judicial control from the authorial voices in a corpus-based study. Cheng and Cheng (2014) examine how epistemic modality is employed in civil judgments to construct legal facts and indicate legal probability. From a semiotic perspective, Wu and Cheng (2020) construct a model of evidentiality in Chinese court judgments. Yu (2021) argues that the reporting verbs reflect how judges identify the evidence of different documents in Chinese court judgments. Alghazzawi et al. (2022) use an LSTM + CNN neural network model with an optimal feature set to predict court judgments efficiently. Overall, these studies analyze the relationship between language and the law. From the perspective of CDA, the language of the law classifies the world and represents identities and human agency (Cheng and Machin, 2022).

Data and methods

Corpus linguistics (hereafter CL) is a methodology for studying the use of language. A corpus-based approach looks at the tangible evidence of the corpus and analyses the evidence to find out the probabilities, trends, patterns, and co-occurrences of elements, features, or groupings of features (Teubert and Krishnamurthy, 2007). The present study is situated within a corpus-based CDA since the non-obvious meaning is not accessible to the naked-eye and direct observation (Partington et al., 2013). CL helps the interpretations to be more trustworthy (Subtirelu and Baker, 2018), which runs against subjectivity and over-generalization (Hart and Cap, 2014). The integration of CL and CDA can be traced back to the 1990s (Subtirelu and Baker, 2018), which has been extensively utilized to approach the discourse of law (Wu and Sun, 2019; Zhao et al., 2021; Cheng and Machin, 2022), media (van Dijk, 2021; Pei et al., 2022). This paper conducts the socio-cognitive discourse analysis, and the corpora in the present study contain 1.74 million words, randomly selected from China Court Network¹. A corpus-based method not only accounts for a much broader range of data than introspective approaches but also produces more exact results by mechanical retrieval (Stefanowitsch and Gries, 2006).

The corpus search and analysis methods are word frequency profile, concordances, semantic analysis, and linear regression model (LRM). Firstly, the expressions indicating the cognitive source and social function are first retrieved and listed by the corpus tool, and Wmatrix3.0 is employed as a corpus analysis tool for identifying these expressions. Wmatrix3.0 functions for automatic semantic analysis of words in discourses are based on the tagset in UCREL Semantic Analysis System (USAS) (Rayson, 2008). Then, the word frequency lists and the co-text in relevant concordances are examined regarding the cognitive and social expressions by ConcGram1.0, a phraseological search program used to generate single word frequency lists from the corpus (Cheng et al., 2009; Greaves, 2009). In the analysis of the corpora above, Wmatrix3.0 and ConcGram1.0 are the primary tools, supplemented by manual labeling and calculation. As the key variables, some discourse components with cognitive sources and social functions are manually summarized and shown in Table 1. Next, based on LRM, Stata16.0 is used to explore the possible quantitative relationships between the variables. LRM is a mathematical model to determine the correlation between the variables. Finally, the discourse of Chinese court judgments is interpreted at cognitive and social levels from the perspective of van Dijk’s Discourse–Cognition–Society.

Based on LRM, this paper uses stata16.0 to explore the possible quantitative relationships between discourse component, cognitive source, and social function, and the models are shown in Eqs. (1)–(4) for details:

$$\text{Cit} = \alpha + \beta_1\text{Fai} + \beta_2\text{Ind} + \beta_3\text{Par} + \beta_4\text{Inf} + \beta_5\text{Wor} + \beta_6\text{Phr} + \beta_7\text{Sen} \quad (1)$$

$$\text{Dep} = \alpha + \beta_1\text{Fai} + \beta_2\text{Ind} + \beta_3\text{Par} + \beta_4\text{Inf} + \beta_5\text{Wor} + \beta_6\text{Phr} + \beta_7\text{Sen} \quad (2)$$

$$\text{Dis} = \alpha + \beta_1\text{Fai} + \beta_2\text{Ind} + \beta_3\text{Par} + \beta_4\text{Inf} + \beta_5\text{Wor} + \beta_6\text{Phr} + \beta_7\text{Sen} \quad (3)$$

$$\text{Sum} = \alpha + \beta_1\text{Fai} + \beta_2\text{Ind} + \beta_3\text{Par} + \beta_4\text{Inf} + \beta_5\text{Wor} + \beta_6\text{Phr} + \beta_7\text{Sen} \quad (4)$$

The study takes the social functions (citation, depiction, distance, and summary) as the explained variables, and cognitive sources (faith, induction, paraphrase, and inference) and discourse components (vocabulary, phrases, and sentences) as

explanatory variables. The relationships between the four explained variables and seven explanatory variables are analyzed in the four models, and β_i is the influence coefficient and α is a constant item. The regression results are shown in Table 2, which displays the data relationship between the explained and explanatory variables in CCJ. In Model (1), the regression results indicate that Fai has significant positive effects on Cit, and Phr also has significant positive effects on Cit. In Model (2), Ind and Par have significant positive effects on Dep. In Model (3), Par significantly positively affects Dis. In Model (4), Inf has a significant positive impact on the Sum; also, Phr and Sen have a significant positive impact on the Sum. On the whole, different cognitive sources influence the corresponding social functions, while Phr and Sen in discourse components impact some social functions.

In order to avoid the co-linearity among the research variables and distortion of the regression model, this paper uses the variance inflation factor (VIF) to test the degree of multi-collinearity among the observed variables with the help of Spss23.0. When $VIF > 10$, it indicates high multi-collinearity among the variables. The test results are shown in Table 3, and the VIF of each variable is < 10 , indicating that there is no multi-collinearity problem among variables, and the model results are valid.

In order to further study the citation function, Models (5)–(7) indicate the possible relationship between the discourse components and the specific segments of law, regulation, and evidence in the cognitive source of faith, respectively. The regression results are shown in Table 4. In Models (5)–(7), as the primary knowledge sources of faith, Law, Reg, and Evi all have a significant positive effect on Cit, while Phr and Sen have a significant positive effect on Cit. In Models (8)–(10), the three explanatory variables as Law, Reg, and Evi are tested simultaneously, and the results are unchanged.

$$\text{Cit1} = \alpha + \beta_1\text{Law} + \beta_2\text{Wor} + \beta_3\text{Phr} + \beta_4\text{Sen} \quad (5)$$

$$\text{Cit2} = \alpha + \beta_1\text{Reg} + \beta_2\text{Wor} + \beta_3\text{Phr} + \beta_4\text{Sen} \quad (6)$$

$$\text{Cit3} = \alpha + \beta_1\text{Evi} + \beta_2\text{Wor} + \beta_3\text{Phr} + \beta_4\text{Sen} \quad (7)$$

$$\text{Cit1} = \alpha + \beta_1\text{Law} + \beta_2\text{Reg} + \beta_3\text{Evi} + \beta_4\text{Wor} + \beta_5\text{Phr} + \beta_6\text{Sen} \quad (8)$$

$$\text{Cit2} = \alpha + \beta_1\text{Law} + \beta_2\text{Reg} + \beta_3\text{Evi} + \beta_4\text{Wor} + \beta_5\text{Phr} + \beta_6\text{Sen} \quad (9)$$

$$\text{Cit3} = \alpha + \beta_1\text{Law} + \beta_2\text{Reg} + \beta_3\text{Evi} + \beta_4\text{Wor} + \beta_5\text{Phr} + \beta_6\text{Sen} \quad (10)$$

Next, vocabulary, phrases, and sentences in the discourse components are further subdivided to explore the quantitative relationship between different social functions and cognitive sources in CCJ. In Models (11)–(14), the social functions (citation, depiction, distance, and summary) are still taken as the explained variables, and there are ten explanatory variables: (1) Fps+V: First-person subject+verb; (2) Tps+V: third person subject+verb; (3) Ups+V: unknown person subject+verb; (4) H-ad: adverb with high confidence; (5) M-ad: adverb with medium confidence; (6) L-ad: adverb with low confidence; (7) Pre-P: prepositional phrase; (8) Ver-P: verb-object phrase; (9) Con-C: conditional clause; (10) Cau-C: causal clause. The data could reflect how discourse components influence social functions. Models (15)–(18) respectively indicate how the cognitive sources of faith, induction, paraphrase, and inference are affected by different discourse components. The regression results are shown

Table 1 Definition of key variables.

