

Политическая лингвистика. 2024. № 1 (103).  
Political Linguistics. 2024. No 1 (103).

УДК 81'25:81'42  
ББК Ш118+Ш105.51

ГРНТИ 16.31.21; 16.31.41

Код ВАК 5.9.8

Чжу Хуэй<sup>1</sup>, Виктор Павлович Захаров<sup>2</sup>

<sup>1</sup> Даляньский университет иностранных языков, Далянь, Китай, zhuhui1230@qq.com, <https://orcid.org/0009-0003-2922-8156>

<sup>2</sup> Санкт-Петербургский государственный университет, Санкт-Петербург, Россия, vz1311@yandex.ru, SPIN-код: 5834-5012

## Корпусное сравнение языка китайских и российских политических текстов

**АННОТАЦИЯ.** Политические тексты имеют свои специфические лингвистические характеристики. В данной статье анализируются сходства и различия лингвистических характеристик переводных и оригинальных политических текстов на материале переводов китайских политических текстов на русский язык (переводные тексты) и сопоставляемых с ними русских текстов (оригинальные тексты). Исследование выполнено с помощью корпусных методов и инструментов. Выявлены общие лингвистические характеристики политических текстов, сравниваются различия в тематике китайских и российских политических текстов. Исследование показало, что переводные и оригинальные тексты различаются по лексическому разнообразию, средней длине предложения и частоте употребления частей речи. Было также обнаружено, что и переводные, и оригинальные тексты имеют более высокую среднюю длину слов и предложений по сравнению с текстами других сфер на русском языке. Выявляются дифференцирующие темы китайских и российских политических текстов. В китайских политических текстах ключевые тематические блоки — это китайская специфика, ликвидация бедности, сельское хозяйство, социальное обеспечение, а в российских текстах основными темами являются прирост населения, международная безопасность, военное строительство, экономический союз. Предполагается продолжение исследований по сравнению текста на исходном и целевом языке перевода с целью выявления универсалий, существующих при переводе политических текстов. Также перспективным видится объединение методов корпусной и контрастивной лингвистики, переводоведения и других дисциплин с целью выработки единого комплекса измерительных параметров и средств.

**КЛЮЧЕВЫЕ СЛОВА:** переводоведение, переводческая деятельность, политический дискурс, политические тексты, перевод политических текстов, китайский язык, русский язык, русские переводы, переводные тексты, оригинальные тексты, лингвистические характеристики.

**ИНФОРМАЦИЯ ОБ АВТОРАХ:** Чжу Хуэй, аспирант, Даляньский университет иностранных языков; 116044, China, Liaoning, Dalian; email: zhuhui1230@qq.com.

Захаров Виктор Павлович, кандидат филологических наук, доцент кафедры математической лингвистики, филологический факультет, Санкт-Петербургский государственный университет; ведущий научный сотрудник, Институт лингвистических исследований РАН; 199034, Россия, Санкт-Петербург, Университетская наб., 7–9; email: vz1311@yandex.ru.

**БЛАГОДАРНОСТИ.** Работа выполнена при поддержке научного проекта по теме «Стратегии перевода политических культурных концептов», утвержденного ДУИЯ, номер объекта: YJSCX2022-009.

(本文系大连外国语大学研究生创新项目YJSCX2022-009《习近平谈治国理政》<第三卷>俄译本中核心政治文化观念翻译策略研究的阶段性成果)。

**ДЛЯ ЦИТИРОВАНИЯ:** Чжу, Хуэй. Корпусное сравнение языка китайских и российских политических текстов / Чжу Хуэй, В. П. Захаров. — Текст : непосредственный // Политическая лингвистика. — 2024. — № 1 (103). — С. 115–128.

Zhu Hui<sup>1</sup>, Viktor P. Zakharov<sup>2</sup>

<sup>1</sup> Dalian University of Foreign Languages, Dalian, China, zhuhui1230@qq.com, <https://orcid.org/0009-0003-2922-8156>

<sup>2</sup> St. Petersburg State University, St. Petersburg, Russia, vz1311@yandex.ru, SPIN: 5834-5012

## A Corpus-Based Linguistic Comparison of Chinese and Russian Political Texts

**ABSTRACT.** Political texts have their own specific linguistic characteristics. This article analyzes the similarities and differences in the linguistic characteristics of translated and original political texts based on the material of translations of Chinese political texts into Russian (translated texts) and the Russian texts compared with them (original texts). The study employs corpus methods and tools. It reveals the common linguistic characteristics of political texts and compares the differences in the topics of Chinese and Russian political texts. The comparison has shown that translated and original texts differ in lexical diversity, average sentence length and frequency of use of parts of speech. It has been also found that both translated and original texts have on average longer words and sentences compared to Russian texts in other spheres. The analysis reveals the differen-

*tiating topics of Chinese and Russian political texts. The key thematic blocks in Chinese political texts include Chinese specificity, poverty eradication, agriculture, and social security; and in Russian texts, the main topics are population growth, international security, military construction, and economic union. The authors plan to continue research on comparing texts of the source and target languages of translation in order to identify universals that exist in the translation of political texts. The combination of the methods of corpus and comparative linguistics, translation studies and other disciplines in order to develop a single set of measuring parameters and tools seems to be a promising area of further investigation.*

**KEYWORDS:** translation studies, translation, political discourse, political texts, translation of political texts, Chinese language, Russian language, Russian translations, translated texts, original texts, linguistic characteristics.

**AUTHOR'S INFORMATION:** Zhuhui, Post-Graduate Student, Dalian University of Foreign Languages, Dalian, China.

Zakharov Viktor Pavlovich, Associate Professor of Department of Mathematical Linguistics, Philological Faculty, Leading Researcher of the Russian Academy of Sciences, Russia, St. Petersburg.

**ACKNOWLEDGMENTS.** The study has been accomplished with financial support of the 2022 scientific project of Dalian University of Foreign Studies “Translation strategies of political cultural concepts”, Project No. YJSCX2022-009.

**FOR CITATION:** Zhu Hui, Zakharov V. P. (2024). A Corpus-Based Linguistic Comparison of Chinese and Russian Political Texts. In *Political Linguistics*. No 1 (103), pp. 115-128. (In Russ.).

## ВВЕДЕНИЕ

Язык — основа развития человеческой цивилизации. Многообразие языков приводит к языковым преградам в межкультурном общении. В таких случаях особенно важна роль перевода. Языки в различных областях имеют свои собственные характеристики и нормы, например, политический язык, научно-технический язык и т. д. Как особая форма языка, политический язык имеет свои уникальные характеристики и особый стиль. Политический текст отличается актуальностью, информативностью и т. п. Изучая политический язык, мы можем лучше понять природу и характер политической коммуникации [Чудинов 2006].

В настоящее время большинство исследований, посвященных проблеме лингвистических особенностей политических текстов, носят качественный содержательный характер, а количественные исследования встречаются редко. Исследование лингвистических характеристик переводных политических текстов на основе корпуса требует сопоставления их с оригинальными (непереводными) политическими текстами на целевом языке. В данной статье анализируются лингвистические и тематические характеристики политических текстов русского языка на основе корпуса переводов Доклада о работе правительства Китая на русский язык и сопоставимого корпуса Посланий Президента Федеральному Собранию России.

### 1. ОБЗОР ЛИТЕРАТУРЫ

Языковые и дискурсивные характеристики политических текстов существенно отличаются от иных текстов. Политические тексты характеризуются формальностью и официальностью, что способствует созданию впечатления авторитетности [Fairclough 1989]. Политические тексты содержат боль-

шое количество лексики специализированных терминов [Hatim, Mason 1990], которые зачастую тесно связаны с политической системой каждой страны. Кроме того, политические тексты характеризуются сильной логикой и часто используют риторические приемы, такие как метафора, сравнение и т. д., чтобы передать сложные политические взгляды [Чудинов 2012]. Однако большинство этих исследований основано на оригинальных текстах и не рассматривают лингвистические характеристики переводных политических текстов.

Мона Бейкер (*Mona Baker*) первой отметила, что «с помощью корпуса можно изучать переводческие универсалии (*translation universals*)» [Baker 1993: 243–248]. С этого времени исследования языковых особенностей перевода на основе корпусов быстро развиваются во всем мире. Исследования сосредоточены на изучении языковых особенностей переводного английского языка и проверке наличия переводческих универсалий [Laviosa 1998; Mauranen 2000; Kenny 2000]. В Китае ученые использовали аналогичные методы исследования и обращали внимание на языковые характеристики переводного китайского и английского языков [Ху Сяньяо 2007; Ван Кэфэй, Цинь Хунву 2009; Ли Сяоцян, Ху Кайбао 2017; Лю Мяо 2018]. Языковым особенностям перевода китайских политических текстов на русский язык уделялось недостаточное внимание в Китае. Только отдельные ученые исследовали особенности лексики и синтаксиса в переводе политических текстов на русский язык [Цзян Ямин, Ли Мэнъя 2016; Ван Минъю, Лю Цзые 2016], однако в основном это были исследования на содержательном уровне, количественные исследования китайско-русского перевода с помощью корпусов по-прежнему редки.

## 2. ТЕОРЕТИЧЕСКАЯ БАЗА

### Характеристика и сложность текстов.

Понятие «характеристика текста» тесно связано с понятием «сложность текста» [Кисельников 2015; Ляшевская 2016]. В настоящее время исследования в области корпусной лингвистики, посвященные сложности русскоязычного текста, обращают особое внимание на лексические метрики и автоматизированную оценку сложности текста. Несмотря на то, что существует много инструментов для измерения уровня сложности русскоязычного языка, например, «Визуализация русского языка» [Clancy, 2014-2022], «Текстомер» [Лапошина 2021] и др., ученые до сих пор не пришли к единому мнению относительно лексической и грамматической сложности текста.

По мнению Л. Г. Флюдченко, понятие «сложность текста» — это фактор, на который влияет четыре группы параметров: информативность, сложность предложений, ясность структуры и абстрактность, от которых зависит качество восприятия и интерпретации текста [Федюченко 2010]. О. Н. Ляшевская связывает понятие «сложность текста» с индексами удобочитаемости. Она отмечает, что оценка сложности текста не зависит от сложности самого содержания, а связана с лингвистическими особенностями текста, его организацией, языковым опытом и мотивацией читателя [Ляшевская 2016]. А. С. Кисельников понимает текстовую сложность как текстовые признаки, вычисляемые на основе внутренних текстовых параметров, которые делятся на количественные и качественные [Кисельников 2015]. Что касается метрик сложности текста, для их описания чаще всего используются следующие конкретные параметры: коэффициент лексического разнообразия (TTR); количество букв, слов, предложений; средняя длина слова и предложения и т. д.

**Тематические параметры текста.** В исследованиях по корпусной лингвистике тематическая компрессия текстов (реферирование, выявление ключевых слов и словосочетаний) представляет собой важную тему исследований. Ключевое слово — это слово, называющее понятие, которому в тексте уделяется особое внимание. Оно часто выбирается в зависимости от вопросов и интересов исследователя. Ключевые слова должны отражать тему и основное содержание текста. Под ключевым словом и словосочетанием понимают слова и словосочетания, обладающие следующими характеристиками: важность для восприятия инфор-

мации в тексте, богатство содержания, способность структурировать и компактно хранить смысл текста, культурная значимость [Ванюшкин, Гращенко 2017]. В данном исследовании будет проведен сравнительный анализ ключевых слов и словосочетаний в переводах китайских политических текстов на русский язык и в российских политических текстах.

## 3. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

**Цели исследования.** Данное исследование направлено на решение следующих вопросов:

1) Каковы сходства и различия между лингвистическими характеристиками переводов китайских политических текстов на русский язык и российскими политическими текстами?

2) Каковы лингвистические характеристики политических текстов?

3) Чем отличается тематика китайских и российских политических текстов?

**Материал исследования и обработка данных.** В ходе исследования был самостоятельно составлен корпус «Доклады о работе правительства в 2012–2022 гг.» (переводы с китайского на русский язык; далее сокращенно именуется корпусом ДРП). Как политический документ «Двух сессий» Китая, ежегодный «Доклад о работе правительства» является важной частью китайских политических документов. Его содержание отражает состояние развития экономики, общества и других сфер Китая и имеет высокую политологическую ценность. Кроме того, «Доклады» дают нам большой материал для исследования состава политической лексики.