Variable	Definition
<i>Discourse components</i>	
Vocabulary	the words indicating the cognitive source and social function include verbs and adverbs in CCJ, among which the verbs consist of speech verbs, sensory verbs, cognitive verbs and modal verbs, and the adverbs are categorized with the degree of value
Fps+V	First-person subject + speech verb/sensory verb/cognitive verb/modal verb
Tps+V	Third person subject + speech verb/sensory verb/cognitive verb/modal verb
Ups+V	Unknown person subject + speech verb/cognitive verb/modal verb
H-ad	Adverbs with high confidence, which reflects the speaker's commitment to the authenticity of the information (Cheng and Sin, 2011; Wu and Cheng, 2020), such as "fundamentally, certainly, and affirmatively"
M-ad	Adverbs with medium confidence, such as "possibly, approximately, generally"
L-ad	Adverbs with low confidence, such as "likely, maybe, as if"
Phrase	Discourse is a hierarchical phenomenon consisting of the form, meaning and action (van Dijk, 2008), and through semantic analysis with Wmatrix3.0 and ConcGram1.0, the prepositional phrase (Pre-P) and verb-object phrase (Ver-P) are used to indicate the cognitive source and social function in CCJ.
Pre-P	Prepositional phrase, as "according to..."
Ver-P	Verb-object phrase, as "identify the fact that, determine the fact that..."
Sentence	Similar to Saussure's utterance, van Dijk's (2008) sentence is a part of social practices, and through semantic analysis with Wmatrix3.0 and ConcGram1.0, the causal clause (Cau-C) and a conditional clause (Con-C) are used to indicate the cognitive source and social function in CCJ
Con-C	A hypothetical presupposition is built on the conditional clause, and the result clause reflects the subjective speculation of the speaker after the achievement of the presupposition (Chen, 2009, p. 133)
Cau-C	the result clause puts a fact or makes an inference, and the reason clause provides evidence to support it (Chen, 2009, p. 137)
<i>Cognitive sources</i>	
Fai (faith)	The individual, institution, and social-cultural experience that is stored in and can be reproduced in a culture (Hu, 1994, p. 12). As the cognitive source of faith, there are mainly laws, regulations, and evidence in the court judgments (Wu and Cheng, 2020)
Law	The enforceable body of rules that govern any society, or one of the rules making up the body of law (Oxford Law Dictionary, p. 280)
Reg (regulation)	The normative documents are formulated and issued with administrative force and moral code to conduct the activities and behaviors of social organizations, enterprises, and institutions (from Administrative Laws and Regulations of China)
Evi (evidence)	It tends to prove the existence or non-existence of some fact, which may consist of testimony, documentary evidence, objective evidence, and hearsay evidence when admissible (Oxford Law Dictionary, p. 187)
Ind (induction)	A kind of reasoning process in the cognitive process, human makes further rational judgments through the direct perception of the sensory organs (Hu, 1995, p. 15)
Par (paraphrase)	The information source expressed by the speaker is not got through one's perception, but comes from another's utterance or writing (Hu, 1995, p. 18)
Inf (inference)	Inference is pushed from general to unique reasoning, and the speaker conducts reasoning and estimation based on known or existing information (De Haan, 2001). There are mainly conditional reasoning and causal reasoning in the court judgments (Wu and Cheng, 2020)
<i>Social functions</i>	
Cit (citation)	Citation from cultural belief plays an essential social function in a court judgment, since the speaker's personal experience is limited, and it is vital to quote other information to support one's ideas (Zhang and Yu, 2003)
Dep (depiction)	Sensory experience in court judgments refers to the objective information from a speaker who has personally experienced or witnessed (Wu and Cheng, 2020), which conducts a social function of depiction
Dis (distance)	The speech behavior that marks the other voice appearing in the discourse (Thompson, 1996), which means the speaker distances the quoted discourse away oneself through verbal retelling
Sum (summary)	In court judgments, the logical and reasonable judicial conclusion is derived from one or several known facts or assumptions (Wu and Cheng, 2020), and thus the result of the judgments is a summary of the social behavior

Table 2 Regression result of Model 1 to 4.

	(1) Cit	(2) Dep	(3) Dis	(4) Sum
Fai	0.07** (0.011)	-0.022 (0.014)	-0.001 (0.020)	0.002 (0.015)
Ind	0.016 (0.030)	0.088*** (0.028)	0.014 (0.041)	0.025 (0.031)
Par	0.041 (0.045)	0.040** (0.019)	0.135*** (0.028)	0.026 (0.021)
Inf	-0.548 (0.484)	-0.724 (0.460)	-0.797 (0.667)	0.091*** (0.007)
Wor	-0.465 (0.319)	0.107 (0.304)	-0.458 (0.441)	-0.359 (0.328)
Phr	1.312*** (0.325)	0.723 (0.709)	0.698 (0.448)	0.651* (0.334)
Sen	0.032 (0.019)	0.012 (0.018)	0.029 (0.027)	0.075*** (0.020)
_cons	1.243** (0.573)	-0.009 (0.545)	0.506 (0.790)	0.458 (0.588)
N	116	116	116	116
R ²	0.601	0.635	0.669	0.444

***p < 0.01, **p < 0.05, *p < 0.1.

in Table 5:

$$\text{Cit} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{11}$$

$$\text{Dep} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{12}$$

$$\text{Dis} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{13}$$

$$\text{Sum} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{14}$$

$$\text{Fai} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{15}$$

$$\text{Ind} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{16}$$

$$\text{Par} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{17}$$

$$\text{Inf} = \alpha + \beta_1\text{Fps} + \beta_2\text{Tps} + \beta_3\text{Ups} + \beta_4\text{H-ad} + \beta_5\text{M-ad} + \beta_6\text{L-ad} + \beta_7\text{Pre-P} + \beta_8\text{Ver-P} + \beta_9\text{Con-C} + \beta_{10}\text{Cau-C} \tag{18}$$

In Model (11), the regression results show that Pre-P and Cau-C significantly positively affect Cit. In Models (12) and (15), Fps and H-ad have significant positive effects on Dep and Ind, while Ups has significant adverse effects on Dep. Model (13) indicates Tps has significant positive effects on Dis, while Con-C has significant adverse effects on Dis. In Model (14), Pre-P, Ver-P and Cau-C all have significant positive effects on Sum, while Pre-P and Ver-P have significant positive effects on Fai in Model (15).

In Model (17), Tps, Ups, H-ad and M-ad all significantly positively affect Par. Model (18) shows Pre-P, Ver-P, Cau-C and Con-C all have significant positive effects on Inf. In general, the vocabularies marked in CCJ exert a specific influence on the cognitive source of paraphrase and the social function of depiction, and the phrases and sentences marked in CCJ have a particular impact on the cognitive source of inference and social function of summary.