В данном исследовании в качестве сопоставимого корпуса будут использованы «Послания Президента Российской Федерации Федеральному Собранию РФ 2011–2021 гг.» (далее сокращенно именуемые корпусом ППР). Ежегодные «Послания Президента Российской Федерации Федеральному собранию» являются важным политическим документом в России, они представляют собой типичный политический текст и интересуют нас как материал для лингвистического исследования. Содержание ДРП и ППР сходно и в значительной степени сопоставимо.

В данном исследовании для обработки данных в основном используется корпусный менеджер *Sketch Engine*. Это мощный инструмент для автоматизированного анализа корпусов. Основные параметры корпусов ДРП и ППР приведены в таблице 1.

Таблица 1

Основные параметры (ДРП и ППР)

	Количество токенов	Количество лексических токенов	Количество словоформ	Количество лемм
ДРП	155130	133636	17630	7337
ППР	108324	89790	16665	7076

Таблица 2

Значения TTR (ДРП и ППР)

	Количество лексических токенов (words)	Количество лемм	Коэффициент лексического разнообразия (TTR)
ДРП	133636	7337	5,49 %
ППР	89790	7076	7,88 %

**Методы исследования.** Данное исследование, основанное на методологии корпусной лингвистики и переводоведения, посвящено изучению лингвистических особенностей ДРП и ППР в сопоставительном аспекте. Исследование проводится в соответствии со следующими этапами. Во-первых, с помощью корпусных менеджеров сравниваются основные параметры корпусов ДРП и ППР, анализируются характеристики двух корпусов с точки зрения TTR, средней длины слова, части речи и т. д. Во-вторых, с опорой на данные корпусного менеджера, анализируются ключевые слова и словосочетания ДРП и ППР, чтобы определить тематические характеристики китайских и российских политических текстов. По данным ключевых слов и словосочетаний мы можем обнаружить сходства и различия политической повестки дня в России и в Китае.

#### 4. РЕЗУЛЬТАТЫ АНАЛИЗА

##### 4.1. Характеристика и сложность текстов

###### 4.1.1. Коэффициент лексического разнообразия (TTR)

В корпусной лингвистике значение TTR — это самый простой показатель оценки богатства лексики, поэтому его также называют коэффициентом лексического разнообразия. В общем случае TTR рассчитывается путем деления общего количества словоформ (types) в корпусе на количество токенов (tokens), т. е.  $TTR = \text{types} / \text{tokens}$ . Под токеном можно понимать все элементарные вхождения, т. е. слово, знак препинания или число — это все токены (tokens). С другой

стороны, под токеном может пониматься лексический токен (в терминологии *Sketch Engine* это «words», последовательность символов, начинающаяся с буквы). При расчете TTR мы будем использовать последнее понимание токена. Нам представляется, что для измерения коэффициента лексического разнообразия русских текстов правильнее учитывать леммы, т. е. делить количество лемм на количество лексических токенов. Эти данные приведены в таблице 2.

Значение TTR для ППР выше примерно на два процентных пункта, чем в ДРП, что свидетельствует о том, что в политических текстах лексическое разнообразие переводных текстов ниже, чем оригинальных. Возможно, это связано именно с характером переводных текстов. Наша гипотеза заключается в том, что переводчик при переводе ограничен исходным текстом и часто использует метод дословного перевода, использует ближайший перевод эквивалент, и это приводит к меньшему лексическому разнообразию в переводном тексте.

###### 4.1.2. Средняя длина слова

Для измерения сложности текста часто используется параметр «средняя длина слова» и её стандартное отклонение. Стандартное отклонение относится к степени разницы между длиной каждого слова и средней длиной слова. С помощью инструмента *LancsBox*<sup>1</sup> были подсчитаны значения средней длины слова и стандартного отклонения длины слова для корпусов ДРП и ППР (таблица 3).

<sup>1</sup> <http://corpora.lancs.ac.uk/lancsbox>

Таблица 3

Значения средней длины слова и стандартного отклонение (ДРП и ППР)

Корпус	Средняя длина слова	Стандартное отклонение длины слова	Средняя длина русского слова <sup>1</sup>
ДРП	7.39	4.17	5.28
ППР	6.85	3.71	

Таблица 4

Значения средней длины предложения и стандартного отклонения (ДРП и ППР)

Корпус	Средняя длина предложения (в словах)	Стандартное отклонение длины предложения	Средняя длина предложения в русском языке <sup>2</sup>
ДРП	18.73	12.79	10.38
ППР	14.42	9.33	

Статистика показала, что средняя длина слова и стандартное отклонение длины слова ДРП и ППР почти одинаковы. Мы сравнили эти данные с данными С. А. Шарова о средней длине слова русского языка и обнаружили, что средняя длина слова в политических текстах больше, чем средняя длина слова русского языка в целом. Причина этого может заключаться в том, что в политических текстах встречается много производных слов. Например, *информатизация, производительность, продолжительность, профессиональное, стратегическое, функционирование, распределение* и т. д.

#### 4.1.3. Средняя длина предложения

Средняя длина предложения и стандартное отклонение длины предложения могут определять уровень сложности предложений в корпусе. Значения средней длины предложения и стандартного отклонения длины предложения ДРП и ППР приведены в таблице 4.

Видно, что средняя длина предложения и стандартное отклонение в ДРП больше, чем в ППР. На наш взгляд, причина этого заключается в том, что длина предложений переводных текстов зависит от синтаксической структуры исходного текста. Некоторые исследования показали, что в китайских политических текстах много длинных предло-

жений [Цзян Ямин, Ли Мэнья 2016], и при переводе переводчик пытается сохранить эквивалентность на синтаксическом уровне. Если сравнить эти данные со средней длиной предложения общего русского языка, видно, что предложения в политических текстах длиннее, что можно считать общей характеристикой всех политических текстов.