Robustness analysis. In order to ensure the validity and applicability of the empirical model, this study changes the original regression method to use the Tobit method to test the robustness of the model. The robustness analysis of Models (1)–(4) is shown in Table 6, and that of Models (11)–(18) in Table 7. Data show that the significance of variables has not changed, which is consistent with the regression results in Tables 2 and 5.

Findings and analysis. CDA holds a “critical perspective, position or attitude within the discipline of multi-disciplinary Discourse Studies” (van Dijk, 2016, p. 62), focusing on various forms of the complex relations between social structures and discourse structures (van Dijk, 2018). This study draws upon van Dijk’s SCA to link discourse components to cognitive sources and social functions through a complex socio-cognitive interface.

Discourse dimension. On the one hand, similar to Saussure’s “utterance”, van Dijk’s (2008) discourse refers to the language in practical use, which is a complex and hierarchical phenomenon consisting of the form (vocabulary, phrase, and sentence), meaning, and action. On the other hand, as a part of social practices, discourse is constructively connected with the context and discourse subject, further developing Foucault’s (1995) discourse view. Through the statistical analysis of CCJ, the vocabularies indicating the cognitive source and social function include verbs and adverbs, among which the verbs consist of speech verbs, sensory verbs, cognitive verbs, and modal verbs, and the adverbs are categorized with the degree of value. The verb is the core of Chinese syntactic and semantic structure (Wu and Cheng, 2020). The general use of different verbs shows the cognitive source of the discourse subject and the attitude to the information mentioned. In the linguistic study of the Chinese Language, speech verb means “speaking”, that is, expressing meaning with words. “Speaking” is a collective concept, the most crucial subcategory in the semantic field of speech (Wang, 2004). The top 3 high-frequency speech verbs in CCJ are “shuo1 (say)” (1276 items), “cheng1/ sheng1cheng1 (express)” (741 items), and “zhu3zhang1 (claim)” (496 items). In Extract 1, witness Cui uses

Table 3 VIF test result.

variable	Fai	Ind	Par	Inf	Wor	Phr	Sen
VIF	1.973	1.161	1.996	1.431	2.120	1.460	2.042

Table 4 Regression result of Model 5 to 10.

	(5) Cit1	(6) Cit2	(7) Cit3	(8) Cit1	(9) Cit2	(10) Cit3
Law	0.024* (0.013)			0.020* (0.013)	0.030* (0.017)	-0.039 (0.026)
Reg		0.061*** (0.010)		-0.014 (0.008)	0.071*** (0.011)	-0.050 (0.017)
Evi			0.044*** (0.017)	-0.008 (0.008)	0.021* (0.011)	0.019*** (0.017)
Wor	-0.127 (0.149)	0.003 (0.205)	0.018 (0.341)	-0.133 (0.146)	0.051 (0.193)	-0.016 (0.300)
Phr	0.116 (0.162)	-0.023 (0.240)	1.132*** (0.373)	0.250 (0.176)	-0.132 (0.233)	1.572*** (0.362)
Sen	0.490** (0.226)	0.375 (0.313)	-0.719 (0.525)	0.384 (0.232)	0.480 (0.307)	-1.004 (0.477)
_cons	1.031*** (0.256)	0.094 (0.339)	0.917 (0.617)	1.133*** (0.284)	-0.388 (0.375)	1.361** (0.583)
N	116	116	116	116	116	116
R ²	0.285	0.601	0.494	0.366	0.679	0.642

***p < 0.01, **p < 0.05, *p < 0.1.

Table 5 Regression result of Model 11 to 18.

	(11) Cit	(12) Dep	(13) Dis	(14) Sum	(15) Fai	(16) Ind	(17) Par	(18) Inf
Fps	0.001 (0.043)	0.319** (0.138)	-0.009 (0.048)	-0.030 (0.037)	0.148 (0.137)	0.19*** (0.021)	0.006 (0.006)	-0.105 (0.108)
Tps	0.009 (0.104)	0.046 (0.069)	0.310** (0.116)	0.042 (0.089)	-0.168 (0.283)	0.006 (0.068)	0.327*** (0.043)	-1.119 (0.822)
Ups	-0.054 (0.120)	-0.149* (0.080)	0.057 (0.134)	-0.025 (0.103)	0.086 (0.338)	0.026 (0.081)	0.163** (0.051)	0.690 (0.984)
H-ad	-0.007 (0.009)	0.014** (0.006)	-0.007 (0.010)	-0.001 (0.007)	0.006 (0.024)	0.203*** (0.006)	0.001 (0.004)	0.024 (0.070)
M-ad	-0.018 (0.114)	0.048 (0.076)	-0.041 (0.128)	-0.086 (0.098)	0.160 (0.235)	-0.033 (0.056)	0.406*** (0.035)	0.826 (0.683)
L-ad	0.062 (0.090)	-0.012 (0.059)	0.081 (0.100)	0.077 (0.077)	-0.031 (0.032)	0.021 (0.074)	0.643*** (0.024)	-0.006 (0.008)
Pre-P	0.196*** (0.011)	-0.007 (0.008)	0.011 (0.014)	0.101** (0.031)	0.092*** (0.027)	-0.005 (0.061)	0.002 (0.020)	0.136** (0.042)
Ver-P	-0.002 (0.011)	0.007 (0.007)	-0.005 (0.013)	0.210*** (0.010)	0.079*** (0.021)	-0.003 (0.049)	-0.002 (0.016)	0.097* (0.040)
Con-C	-0.162 (0.146)	-0.172 (0.150)	-0.296** (0.118)	-0.043 (0.090)	-0.240 (0.216)	-0.806 (0.498)	-0.116 (0.160)	0.875*** (0.012)
Cau-C	0.314*** (0.024)	-0.018 (0.016)	-0.011 (0.027)	0.059*** (0.021)	-0.065 (0.050)	-0.054 (0.116)	-0.001 (0.037)	0.923*** (0.018)
_cons	2.608** (0.991)	1.145* (0.656)	1.308 (1.106)	1.286 (0.845)	0.790 (1.789)	-0.597 (0.707)	0.336 (0.455)	0.005 (0.005)
N	116	116	116	116	116	116	116	116
R ²	0.391	0.731	0.669	0.415	0.788	0.886	0.799	0.897

***p < 0.01, **p < 0.05, *p < 0.1.

Table 6 Robustness analysis of Model 1 to 4.

	(1) Cit	(2) Dep	(3) Dis	(4) Sum
Fai	0.125** (0.346)	-0.760 (-1.744)	-0.675 (-1.717)	-0.027 (-0.060)
Ind	-0.171 (-0.228)	2.652** (2.920)	0.426 (0.520)	0.139 (0.146)
Par	0.387 (0.770)	1.200* (1.977)	1.616** (2.952)	0.512 (0.803)
Inf	0.286 (0.587)	-0.101 (-0.172)	0.509 (0.962)	2.274** (3.683)
Wor	-4.291 (-0.534)	3.937 (0.405)	-14.937 (-1.705)	-4.203 (-0.411)
Phr	2.484** (3.036)	1.813 (1.835)	2.734 (3.067)	2.084* (2.006)
Sen	1.317 (1.081)	2.569 (1.747)	2.323 (1.751)	1.042*** (0.674)
_cons	-0.066 (-0.455)	0.088 (0.505)	0.107 (0.684)	-0.185 (-1.009)
N	116	116	116	116

Z values are in brackets.