#### 4.1.4. Части речи

Частью речи называется категория слов языка, которая определяется морфологическими и синтаксическими признаками. Все языки имеют схожие категории слов, но в разных языках состав частей речи и деление на них могут отличаться. Представленность частей речи является одним из параметров измерения текстов. Используя статистические данные *Sketch Engine* и подсчитав проценты, мы получили данные, представленные в таблице 5. Там же мы приводим данные о частотности частеречных классов общего русского языка [Ляшевская, Шаров 2009]. Следует отметить, что данные по глаголам в таблице 5 включают глаголы, причастия и деепричастия; данные по прилагательным включают полные и краткие прилагательные и порядковые числительные; под местоимениями имеются в виду местоимение-существительное, местоимение-прилагательное и местоименное наречие.

Таблица 5

Процентное содержание частей речи (ДРП и ППР)

Части речи	ДРП		ППР		Общая русская лексика
	Частота	Доля(%)	Частота	Доля(%)	Доля(%)
Существительное	56399	41.60	29551	32.63	28.90
Прилагательное	24426	18.02	12933	14.28	9.00
Числительное	2184	1.61	1768	1.95	2.00
Глагол	17234	12.71	12084	13.34	16.70
Союз	9859	7.27	6356	7.02	7.70
Местоимение	5477	4.04	10807	11.93	14.70

<sup>1,2</sup> Данные получены от С. А. Шаров ([http://lingvisto.org/artikoloj/ru\\_stat.html](http://lingvisto.org/artikoloj/ru_stat.html))

Части речи	ДРП		ППР		Общая русская лексика
	Частота	Доля(%)	Частота	Доля(%)	Доля(%)
Наречие	5138	3.79	4553	5.03	5.10
Предлог	14158	10.44	10225	11.29	10.40
Частица	704	0.52	2278	2.52	4.60
Междометие	0	0	1	0.001	0.10

На основании этих данных мы можем сделать следующие выводы:

1) Процент существительных и прилагательных в переводных текстах значительно выше, чем в оригинальных, в то время как местоимения, частицы, наречия и предлоги используются в них реже, чем в оригинальных. Особо следует отметить, что местоимения в переводных текстах используются в примерах значительно реже, чем в оригинальных. Это говорит, видимо, о том, что по сравнению с оригинальным текстом переводной текст характеризуется большей ясностью (*explicitation*), т. е. переводчики вместо местоимения часто предпочитают использовать основную номинацию (само существительное), что в определенной степени также подтверждает гипотезу присутствия переводческой универсалии экспликации.

2) По сравнению с оригинальным текстом переводной текст содержит меньше частиц. В оригинальном тексте докладчик, вероятно, использует их, чтобы усилить выразительность речи и приблизить аудиторию. В переводном тексте эти слова используются редко, что может быть связано, с одной стороны, с тем, что они редко встречаются в исходном тексте, а с другой — с переводческим приемом переводчика.

3) Мы видим характерные особенности политических текстов (как переводных, так и оригинальных) по сравнению с представленностью частей речи в общей русской лексике. Существительные и прилагательные используются в политических текстах чаще, а глаголы, местоимения, союзы, частицы реже, чем в общем русском языке. Политическая лексика имеет более номинативный характер.

## 4.2. Тематические параметры текста

### 4.2.1. Ключевое слово

Ключевое слово (КС) — это слово, называющее понятие, которое может отражать тему и основное содержание текста. В корпусной лингвистике применяются методы автоматического выявления ключевых слов. Для этого требуется два корпуса: наблюдаемый корпус и справочный корпус. В этом исследовании наблюдаемыми корпусами являются ДРП и ППР, а справочным корпусом — Russian Web 2017 (ruTenTen17). Мы использовали функцию «Ключевое слово» системы *Sketch Engine* для извлечения ключевых слов. При этом исследователю необходимо определить критерий извлечения ключевых слов в терминах управляющего параметра «Focus on rare or common». При всех параметрах выдаются КС и ключевые словосочетания с повышенным коэффициентом «важности» («ключевости»), в основе которого лежит сравнение относительных частот (*ipm*) в данных корпусах. Мы пришли к выводу, что наиболее разумно установить критерий извлечения на уровне 100.

*Sketch Engine* сформировал список КС в соответствии с уровнем ключевости (*Score*), затем мы вручную обработали список, исключив характерные для наших корпусов слова, не имеющие, однако, тематического значения, такие как *миллион, уважаемый, всесторонне, важно, страна* и т. д., а также слова, имманентно присущие китайским и российским политическим текстам, например, в ДРП выкинули такие слова, как *Китай, КПК, ЦК, юань* и т. д., а в ППР — *Россия, рубль, Дума, коллега* и т. д. После этих процедур мы отобрали 50 первых КС из двух корпусов в порядке уровня ключевости (*Score*) и относительной частоты (*relative frequency*) от высокой к низкой (табл. 6 и 7).

Таблица 6

Фрагмент списка ключевых слов (ДРП и ППР)  
(*sort by score*)

	ДРП		ППР	
	КС	Score	КС	Score
1	реформа	31.44	экономика	6.92
2	совершенствовать	19.14	процент	6.39
3	усиливать	19.08	стратегический	5.60
4	стимулировать	18.03	здравоохранение	4.83

Окончание таблицы 6

	ДРП		ППР	
	КС	Score	КС	Score
5	сельский	17.60	сфера	4.80
6	экономика	12.77	инициатива	4.77
7	продвигать	12.40	демографический	4.71
8	углублять	11.54	модернизация	4.65
9	содействовать	10.28	потенциал	4.63
10	ускорять	9.51	региональный	4.33
11	стимулирование	8.54	ответственность	4.25
12	открытость	8.42	глобальный	4.18
13	страхование	8.25	семья	4.13
14	финансовый	8.22	доход	4.12
15	наращивать	8.17	ракета	4.08
	...	...	...	...

Таблица 7

Фрагмент списка ключевых слов (ДРП и ППР)  
(*sort by Relative Frequency*)

	ДРП		ППР	
	КС	RF	КС	RF
1	реформа	4280.28	регион	2150.95
2	экономика	2926.58	решение	1929.40
3	сельский	2868.56	экономика	1541.67
4	предприятие	2713.85	семья	1523.21
5	мера	2456.00	социальный	1403.20
6	усиливать	2404.44	сфера	1320.11
7	рост	2320.63	процент	1292.42
8	финансовый	2223.94	общество	1144.71
9	стимулировать	2217.50	рост	1135.48
10	экономический	2191.71	экономический	1080.09
11	совершенствовать	2017.66	бизнес	1070.86
12	контроль	1940.31	ответственность	867.77
13	строительство	1933.86	интерес	849.30
14	общественный	1676.01	доход	840.07
15	продвигать	1431.06	региональный	784.68
	...	...	...	...

Как видно, упорядоченные по силе *Score* и по относительной частоте списки явно пересекаются, что свидетельствует о высокой репрезентативности этих ключевых слов. С точки зрения частеречной принадлежности, ключевые слова ДРП и ППР представлены в основном существительными, прилагательными и глаголами. Что характерно, это большое число глаголов в качестве ключевых слов в китайском корпусе. Более того, многие ключевые слова образуют гнездо близких по семантике слов, относящихся к разным частям речи. Например, *стимулировать* / *стимулирование*, *усиливать* / *усилие*, *улучшать* / *улучшение*, *экономический* / *экономика* (ДРП); *вооружение* / *вооружен-*

*ный*, *экономический* / *экономика*, *регион* / *региональный* (ППР).

#### 4.2.2. Ключевые словосочетания

Когда мы говорим о ключевых словах, то на самом деле имеются в виду ключевые понятия. В текстах ключевые понятия очень часто выражаются словосочетаниями (КСС). В *Sketch Engine* имеются инструменты вычисления *N-gram*, выявления статистически значимых коллокаций (*Collocation*), лексико-синтаксических сочетаний (*Sketch*). В этом исследовании используется инструмент *Keywords*, который вместе с ключевыми словами выдает и ключевые словосочетания (*Terms*) (см. табл. 8 и 9).

Таблица 8

Фрагмент списка ключевых словосочетаний (ДРП)

ранг	КСС	Frequency	Score
1	сельское хозяйство	119	6.84
2	экономический рост	89	6.31
3	сектор экономики	85	6.22
4	китайская специфика	72	5.64
5	новейший тип	75	5.54
6	сельское население	63	4.99
7	свободная торговля	60	4.79
8	ликвидация бедности	57	4.68
9	реформа системы	57	4.66
10	основной вид	62	4.62
11	центральный бюджет	53	4.42
12	экономическое развитие	59	4.33
13	пилотный проект	48	4.09
14	зона свободной торговли	48	4.07
15	здоровое развитие	47	4.02
	...	...	...

Таблица 9

Фрагмент списка ключевых словосочетаний (ППР)

ранг	КСС	Frequency	Score
1	социальная сфера	33	3.81
2	Дальний Восток	25	2.87
3	материнский капитал	20	2.74
4	прожиточный минимум	18	2.56
5	жизнь людей	18	2.53
6	национальный проект	17	2.51
7	медицинская помощь	21	2.48
8	сельское хозяйство	23	2.46
9	качество жизни	18	2.46
10	правоохранительный орган	20	2.45
11	экономический рост	17	2.41
12	детский сад	22	2.30
13	гражданское общество	16	2.28
14	продолжительность жизни	15	2.26
15	особое внимание	20	2.25
	...	...	...

Эти ключевые словосочетания можно разделить на две основные категории: 1) КСС, которые характерны не столько для китайских и российских политических текстов, сколько для политической лексики в целом. К примеру, *новейший тип, основной вид, высочайший уровень (ДРП), особое внимание, новейший уровень, полнейшая мера (ППР)*. 2) КСС, которые относятся к фиксированным терминам или понятиям, имеющим фиксированные коннотации и значения в китайских и российских политических документах. Тот факт, что эти ключевые понятия либо совпадают, либо различаются в двух корпусах, связан с основными тематиками, представляющими интерес для китайского и российского правительств. Например, имеются КСС, общие для двух корпусов: *экономический рост, доход населения* и т. д.

Как примеры специфических сочетаний можно привести следующие: *китайская специфика, ликвидация бедности, среднезажиточное общество (ДРП), материнский капитал, продолжительность жизни, стратегическое оружие (ППР)*.

#### 4.2.3. Лексико-семантические группы КС

Инструмент *Thesaurus* в системе *Sketch Engine* автоматически создает для заданных КС списки синонимов или слов, относящихся к одной и той же категории (семантическому полю). Мы сформировали списки по 30 семантически близким словам для каждого слова из первой десятки базовых КС (существительных) и затем удалили те слова, которые в русском языке выражают слишком широкое значение, такие как *система, услуга, имя, наука, время* и т. д. Будем называть эти слова КС второго уровня или КС вторич-



ной номинации. Фактически их нужно рассматривать как продолжение списка КС. Сводный список КС второго уровня для ДРП и ППР см. в таблицах 10 и 11.

Сравнивая список КС (как первого, так и второго уровня) со списком КСС, можно заметить, что некоторые ключевые слова являются частью ключевых словосочетаний, например, *реформа — реформа системы, углубление реформ; сельский — сельское хозяйство, сельское население; экономика — сектор экономики, функционирование экономики (ДРП); ракета — противоракетная оборона, ракетный комплекс; экономический — евразийский экономический союз, экономическое пространство; рынок — внешний рынок, рынок труда (ППР).*

Возникает вопрос, что является идентификатором ключевых понятий. Однозначного ответа нет. Как известно, понятия часто находятся в отношении иерархического включения (род — вид, часть — целое и т. п.), причем понятия часто выражаются словосочетаниями. В данном исследовании мы решаем этот вопрос в пользу видовых понятий (терминов-сочетаний).

#### 4.2.4. Тематическая дифференциация ДРП и ППР

В ДРП и ППР встречаются как одинаковые, так и разные КС и КСС, что говорит о том, что эти два корпуса имеют как схожие, так и отличающиеся тематические компоненты (см. табл. 12, 13 и 14).

Таблица 10

Фрагмент списка ключевых слов второго уровня (ДРП)

КС	КС-II	Score	Frequency
бедность	эпидемия	0.19	69
	опора	0.18	25
	опустыниванием	0.16	10
	нищета	0.16	11
	загрязнение	0.16	88
	...	...	...
модернизация	трансформация	0.19	74
	регулирование	0.18	111
	реконструкция	0.15	51
	размещение	0.15	35
	интеграция	0.15	35
	...	...	...