“shuo1 (say)” to paraphrase Xu’s speech content, which means the information source expressed by the speaker comes from another’s utterance or writing (Hu, 1995), and thus the speech verb “said” marks the cognitive source of paraphrase. Sensory verbs obtain external information through human senses, and the top 2 high-frequency sensory verbs in CCJ are “kan4/kan4jian4/kan4dao4 (see)” (348 items) and “ting1/ting1jian4/ting1dao4 (listen)” (182 items). In Extract 2, the sensory verb “kan4dao4 (saw)” marks the cognitive source of induction, since human makes further rational judgments through the direct perception of the sensory organs (Hu, 1995). The cognitive verb is related to the psychological activities of human beings (Kellogg, 2003), and it reflects subjective human cognition, which expresses psychological activities or states in meaning, mainly including “thinking, feeling and knowing” (Wu and Cheng, 2020). The top 2 high-frequency cognitive verbs in CCJ are “ren4wei2 (hold)” (912 items) and “zhi1dao4 (know)” (336 items). In Extract 3, the court of the second instance recites the judicial view of “the court of the first instance”, and uses the cognitive verb “ren4wei2 (held)” to mark the cognitive source of paraphrase. The modality refers to the speaker’s commitment to the proposition (Katriel and Dascal, 1989), and modal verbs show the speaker’s attitude toward the possibility of the information (Le, 2014). The top high-frequency modal verbs in CCJ are “ying1gai1 (should)” (2512 items). In Extract 4, the modal verb “ying1gai1 (should)” combines with the prepositional phrase “gen1ju4 (according to)” to indicate the cognitive source of inference and belief, since “contract” is a kind of documentary evidence, and the speaker conducts inference based on the existing information (De Haan, 2001).

Extract 1 *Testimony of witness Cui: in the summer of 2008, Xu said (shuo1) that there was an office building project in a tourism service area of County Jing, and the cooperation had already been negotiated. If I paid the deposit, the contract could be signed.*

Extract 2 *Tao privately saw (kan4dao4) that Xue had more than 10 million yuan in a passbook.*

Extract 3 *The court of first instance held (ren4wei2) that: the defender of Jing, Xue, and Wang believes that the parties have conducted a civil lawsuit on the several facts of this case, and thus it should not be identified as a criminal case.*

Extract 4 *There should (ying1gai1) all be civil disputes, according to (gen1ju4) the contract signed by the appellant and Yang, the contract signed with Xishi Company, the contract signed with Hongda Company, the contract signed with the Nanjing branch of Shanghai Baoye Company, and the contract signed by Huahong Municipal Company.*

In SCA, the participant is an important category, mainly referring to individuals in discourse. The relationship among speakers, listeners, and other discourse participants is dynamically constructed in the social context (van Dijk, 2008). In CCJ, the syntactical

structure “third person subject+verb (Tps+V)” is generally used, accounting for 90.88% of all verb constructs that convey cognitive and social meaning, such as “the court of the first instance held” in Extract 3, which marks the information comes from a clear participant. The structure “first-person subject+verb (Fps+V)” accounts for 3.64%, as “wo3jue2de2 (I thought)” in Extract 5, yet the pronoun “I” does not refer to the court judge, but a direct reference to the testimony. The structure “unknown person subject+verb (Ups+V)” accounts for 5.48%, for example, “nong2hang2 peng2you3 (a staff from Agricultural Bank)” is an unknown person subject in Extract 6, and the predicate “da3 dian4hua4 jiang3(called to tell)” marks the cognitive source of paraphrase.

Extract 5 *On the night of April 29th, Lu volunteered to babysit for my son, but I thought (wo3jue2de2) it was the first day we met, so I refused.*

Extract 6 *The statement from victim Yang: after I had paid the deposit, a staff from Agricultural Bank called to tell (nong2hang2 peng2you3 da3dian4hua4jiang3) me that 500,000 RMB had been transferred, so I felt it was probably a cheat, and asked Jing to give the deposit back to the account to continue the project. Otherwise, he returned the deposit and canceled the project.*

Personal cognition refers to the understanding mode of social group members subjectively to the discourse, which is manifested as the unique psychological representations of specific situations, events, and actions (van Dijk and Kintsch, 1983). Some adverbs indicating degree can show the personal cognition of the speaker to the confidence in court judgments, which can be divided into types of high confidence as “surely”, “certainly”, middle confidence as “about”, “probably” and low confidence as “like”, “seemly” (Cheng and Sin, 2011). In CCJ, the top high-frequency adverb with high confidence is “ken3ding4 (certainly)” (36 items), accounting for 25% of all adverbs for cognitive and social meaning. In Extract 7, “ken3ding4 (indeed)” indicates the high probability of information in Zhou’s cognition. The top 3 high-frequency adverbs with middle confidence are “ke3neng2 (perhaps)” (124 items), “da4gai4 (probably)” (56 items), and “da4yue1 (about)” (32 items), accounting for 63.10%. “Ke3neng2 (probably)” in Extract 8 marks the cognitive source of inference with middle confidence. The top high-frequency adverb with low confidence is “hao3xiang4 (seemly)” (36 items), accounting for 11.90%. In Extract 9, the speaker uses “hao3xiang4 (seemly)” to show low confidence in the contract amount. In SCA, social context is a process of individual construction and perception of the communicative situation (van Dijk, 2008). The court judgments dynamically construct the relationship between the speakers, listeners, and other discourse participants through verb structures and adverbs.

Extract 7 *Zhou provided two receipts to the police office, which could prove the deposit was received indeed (ken3ding4).*

Table 7 Robustness analysis of Model 11 to 18.

	(11) Cit	(12) Dep	(13) Dis	(14) Sum	(15) Fai	(16) Ind	(17) Par	(18) Inf
Fps	0.566 (0.713)	0.648** (0.837)	0.647 (0.641)	-0.742 (-0.696)	0.698 (4.123)	0.161** (1.181)	-0.042 (-0.725)	-0.078 (-0.977)
Tps	-3.537 (-1.836)	4.600 (2.448)	2.993** (1.222)	2.958 (1.144)	0.175 (0.426)	-0.126 (-0.380)	0.285** (2.024)	-0.008 (-1.726)
Ups	1.062 (0.478)	-9.083* (-4.193)	-2.737 (-0.969)	-2.824 (-0.948)	-0.534 (-1.129)	0.285 (0.749)	0.197** (1.215)	-0.006 (-1.306)
H-ad	-0.124 (-0.781)	0.344** (2.224)	-0.207 (-1.026)	-0.117 (-0.549)	-0.025 (-0.742)	1.139*** (41.859)	-0.004 (-0.346)	0.015 (2.147)
M-ad	1.997 (0.944)	-0.002 (-0.001)	-3.824 (-1.423)	-2.975 (-1.048)	0.559 (1.241)	-0.277 (-0.763)	1.482** (9.580)	0.181 (1.886)
L-ad	2.147 (1.293)	-0.246 (-0.152)	3.093 (1.465)	1.636 (0.734)	-0.650 (-1.836)	0.268 (0.941)	0.036*** (0.294)	0.236 (1.865)
Pre-P	0.026 (0.111)	-0.134 (-0.592)	-0.087 (-0.296)	0.188** (0.603)	0.188** (3.800)	0.041 (1.039)	-0.006 (-0.340)	0.197** (2.770)
Ver-P	-0.138 (-0.665)	0.171 (0.844)	-0.073 (-0.275)	0.097*** (0.348)	0.126** (2.856)	-0.060 (-1.679)	0.018 (1.194)	0.074* (1.262)
Con-C	-0.059 (-3.000)	-0.042 (-2.213)	-0.063* (-2.538)	0.005 (0.188)	-0.005 (-1.184)	0.002 (0.608)	-0.002 (-1.455)	0.301*** (2.373)
Cau-C	0.008*** (1.735)	-0.008 (-1.780)	-0.006 (-0.974)	0.019*** (3.184)	-0.001 (-1.414)	-0.001 (-1.245)	0.000 (0.465)	0.261** (3.589)
_cons	0.175** (0.958)	0.349 (1.954)	0.182 (0.781)	-0.014 (-0.057)	-0.182 (-4.662)	-0.033 (-1.050)	0.012 (0.919)	0.091 (0.587)
N	116	116	116	116	116	116	116	116

Z values are in brackets.

Extract 8 *The witness Bo has a mental disorder, so her ability to testify is in doubt. All of her testimony is made in the death sentence's probation period, which may affect the testimony's authenticity because she is probably (ke3neng2) under some extraordinary pressure or for the sake of her meritorious service.*

Extract 9 *Wang's confession and defense: I remembered that I going Gaochun with Ma, Jing and a driver and signing a contract with a local company. The company seemly (hao3xiang4) paid a deposit of 200,000 or 300,000 RMB.*

Through the semantic analysis with Wmatrix3.0 and ConcGram1.0, the present study holds the syntactic structures as the prepositional phrase (Pre-P), verb-object phrase (Ver-P), the causal clause (Cau-C) and a conditional clause (Con-C) are used to indicate the cognitive source and social function in the Chinese court judgments. In CCJ, the top high-frequency Pre-P and Ver-P are “gen1ju4/an4zhao4 (in accordance with)” (5004 items) and “ren4ding4/ zheng4ming2...shi4shi2 (identify the fact that...)” (3880 items). In addition, Pre-P and Ver-P have significant positive effects on Fai and Inf in Model (15) and (18), and Pre-P accounts for 55.26% of the whole phrases that convey socio-cognitive meaning, but Ver-P for 44.74%. As a knowledge source of faith (Wu and Cheng, 2020), the law is the enforceable body of rules that govern any society (Martin, 2009, p. 280), and the Pre-P “gen1ju4 (in accordance with)” in Extract 10 cites the legal provisions of the PRC Criminal Law, marking the cognitive source of faith. In Extract 11, combined with “To sum up” and the evidence in that court judgment, the Ver-P “ren4ding4...shi4shi2 (identify the fact that)” conducts the estimation based on the known information (De Haan, 2001) with high confidence, indicating the cognitive source of inference.

In CCJ, the top 3 high-frequency Cau-C are “yin1/yin1wei2 (because...)” (1592 items), “gu4 (therefore...)” (740 items) and “yin1ci3 (hence...)” (363 items), which accounts for 87.49% of the total clauses with socio-cognitive meaning, and the top high-frequency Con-C is “ru2guo3 (if...)” (456 items), accounting for 12.51%. Cau-C is generally used in Chinese court judgments because the result clause puts a fact or makes an inference, and the reason clause provides evidence to support it (Chen, 2009, p. 137). In Extract 12, the “zheng4ju4 (evidence)” is provided after the investigation in the reason clause, clause conjunction “yin1ci3 (hence)” leads to the conclusion and marks the cognitive source of inference. In Con-C, a hypothetical presupposition is built on the conditional clause (Chen, 2009, p. 133). As the conditional conjunction “ru2guo3 (if)” is marked in Extract 13, the result clause reflects the reasonable speculation after achieving presupposition.

Extract 10 *This court made overall considerations about the harmful consequences to the criminal acts of defendant He, and He's attitude toward the admission of guilt. In accordance with (gen1ju4) paragraph 1 and paragraph 3 of Article 239, Article 52, Article 53, and Article 64 of the Criminal Law of the People's Republic of China, the judgment is as follows...*

Extract 11 *To sum up, it's enough to identify the fact that (ren4ding4...shi4shi2) the defendant He stole the infant Li to blackmail money from the parents. Besides, He's defense is inconsistent with the verified facts, so this court will not adopt it.*

Extract 12 *After investigation, the evidence in the case can confirm that the defendant Bo took advantage of his duties to help Tang and Xu's business, hence (yin1ci3) received the money from them.*

Extract 13 *If (ru2guo3) the registered trademark infringes upon the prior rights of others, the prior right holder may directly bring a civil lawsuit without the administrative procedure of revoking the trademark.*

The explicit language structure is inseparable from the implicit social relations since discourse structure in SCA is a logical

analysis aiming to make microscopic connections to the dimensions outside the language (van Dijk, 2009a), such as social structure, situation and cognitive structure. As a realistic legal carrier, the discourse dimension of Chinese court judgments is both society-oriented and cognition-oriented. Unlike Saussure's structuralism analysis, the present study analyzes the linguistic levels marking cognitive sources to establish a connection with the external world.

Cognitive dimension. From a cognitive perspective, CDA aims not only to describe the structural properties of text and talk, but also to account for how the cognitive, social, and political contexts influence the structures, strategies, and functions of text or talk (van Dijk, 2016). In the study of how language expresses a speaker's attitude, Chafe (1986) holds there are four cognitive modes: belief, induction, rumor, and deduction because the knowledge sources of humans are fuzzy or unknowable, evidence, language, and hypothesis. However, Hu (1994) believes any individual, institution, or social-cultural experience is eventually stored in the culture, and all experiences can be reproduced. so the knowledge sources are revised into culture, sense, language, and hypothesis in Chinese. Therefore, the expressions indicating knowledge sources as culture, sense, rumor, and hypothesis are retrieved and listed by Wmatrix3.0 in CCJ. Then the cognitive sources of Chinese court judgments are classified into four types: faith (Fai), induction (Ind), paraphrase (Par), and inference (Inf).

Firstly, as the individual and social-cultural experience is stored in and can be reproduced in a culture (Hu, 1994), Fai refers to the common sense of the public and also the ordinary cognition of a particular speech community. As the knowledge sources of Fai, there are mainly laws, regulations (Reg), and evidence (Evi) in the court judgments (Wu and Cheng, 2020), which provides the social context for the trial result. In SCA, Fai is a part of social cognition, which includes socially shared emotions, attitudes, ideology, and memory structure (van Dijk, 2003). Law is the enforceable body of rules that govern society (Martin, 2009). The regulations are formulated and issued to conduct the behaviors of social organizations in the Administrative Laws and Regulations of China. In Extracts 14 and 15, the basis of the court decision is derived from “fan3bu2zheng4dang1 jing4zheng1fa3 (Anti-Unfair Competition law)” and “cheng2zhen4 qi3ye4 zhi2gong1 yang3lao3bao3xian3 zhuan3yi2 zan4xing2ban4fa3 (Interim Measures of Basic Retirement Security for Employees of Urban Enterprises)”, which marks the cognitive sources of faith. Evidence proves the existence or non-existence of some fact (Martin, 2009), and “ji1dong4che1 jiao1tong1shi4gu4 qiang2zhi4bao3xian3 (the traffic insurance policy of the motor vehicle)” is documentary evidence in Extract 16, which is confirmed and adopted by the court, also indicating the cognitive sources of faith. Due to the plenty of citations of Law, Reg, and Evi, Pre-P and Ver-P have significant positive effects on Fai in Model (15). In addition, the articles of the law are cited for 4.81% of the total words, and regulation for 1.6%, but evidence for 15.68% in CCJ. Secondly, induction is the cognitive process of reasoning, especially when a human makes further rational judgments through the direct perception of the sensory organs (Hu, 1995). Because Fps and H-ad positively affect Ind in Model (16), the cognitive source of induction in CCJ is mainly manifested as visual and auditory perception in the first-person perspective, and the direct perception from the sensory organs is given high confidence with H-ad. Thus, the sensory verb and the first-person subject are the primary markers, as shown in Extract 17, combined with the first-person subject “wo3 (I)”, the auditory verb “ting1dao4 (heard)” and the visual verb “kan4dao4 (saw)” mark the cognitive source of induction.