Таблица 11

Фрагмент списка ключевых слов второго уровня (ППР)

КС	КС-II	Score	Frequency
доход	инфляция	0.16	32
	пенсия	0.15	24
	продолжительность	0.14	17
	ограничение	0.14	34
	образование	0.14	109
	...	...	...
модернизация	обновление	0.19	43
	укрепление	0.15	39
	независимость	0.12	16
	доля	0.11	17
	компания	0.11	84
	...	...	...

Таблица 12

Одинаковые ключевые лексики (ДРП и ППР)

Ранг в ДРП	КС и КСС	Ранг в ППР
6	экономика	1
17	экономический	17
20	рост	19
21	научно-технический	24
23	модернизация	8
45	общественный/общество	29

Окончание таблицы 12

Ранг в ДРП	КС и КСС	Ранг в ППР
46	потенциал	9
47	доход	14
20	жизнь народа	5
42	доход населения	30
1	сельское хозяйство	8
17	темп роста	22
2	экономический рост	11
12	экономическое развитие	44
27	медицинская помощь	7

Таблица 13

Отличающиеся ключевые слова (ДРП и ППР)

Ранг в ДРП	КС	Ранг в ППР	КС
1	реформа	2	процент
2	совершенствовать	3	стратегический
3	усиливать	4	здравоохранение
4	стимулировать	5	сфера
5	сельский	6	инициатива
7	продвигать	7	демографический
8	углублять	10	региональный
9	содействовать	11	ответственность
10	ускорять	12	глобальный
11	стимулирование	13	семья
12	открытости	15	ракета
13	страхование	16	инфраструктура
14	финансовый	18	бизнес
15	наращивать	20	приоритет
	...		...

Таблица 14

Отличающиеся ключевые словосочетания (ДРП и ППР)

Ранг в ДРП	КСС	Ранг в ППР	КСС
3	сектор экономики	1	социальная сфера
4	китайская специфика	2	Дальний Восток
5	новейший тип	3	материнской капитал
6	сельское население	4	прожиточный минимум
7	свободная торговля	6	национальный проект
8	ликвидация бедности	9	качество жизни
9	реформа системы	10	правоохранительный орган
10	основной вид	12	детский сад
11	центральный бюджет	13	гражданское общество
13	пилотные проект	14	продолжительность жизни
14	зона свободной торговли	15	особое внимание
15	здоровое развитие	16	новейший уровень
16	удельный вес	17	экономический союз
18	функционирование экономики	18	полнейшая мера
19	социально-экономического развитие	19	политическая система
	...		...

Как видно из таблицы, одинаковые КС и КСС в ДРП и ППР имеют разные ранги (т. е. разные коэффициенты значимости). Это также нужно учитывать, описывая приоритеты в китайской и российской политике.

На основе данных можно выявить различия в их тематиках. В ДРП представлено множество концепций развития, таких как *китайская специфика*, *полное построение среднезажиточного общества* и т. д. Кроме

того, в ДРП также подчеркиваются такие тематики, как *региональное развитие, урбанизация, защита окружающей среды* и т. д. В то же время ППР, с одной стороны, уделяет внимание *рождаемости, приросту населения, мерам по поддержке семьи и детей*, а с другой стороны, в ППР активно обсуждаются вопросы безопасности государства, причем делается акцент на таких темах, как *международная безопасность, стратегическое оружие, ракетный комплекс* и т. д.

#### 4.2.5. Ключевая тематика ДРП и ППР

В ходе проведенного корпусного исследования мы попытались провести кластеризацию КС и КСС, с помощью методов искусственного интеллекта и вручную выделить родовые понятия для каждой группы. В государственной политике двух стран имеются совпадения и различия, и это хорошо видно из нижеследующих таблиц (см. таблицу 15, 16).

Таблица 15

Отличающиеся ключевые лексиксы

ДРП	ППР
<b>1. Государственное управление и регулирование Инновация и инвестиция</b>	
автономия актив единое планирование инновационная деятельность концепция развития коррупция модернизация структурное регулирование центральный бюджет	бюджет деловой климат инновационный институт развития инфраструктура материнский капитал ресурс правоохранительные органы
<b>2. Экономическое развитие</b>	
предпринимательство строительство ускоренный темп функционирование экономики сектор экономики	бизнес Дальний Восток национальный проект прибыль промышленность прорыв регион страны руководитель регионов территория фонд
<b>3. Военное дело и внешняя политика</b>	
	авиация армия внешняя политика вооружение вызов международная безопасность международное право противоракетная оборона стратегическое оружие ядерный
<b>4. Здоровье и экология</b>	
выбросы загрязнение здоровое развитие медицинское страхование опустынивание охрана здоровья	медицинская помощь

ДРП	ППР
<b>5. Социальная сфера и благосостояние населения</b>	
беднейший район безработица городское население имущество инвалидность лишение нужда среднезажиточное общество поощрение социальное обеспечение справедливость трудовой ресурс трудоустройство	возраст демографический естественный прирост населения зарплата пенсия поддержка семей прожиточный минимум производительность труда рынок труда семья эпидемия
<b>6. Торговля и международные отношения</b>	
внешняя торговля внутренний спрос зона свободной торговли международные переговоры мир сотрудничество	внешний рынок евразийский экономический союз таможенный союз экспорт
<b>7. Сельское хозяйство</b>	
земледелие сельское население сельскохозяйственный угодье	сельское хозяйство
<b>8. Общество, образование, молодежь</b>	
	вуз детский сад культура молодежь образование профессиональное образование школьник

Таблица 16

Совпадающие ключевые лексики

ДРП	ППР
<b>2. Экономическое развитие</b>	
	прирост потенциал производство предприятие хозяйство темп роста рынок экономический рост
<b>4. Здоровье и экология</b>	
	здравоохранение экология
<b>5. Социальная сфера и благосостояние населения</b>	
	бедность доходы населения благосостояние занятость

В данном исследовании мы выделили восемь основных тематик, как показано в таблицах. Можно заметить, что по каждой теме в ДРП и ППР могут быть как одинаковые, так и разные единицы ключевой лексики, что может показать приоритеты двух стран в различных областях. Например, в тематике «*Экономическое развитие*» в ДРП больше внимания уделяется предпринимательству и строительству, а в ППР — региональному развитию и промышленности. Кроме того, не все тематики заметно выделены в обоих корпусах, например, в ДРП тематики 3 и 8 не являются ключевыми, а в ППР — тематика 4 и 7.