Thirdly, paraphrase in court judgments means the information expressed by the speaker not through his feelings, but from another person's oral or written retelling (Wu and Cheng, 2020). Fairclough (1995) divides paraphrases into four types: direct paraphrase, free direct, indirect and paraphrase without marks, and Tps and Ups have significant positive effects on Par in Model (14), then the paper finds four sentence structures in CCJ: (1) information subject + verbal verb + direct paraphrase, and the direct paraphrase refers to the original statement and tone (Leech and Short, 1981); (2) information subject + verbal verb + indirect paraphrase, and the indirect paraphrase means that the information is expressed in the perspective of an interpreter in the current context (Leech and Short, 1981); (3) speaker + information subject + verbal verb + direct paraphrase; (4) speaker + information subject + verbal verb + indirect paraphrase, as in the case of “wo3ting1shuo1 (I heard from... that)” in Extract 18, and “a fellow villager” is the information subject, marking the cognitive source of paraphrase. Moreover, M-ad and L-ad have significant positive effects on Par in Model (17), which shows the speaker's low confidence in paraphrase. Fourthly, inference is neither personal experience nor hearsay, which is based on reasoning or hypothesis from the known information (De Haan, 2001), and Con-C and Cau-C have significant positive effects on Inf in Model (18), so there are mainly causal inferences and conditional inferences in CCJ. Causal inference is widely used in Chinese court judgments, mainly supporting the verdict conclusion and case facts. In Extract 19, “enjoy the joint rights” shows the causal reasoning of case fact, and “ju4ci3 (therefore)” indicates the cognitive source of inference. On the contrary, conditional inference is rarely used in CCJ, referring to the reasoning of the judicial procedure and case facts. For instance, the conditional clause “(ru2guo3) If you refuse to accept this judgment” is a hypothetical possibility about the judicial procedure in Extract 20.

Extract 14 *Item 3 in Article 5 of the PRC Anti-Unfair Competition Law (fan3bu2zheng4dang1 jing4zheng1fa3) stipulates that an operator shall not use another enterprise's name without authorization, which may mislead the consumers for that enterprise's commodity.*

Extract 15 *According to Article 3, Article 11, and Article 12 in Interim Measures of Basic Retirement Security for Employees of Urban Enterprises (cheng2zhen4 qi3ye4 zhi2gong1 yang3lao3-bao3xian3 zhuan3yi2 zan4xing2ban4fa3), which is jointly issued by Ministry of Human Resources and Social Security and Ministry of Finance, if the employee gets a job in different provinces, his or her insurance should transfer to the latest area.*

Extract 16 *The plaintiff accordingly provides evidence 1, that is, the traffic insurance policy of the motor vehicle (ji1dong4che1 jiao1tong1 shi4gu4 qiang2zhi4bao3xian3), in order to prove that the accident vehicle, semi-trailer tractor (plate number: Min D87263) driven by the defendant Yao, has insured the traffic insurance from the defendant Xiamen branch of Pingan Insurance.*

Extract 17 *On the second floor of my bathroom, I heard (wo3ting1dao4) a woman shout help outside the window, I ran to the lawn outside the wall, saw (kan4dao4) a woman naked, crying in the corner side.*

Extract 18 *I heard from (wo3ting1shuo1) a fellow villager that he lost 80,000 RMB in the supermarket casino.*

Extract 19 *On April 17, 2013, Niu signed a contract with Birdman Company, which agreed that Birdman Company would exclusively invest in the production and promotion of Niu's music albums. Therefore (ju4ci3), Niu and Birdman Company enjoy the joint rights of Niu's performing activities in the market.*

Extract 20 *If (ru2guo3) you refuse to accept this judgment, you may submit an appeal to Beijing Intermediate People's Court, and pay the court acceptance fee within 15 days from the date of the*

judgment. If (ru2guo3) you refuse to pay the fee within the time limit or are overdue, the appeal will be dismissed.

In SCA, social cognition is the mental structure used in the social context, which is shared by all members of social groups and cultures (van Dijk, 2003). The cognitive source of faith in CCJ mainly provides the legal provisions and judicial interpretation as the social background of judicial trial, accounting for 22.12% of the total discourse. However, personal cognition refers to the individual social group member subjectively producing and understanding the discourse, which is manifested as the unique psychological representations of specific situations, events, actions, and characters (van Dijk and Kintsch, 1983). The cognitive sources of induction and paraphrase are typical personal cognition, which proves and debates the testimony, documentary, or hearsay evidence between the sides of the prosecution. In CCJ, the discourse indicates induction and paraphrase separately account for 1.59% and 10.12%. The cognitive source of inference not only refers to the reasoning process in the debate but also provides confidence support for the judge to make a reasonable trial result, accounting for 11.71% in CCJ. Therefore, when personal cognition conforms to social consensus, it could change into social cognition in a court judgment. The study holds that social cognition has become a connection between individual ideas and social group attitudes, and the production of legal discourse is essentially the creation of social cognition.

Social dimension. van Dijk's social dimension is a context of discourse communication, which is the container of objective reality referred to by discourse (Cheng and Wu, 2019). The social dimension involves social and group relationships, especially discourse-related power relations. SCA defines power as manipulation (van Dijk, 2014a), and the representation of manipulation is typical in legal discourse (Cheng and Cheng, 2012). Court judgments, demonstrating power and control (Wagner and Cheng, 2011), apply the rules of law to the case facts in dispute and make judicial interpretation and decisions to the parties. Thus, the present study holds that court judgments achieve multiple social functions by manipulating the discourse components and cognitive sources. Firstly, citation plays an essential social function in CCJ since personal cognition is limited. It is vital to support judicial ideas with other information which could come from social consensus. It is also proved in the regression analysis that the cognitive source of Fai has significant positive effects on Cit in Model (1) in CCJ. In further analysis, Law, Reg, and Evi in Fai have significant positive effects on Cit in Model (5)–(7), and discourse components of Cau-C and Pre-P are primarily used in the citation of law and Evi separately in Model (1). Secondly, Ind and Par have significant positive effects on Dep in Model (2), because the information from a speaker's sensory experience (Wu and Cheng, 2020) and other's utterance or writing could both conduct a social function of depiction. Model (12) shows Fps and H-ad often describe the direct perception of speaker's sensory organs with high confidence, and Ups is the typical structure of Par. Thirdly, verbal paraphrase means the speaker marks the verbal behavior of another in the discourse (Thompson, 1996). In other words, the information provided by the speaker comes from another voice in a court judgment. By verbal paraphrasing, the speaker expresses the distance between oneself and the quoted voices (Luo, 2013), in order to separate oneself from the responsibility for the authenticity of the information in a court judgment and achieves the social function of distance. Besides, the quoted information in Cit is mainly the objective cultural truth, but that in Dis comes from the speaker's

sensory cognition. Models (3) and (13) show that Par's cognitive source has significant positive effects on Dis, while Tps and Con-C have significant positive effects on Dis in CCJ. Fourthly, logical and reasonable judicial decisions are derived from one or several known facts or assumptions (Wu and Cheng, 2020) in court judgments, and thus the result of the judgment is a summary of the social behavior. The data shows that Inf significantly positively affects Sum in Model (4). There are mainly conditional inferences and causal inferences in CCJ. The former starts from a hypothetical statement, then draw the conclusion from the premise, and the latter summarizes the general conclusions from a series of specific facts (Zhang and Yu, 2003); thus, the social function of summary makes a general explanation of the reasoning mentioned above. Corresponding to the reasoning process, Con-C and Cau-C significantly positively affect Sum. At the same time, Pre-P and Ver-P are often used in the reasoning process, which also significantly positively affects Sum.

At the micro level, the social dimension of SCA is realized in the discourse behavior of text and conversation, which is caused by the social group relations and their common cognition at the macro level (van Dijk, 2009b). Therefore, as a particular type of judicial discourse for legal professionals (Cheng and He, 2016), the social function of the court judgments correspond to the cognitive source and builds the surface structure with various discourse components. In CCJ, faith with citation function provides a legal basis for judgment analysis, and induction with depiction function offers the information of direct experience for evidence. Paraphrase with distance function transfers others' information to alleviate the legal responsibility for the authenticity of the information. In contrast, inference with summary function explains the reasoning process of court resolution.

Conclusion

SCA goes deep into the primary factors of language and society and plays a critical role in description and interpretation (Wu and Sun, 2019). From this perspective, this study aims to describe and explore the Chinese court judgments in a corpus method, concentrating on the interactions among the discourse, cognitive and social dimensions, which are deemed as the essential questions of CDA. This study finds that: (1) as a multi-dimensional social phenomenon, discourse needs to be explained reasonably in society, while human cognition is the intermediary factor (van Dijk, 2014b), and thus the discourse components could mark the cognition source in a court judgment. There are some verb structures such as Fps+V, Tps+V, and Ups+V, and the adverbs as H-ad, M-ad L-ad. Moreover, the phrases as Pre-P/Ver-P and sentences as Cau-C/Con-C are typical structures to indicate the cognition sources in CCJ. The relationship between discourse and cognition is the mapping relationship between language and psychology. (2) The cognitive source of Fai is a part of social cognition on Law and Reg, and that of Ind and Par provides the personal cognition to testimony, documentary, or hearsay evidence. In a court judgment, when personal cognition conforms to a social consensus (Xin and Liu, 2017), Inf could transfer the individual cognition to a part of the social cognition in the reasoning process. (3) Because of the opaque language and self-referential nature of the legal discourse (Cheng and Machin, 2022), the social functions of court judgments correspond to their cognitive sources and build the surface structure with various discourse components. The present study investigates the relationship between discourse, cognitive, and social dimensions in Chinese court judgments. However, further studies of the cognitive dimension are of critical value. In personal cognition, different audiences are ignored, such as the parties to the judgment case, legal practitioners, and the public. However, audiences'

understanding and acceptance are pivotal factors for court judgments. In other words, inter-semiotic operations are the core of interpreting court judgments (Cheng and Cheng, 2014). Moreover, most present studies on court judgments adopt a self-built corpus with insufficient data, which affects the effectiveness of results to a certain extent. Therefore, it is necessary to build a large shared corpus in the future to comprehensively and thoroughly examine the intersemiotic operation in court judgments.

Data availability

The datasets analyzed during the current study are available in the China Judgments Online:<https://wenshu.court.gov.cn>.

Received: 25 August 2022; Accepted: 13 December 2022;

Published online: 29 December 2022

Note

¹ <https://wenshu.court.gov.cn>.

References

- Alhazzawi D, Bamasag O, Albeshri A, Sana I, Ullah H, Asghar MZ (2022) Efficient prediction of court judgments using an LSTM+ CNN neural network model with an optimal feature set. *Mathematics* 10(5):683
- Bell D, Pether P (1998) Rewriting skill training in Law Schools-legal literacy revisited. *Legal Educ Rev*. <https://doi.org/10.53300/001c.6052>
- Bhatia VK (1993) *Analyzing genre: language use in professional settings*. Longman, London
- Bhatia VK (2004) *Worlds of written discourse: a genre-based view*. Continuum, London
- Chafe W (1986) Evidentiality in English conversation and academic writing. In: Chafe W, Nichols J (eds) *Evidentiality: the linguistic coding of epistemology*. Ablex, Norwood, NJ, pp. 261–272
- Cheng L (2007a) A discursive contrast of court judgments in mainland China and Hong Kong. In Gail Stygall (ed). *Proceedings of 8th Biennial Conference on Forensic Linguistics/Language and Law*, University of Washington, Seattle, 12–15 July 2007
- Cheng L (2007b) Identification and appraisal of intergenericity in judgments. *J Zhejiang Gongshang Univ* 83(2):32–37
- Cheng L (2010) A semiotic interpretation of genre: court judgments as an example. *Semiotica* 182:89–113. <https://doi.org/10.1515/semi.2010.053>
- Cheng L (2012) Attribution and judicial control in Chinese court judgments: a corpus-based study. *Int J Speech Language Law* 19(1):27–49. <https://doi.org/10.1558/IJSL.V19I1.27>
- Cheng L, Cheng W (2012) Legal interpretation: meaning as social construction. *Semiotica* 192:427–448. <https://doi.org/10.1515/sem-2012-0086>
- Cheng L, He LZ (2016) Revisiting judgment translation in Hong Kong. *Semiotica* 209:59–75. <https://doi.org/10.1515/sem-2016-0007>
- Cheng L, Machin D (2022) The law and critical discourse studies. *Crit Discourse Stud*. <https://doi.org/10.1080/17405904.2022.2102520>
- Cheng L, Sin KK (2011) A sociosemiotic interpretation of linguistic modality in legal settings. *Semiotica* 185:123–146
- Cheng L, Sin KK, Zheng YL (2008) Contrastive analysis of Chinese and American court judgments. *US-China Law Rev* 5:49–58
- Cheng L, Wu JJ (2019) Multi-dimensional interpretation of “Discourse-Cognition-Society” in Administrative Law enforcement. Paper presented at the Proceedings of 2019 youth academic forum on linguistic, literature, translation and culture, 8–16 May. The American Scholars Press, Georgia
- Cheng W, Greaves C, Sinclair J, Mc H, Warren M (2009) Uncovering the extent of the phraseological tendency: towards a systematic analysis of congrams. *Appl Linguist* 30(2):236–252. <https://doi.org/10.1093/applin/amm039>
- Cheng W, Cheng L (2014) Epistemic modality in court judgments: a corpus-driven comparison of civil cases in Hong Kong and Scotland. *English Specif Purp* 33:15–26. <https://doi.org/10.1016/j.esp.2013.07.006>
- Chen Y (2009) *Study on modern Chinese evidentiality*. China Social Sciences Press, Beijing
- De Haan F (2001) The relation between modality and evidentiality. *Linguist Ber* 9:201–216
- Fairclough N (1995) *Critical Discourse Analysis: The Critical Study of Language*. Longman, London/New York <https://doi.org/10.2307/329335>
- Fairclough N, Wodak R (1997) *Critical discourse analysis*. In: van Dijk (ed.) *Discourse studies: a multidisciplinary introduction*. Sage Publications, London