Основываясь на этих данных, мы можем обобщенно представить типичную лексику в политических текстах: *восстановление, обновление, потребление, ускорение, сила, структура, организация, решение, достижение, наращивание, укрепление, категория, улучшение, увеличение, программа, стратегия, новейший уровень, высочайший уровень* и т. д.

### ЗАКЛЮЧЕНИЕ

Преимущество корпусных исследований заключается в том, что выводы, основанные на масштабных корпусах и статистических методах, более объективны. Наше исследование показало, что политические тексты, как переводные, так и непереводные, характеризуются повышенной сложностью по сравнению с общим русским языком. Лексическое разнообразие переводных текстов ниже, чем у непереводных текстов, а сложность более высокая. Процент существительных и прилагательных в переводных текстах значительно выше, а в непереводных текстах выше процент местоимений, наречий, предлогов и частиц. Исследование влияния входного языка на целевой переводной должно быть продолжено.

Сравнение тематики двух корпусов показало как сходство, так и различия. Государственное управление, международные отношения, экономическое развитие и благосостояние населения являются наиболее важными родовыми понятием в обоих корпусах, но их конкретизация в нижележащих видовых терминах демонстрирует различия в инициативах и мерах. Сельское хозяйство — важная тематика в ДРП, в то время как военное дело, внешняя политика, образование и молодежь являются более важными, ключевыми тематиками в ППР.

Данное исследование имеет положительное значение для развития корпусных исследований на материале русского языка в Китае. Проблему лингвистических харак-

теристик политических текстов предполагается в дальнейшем решать исследованием целевого переводного языка в сравнении его с исходным языком и оригинальным целевым с целью выявления при переводе политических текстов универсалий перевода. Будущее направление исследований должно быть нацелено на объединение методов корпусной и контрастивной лингвистики, переводоведения и других дисциплин с целью выработки единого комплекса измерительных параметров и средств.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ванюшкин, А. С. Оценка алгоритмов извлечения ключевых слов: инструментарий и ресурсы / А. С. Ванюшкин, Л. А. Гращенко. — Текст : электронный // Новые информационные технологии в автоматизированных системах. — 2017. — № 20. — URL: <https://cyberleninka.ru/article/n/otsenka-algoritmov-izvlecheniya-klyuchevykh-slov-instrumentariy-i-resursy>.
2. Кисельников, А. С. К проблеме характеристик текста: читабельность, понятность, сложность, трудность / А. С. Кисельников. — Текст : непосредственный // Филологические науки. Вопросы теории и практики. — 2015. — № 11 (53). — С. 79–84.
3. Лапошина, А. Н. Текстометр: онлайн-инструмент определения уровня сложности текста по русскому языку как иностранному / А. Н. Лапошина, М. Ю. Лебедева. — Текст : непосредственный // Русистика. — № 3 (19). — С. 331–345.
4. Ляшевская, О. Н. Индексы удобочитаемости как мера оценки сложности текста / О. Н. Ляшевская. — 2016. — URL: <https://ling.hse.ru/data/2016/12/15/1111563794/Readability%20ta lk.pdf>. — Текст : электронный.
5. Федюченко, Л. Г. Когнитивный подход в описании уровня сложности текста на перевод / Л. Г. Федюченко. — Текст : электронный // Lingua mobilis. — 2010. — № 2 (21). — С. 170–175. — URL: <https://cyberleninka.ru/article/n/kognitivnyy-podhod-v-opisanii-urovnya-slozhnosti-teksta-na-perevod>.
6. Чудинов, А. П. Дискурсивные характеристики политической коммуникации. / А. П. Чудинов. — Текст : непосредственный // Политическая лингвистика. — 2012. — № 2 (40). — С. 53–59.
7. Чудинов, А. П. Политическая лингвистика / А. П. Чудинов. — Москва : Флинта : Наука, 2006. — 254 с. — Текст : непосредственный.
8. Baker, M. Corpus linguistics and translation studies: Implications and applications / M. Baker. — Text: unmediated // Text and Technology: In Honour of John Sinclair / M. Baker, G. Francis, E. Tognini-Bonelli (eds). — Amsterdam/Philadelphia : John Benjamins, 1993. — P. 233–250.
9. Clancy, S. Visualizing Russian. A suite of tools for exploring the data behind the Russian language, 2014–2022. — URL: <https://visualizingrussian.fas.harvard.edu/> (date of access: 10.10.2023). — Text : electronic.
10. Fairclough, N. Language and Power / N. Fairclough. — London : Longman, 1989. — Text : unmediated.
11. Hatim, B. Discourse and the Translator / B. Hatim, I. Mason. — London : Longman, 1990. — Text : unmediated.
12. Kenny, D. Translators at play: Exploitations of collocational norms in German-English translation / D. Kenny. — Text : unmediated // Working with German Corpora / B. Dodd (ed.). — Birmingham : University of Birmingham Press, 2000. — P. 143–160.
13. Laviosa, S. Core patterns of lexical use in a comparable corpus of English narrative prose / S. Laviosa. — Text : unmediated // Meta. — 1998. — № 4. — P. 557–570.
14. Mauranen, A. Strange strings in translated language: A study on corpora / A. Mauranen. — Text: unmediated // Intercultural Faultlines. Research Methods in Translation Studies I: Textual and Cognitive Aspects / M. Olohan (ed.). — Manchester : St Jerome, 2000. — P. 119–141.
15. Ван, Кэфэй. Обсуждение языковых особенностей англо-китайского языка — макроанализ на основе соответствующе-

го корпуса / Ван Кэфэй, Цинь Хунву. — Текст : непосредственный // Журнал иностранных языков — 2009. — № 1. — С. 102–105.

16. Ван, Минью. Анализ частотности слов и перевод китайско-русской политической текстов — сравнение русской версии Доклада о работе правительства Китая за 2013–2014 годы и Послания Президента Федеральному собранию России / Ван Минью, Лю Цзые. — Текст : непосредственный // Вестник Тяньцзиньского университета иностранных языков. — 2016. — № 3. — С. 1–6+80.

17. Ли, Сяоянь. Исследование ключевых слов и их сочетаний в английском переводе Доклада о работе правительства Китая / Ли Сяоянь, Ху Кайбао. — Текст : непосредственный // Иностранные языки в Китае. — 2017. — № 6. — С. 81–89.

18. Лю, Мяо. Исследование лингвистических особенностей художественного перевода на основе русско-китайского параллельного корпуса — на примере переводного языка и деепричастного оборота / Лю Мяо. — Текст : непосредственный // Русский язык в Китае. — 2018. — № 2. — С. 28–37.

19. Ху, Сяньяо. Исследование лексических характеристик переведенных на китайский язык романов на основе корпуса / Ху Сяньяо. — Текст : непосредственный // Преподавание иностранных языков и научные исследования. — 2007. — № 3. — С. 214–220+241.

20. Цзян, Ямин. Анализ структурных характеристик предложений в русском переводе политических текстов — на примерах русского перевода Доклада о работе Правительства Китая / Цзян Ямин, Ли Менгья. — Текст : непосредственный // Вестник Тяньцзиньского университета иностранных языков. — 2016. — № 3. — С. 7–12+80.

#### REFERENCES

1. Vanyushkin, A.S. & Grashchenko, L.A. (2017). Evaluating keyword extraction algorithms: tools and resources [Otsenka algoritmov izvlecheniya klyuchevykh slov: instrumentariy i resursy]. *New information technologies in automated systems*, 20. Retrieved from <https://cyberleninka.ru/article/n/otsenka-algoritmov-izvlecheniya-klyuchevykh-slov-instrumentariy-i-resursy> (In Russ.)

2. Kisel'nikov, A.S. (2015). K probleme kharakteristik teksta: chitabel'nost', ponyatnost', slozhnost', trudnost [To the Issue of Text Characteristics: Readability, Understandability, Complexity, Difficulty]. *Philological Sciences. Issues of theory and practice*, 11(53), 79–84. (In Russ.)

3. Laposhina, A.N., & Lebedeva, M.Yu. (n.d.). Tekstometr: onlayn-instrument opredeleniya urovnya slozhnosti teksta po russkomu yazyku kak inostrannomu [Textometer: an online tool for determining the level of text complexity in Russian as a foreign language]. *Russistics*, 3(19), 331–345. (In Russ.)

4. Lyashevskaya, O.N. (2016). Indeksy udobochitaemosti kak mera otsenki slozhnosti teksta [Readability indices as a measure of text complexity]. Retrieved from <https://ling.hse.ru/data/2016/12/15/1111563794/Readability%20talk.pdf> (In Russ.)

5. Fedyuchenko, L.G. (2010). Kognitivnyy podkhod v opisani urovnya slozhnosti teksta na perevod [Cognitive approach to describing the level of complexity of a text to be translated]. *Lingua mobilis*, 2(21), 170–175. Retrieved from <https://cyberleninka.ru/article/n/kognitivnyy-podhod-v-opisani-urovnya-slozhnosti-ostit-teksta-na-perevod> (In Russ.)

6. Chudinov, A.P. (2012). Diskursivnye kharakteristiki politicheskoy kommunikatsii [Discursive characteristics of political communication]. *Political Linguistics*, 2(40), 53–59. (In Russ.)

7. Chudinov, A.P. (2006). *Politicheskaya lingvistika* [Political Linguistics]. Moscow: Flinta: Science, 254 p. (In Russ.)

8. Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis & E. Tognini-Bonelli (eds). In *Text and Technology: In Honour of John Sinclair* (pp. 233–250). Amsterdam/Philadelphia: John Benjamins. 233–250.

9. Clancy, S. (n.d.). *Visualizing Russian. A suite of tools for exploring the data behind the Russian language, 2014–2022*. Retrieved Nov. 10, 2023, from <https://visualizingrussian.fas.harvard.edu/>

10. Fairclough, N. (1989). *Language and Power*. London: Longman.

11. Hatim, B. & Mason, I. (1990). *Discourse and the Translator*. London: Longman.

12. Kenny, D. (2000). Translators at play: Exploitations of collocational norms in German-English translation. In B. Dodd (ed.), *Working with German Corpora* (pp. 143–160). Birmingham: University of Birmingham Press.

13. Laviosa, S. (1998). Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta*, 4, 557–570.

14. Mauranen, A. (2000). Strange strings in translated language: A study on corpora. In M. Olohan (ed.), *Intercultural Faultlines. Research Methods in Translation Studies I: Textual and Cognitive Aspects* (pp. 119–141). Manchester: St Jerome.

15. Wang, Kefei, & Qin, Hongwu (2009). A Parallel Corpus-based Study of General Features of Translated Chinese. *Foreign Language Research*, 1, 102–105.

16. Wang, Mingyu, & Liu, Ziyi (2016). Word Frequency Statistics and the Chinese-Russian Translation of Political Literature. *Journal of Tianjin Foreign Studies University*, 3, 1–6+80.

17. Li, Xiaoqian, & Hu, Kaibao (2017). Keywords and Their Collocations in the English Translations of Chinese Government Work Reports. *Foreign Language in China*, 6, 81–89.

18. Liu, Miao (2018). Research on the Linguistic Characteristics of Literary Translation Based on Russian-Chinese Parallel Corpus. *Russian in China*, 2, 28–37.

19. Hu, Xianyao (2007). A Corpus-based study on the lexical features of Chinese translated fiction. *Foreign Language Teaching and Research*, 3, 214–220+241.

20. Jiang, Yaming, & Li, Mengya (2016). An Analysis of the Characteristics of the Sentence Structure of the Russia Version of CPC Literature Based on the Russia Translation of Report on the Work of the Government. *Journal of Tianjin Foreign Studies University*, 7–12+80.