- Foucault M (1995) *Discipline and punish: The birth of the prison*. Vintage Books, New York
- Greaves C (2009) *ConcGram 1.0*. John Benjamins, Amsterdam
- Hart C (2010) Critical discourse analysis and cognitive science: new perspectives on immigration discourse. Palgrave, Basingstoke
- Hart C, Cap P (2014) Introduction. In: Hart C, Cap P (eds.) *Contemporary Critical discourse studies*. Bloomsbury, London, pp. 1–15
- Hu ZL (1994) Evidentiality in language. *Foreign Lang Teach Res* 1:9–15
- Hu ZL (1995) Chinese evidentiality and discourse analysis. *J Hubei Univ (Philos Soc Sci)* 2:13–23
- John G (1994) *Language and the Law*. New York: Longman Publishing.
- John G (2003) *Forensic Linguistics: An Introduction to Language in the Justice System*. Oxford: Blackwell Publishing Ltd.
- Kellogg RT (2003) *Cognitive psychology*, vol 2. SAGE, London
- Katriel T, Marcelo D (1989) Speaker's commitment and involvement in discourse. In: Yishai T (eds.) *From sign to text: a semiotic view of communication*. John Benjamins, Amsterdam, pp. 275–295
- Kurzon D (2001) The politeness of judges: American and English judicial behavior. *J Pragmat* 33:61–85. [https://doi.org/10.1016/S0378-2166\(99\)00123-X](https://doi.org/10.1016/S0378-2166(99)00123-X)
- Leech GN, Short M (1981) *Style in fiction: a linguistic introduction to English fictional prose*. Pearson Education, Harlow
- Le Y (2014) The overview of the modern Chinese evidentiality. *Stud Chin Lang* 2:27–34. <https://doi.org/10.2307/3508564>
- Luo GH (2013) *Discourse stance in court interaction*. Central China Normal University, Wuhan
- Martin EA (2009) *A dictionary of law*. OUP, Oxford
- Maley Y (1994) *Language and the Law*. In: John G (ed.) *Language and the Law*. Longman, London
- O'Hallora K (2003) *Critical discourse analysis and language cognition*. EUP, Edinburgh
- Pei JM, Li DN, Cheng L (2022) Media portrayal of hackers in China Daily and The New York Times: a corpus-based critical discourse analysis. *Discourse Commun*. <https://doi.org/10.1177/17504813221099190>
- Partington A, Duguid A, Taylor C (2013) *Patterns and meanings in discourse: theory and practice in Corpus-assisted Discourse Studies (CADS)*. John Benjamins, Amsterdam/Philadelphia
- Rayson P (2008) From key words to key semantic domains. *Int J Corpus Linguist* 13(4):519–549. <https://doi.org/10.1075/ijcl.13.4.06gray>
- Sandefur RL (2008) Access to civil justice and race, class, and gender inequality. *Annu Rev Sociol* 34(1):339–358. <https://doi.org/10.1146/annurev.soc.34.040507.134534>
- Seron C, Munger F (1996) Law and inequality: race, gender ... and, of course, class. *Annu Rev Sociol* 22(1):187–212. <https://doi.org/10.1146/annurev.soc.22.1.187>
- Solan LM (2010) *The language of judges*. University of Chicago Press. <https://doi.org/10.7208/9780226767895>
- Stefanowitsch A, Stefan G (2006) *Corpus-based approaches to metaphor and metonymy*. Mouton de Gruyter, Berlin
- Subtirelu NC, Baker P (2018) *Corpus-based approaches*. In: Flowerdew J, Richardson J (eds.) *The Routledge handbook of critical discourse studies*. Routledge, New York, pp. 106–119
- Sun Y, Cheng L (2017) Linguistic variation and legal representation in legislative discourse: A corpus-based multi-dimensional study. *Int J Legal Discourse* 2(2):315–339. <https://doi.org/10.1515/ijld-2017-0017>
- Teubert W, Krishnamurthy R (2007) *Corpus linguistics (critical concepts in linguistics)*. Routledge, London
- Titscher S, Meyer M, Wodak R, Vetter E (2000) *Methods of text and discourse analysis: in search of meaning*. Sage, London
- Thompson G (1996) Voice in the text: discourse perspectives on language reports. *Appl Linguist* 17:501–530. <https://doi.org/10.1093/applin/17.4.501>
- van Dijk TA, Kintsch G (1983) *Strategies of discourse comprehension*. Academic Press, New York
- van Dijk TA (1998) *Ideology: an interdisciplinary approach*. Sage, London
- Van Dijk, T., A. (2003) *Critical Discourse Analysis*. In D. Schiffrin, D. Tannen, & H. E. Hamilton (ed), *The Handbook of discourse analysis* (pp. 352–371). Maiden, MA: Blackwell.
- van Dijk TA (2008) *Discourse and context: a socio-cognitive approach*. Cambridge University Press, Cambridge
- van Dijk TA (2009a) *Society and discourse: how social contexts influence text and talk*. Cambridge University Press, Cambridge
- van Dijk TA (2009b) *Critical discourse studies: a sociocognitive approach*. *Methods Crit Discourse Anal* 2(1):62–86
- van Dijk TA (2014a) *Discourse and knowledge: a sociocognitive approach*. Cambridge University Press, London
- van Dijk TA (2014b) *Discourse-cognition: current state and prospects of the socio-cognitive approach to discourse*. In: Hart C, Cap P (eds.) *Contemporary studies in critical discourse analysis*. Bloomsbury, London, pp. 121–146
- van Dijk TA (2016) *Critical discourse studies: a sociocognitive approach*. In: Wodak R, Meyer M (eds.) *Methods of critical discourse studies*. SAGE, London, pp. 62–86
- van Dijk TA (2018) *Socio-cognitive discourse studies*. In: Flowerdew J, Richardson JE (eds.) *The Routledge handbook of critical discourse studies*. Routledge, Abingdon, pp. 26–43
- van Dijk TA (2021) *Antiracist discourse: theory and history of a macromovement*. Cambridge University Press, Cambridge
- Wagner A, Cheng L (2011) *Language, power and control in courtroom discourse*. In: Wagner A, Cheng L (eds.) *Exploring courtroom discourse: the language of power and control*. Ashgate, London, pp. 1–10
- Wang WH (2004) Diachronic evolution and synchronic distribution of “speaking words” in Chinese. *Stud Chin Lang* 4:329–342
- Wetter JC (1960) *The style of appellate judicial opinions: a case study in Comparative Law*. A. W. Sythoff, Leyden
- Wu JJ, Cheng L (2020) Evidentiality of court judgments in the people's Republic of China: a semiotic perspective. *Semiotica* 236–237:477–500. <https://doi.org/10.1515/sem-2020-0031>
- Wu JJ, Sun YX (2019) Study on utterance of administrative law enforcement in van Dijk's critical discourse analysis. *Int J Legal Discourse* 4(2):217–236. <https://doi.org/10.1515/ijld-2019-2024>
- Xin B, Liu C (2017) On van Dijk's socio-cognitive discourse analysis. *Foreign Lang Res* 5:14–19. <https://doi.org/10.16263/j.cnki.23-1071/h.2017.05.003>
- Yu W (2021) Reporting verbs in court judgments of the common law system: a Corpus-Based Study. *Int J Semiot Law-Rev Int Sémiot Jurid* 34(2):525–560. <https://doi.org/10.1007/s1196-020-09740-7>
- Zhao JF, Wu JJ, Yang Y (2021) A sociosemiotic exploration of medical legislation reform in China (1990–2021) *Int J Legal Discourse* 6(2):203–228. <https://doi.org/10.1515/ijld-2021-2054>
- Zhang FUC, Yu GW (2003) Study on Chinese evidentiality: take parentheses as an example. *Linguist Sci* 3:50–58

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies requiring ethical approval.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Le Cheng.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022