***П. Чжан***

**Корпусный анализ лексического наполнения концепта «государство» в русской и китайской языковых картин мира**

**Аннотация.** Концепт «государство» является важной составляющей частью концептосферы русского и китайского народа. Статья посвящена исследованию лексического наполнения концептов «государство» на основе сопоставимых корпусов русского и китайского языков, выявлению сходств и различий в содержании дистрибутивных тезаурусов русских и китайских слов.

**Ключевые слова.**концепт, дистрибутивный тезаурус, сопоставительное исследование, языковая картина мира

**1.Введение**

Концепты являются основной единицей языковой картины мира. Языковая картина мира является представлением о мире, сложившимся в обыденном сознании данного языкового коллектива. В концептах аккумулируется культурный уровень каждой языковой личности и всех носителей языка в целом. Концепт репрезентируется в языке лексемами, фразеологизмами, словосочетаниями, схемами предложений, текстами, совокупностями текстов. Сопоставление концептов разных языков способствует нам увидеть общечеловеческие универсалии в окружении окружающего людей мира, и в то же время выявляет специфическое, национальное, а затем групповое и индивидуальное концептов и их структуризации.

**2.** **Концепты и семантические поля**

Концепт является основной единицей языковой картины мира. Концепт как ментальная сущность имеет национально-специфические черты, соотносимые с мировидением, культурой, обычаями, верованиями и историей народа.

Семантические поля, которые принадлежат различным языкам, отражают структурную организацию окружающего мира. Они являются способами репрезентации языковых картин мира.

Семантические поля, с помощью которых можно описывать лексическую систему языка, фактически, являются «овеществленными» отражениями концептов. Впервые термин «семантическое поле» был введен Г. Ипсеном. Поле обычно определяется как совокупность языковых единиц, объединенных каким-то общим семантическим признаком; имеющих некоторый общий компонент значения. Семантический признак, лежащий в основе семантического поля, может также рассматриваться как некоторая понятийная категория (А.В. Бондарко, Л.М. Васильев, И.М. Кобозева). В.Г. Адмони считает, что поле характеризуется наличием инвентаря элементов, связанных системными отношениями. По мнению В.Г. Адмони, в поле можно выделить центральную часть — ядро, элементы которого обладают полным набором признаков, определяющих данную группировку, и периферию, элементы которой обладают не всеми, характерными для поля признаками, но могут иметь и признаки, присущие соседним полям.

**3. Сравнительный анализ лексического наполнения концептов «государство» на основе корпусов русского и китайского языков**

Цель нашего исследования — сопоставление лексического наполнения концептов в корпусах русского и китайского языков, и описание их основных интегральных и дифференциальных признаков. Гипотеза исследования заключается в том, что у концептов государство/国家 существует сходства и различия, что обусловлено универсальными и национально-культурными признаками.

Материалом исследования являются специально созданный нами корпус русского языка по теме «государство» на основе текстов о государстве в русской культуре (3,9 млн. токенов), и корпус китайского языка на теме «国家(государство)» на основе текстов на современном китайском языке о государстве в китайской культуре (3,6 млн. токенов).

В качестве метода применяется в исследовании дистрибутивно-статистический анализ, при котором используется информация о дистрибуции элементов текста и их числовых параметрах.

Инструментом исследования является инструмент «Тезаурус» в системе Sketch Engine. Это инструмент для создания автоматического дистрибутивного тезауруса, который предоставляет нам возможность выявить, какие слова имеют схожую дистрибуцию с заданными словом. В этом случае мы говорим о семантической близости или парадигматическими подобиями слов. Единицы семантического поля обладают общими парадигматическими и синтагматическими свойствами, что показывают их семантическую близость.

С помощью инструмента «Тезаурус» мы построили дистрибутивный тезаурус для ключевых слов «государство» и «国家» на основе корпусов русского и китайского языках по теме «государство/国家» (см. Табл. 1 и 2).

*Таблица 1.* **Дистрибутивный тезаурус для слова «государство» в русском языке**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Lemma** | **Score** | **Freq** | **Lemma** | **Score** | **Freq** |
| общество | 0,31 | 3 839 | организация | 0,179 | 1 926 |
| власть | 0,294 | 5 589 | класс | 0,179 | 1 776 |
| народ | 0,257 | 4 858 | политика | 0,17 | 1 657 |
| страна | 0,242 | 4 141 | демократия | 0,165 | 1 315 |
| человек | 0,241 | 7 233 | церковь | 0,165 | 1 802 |
| система | 0,227 | 2 450 | князь | 0,165 | 4 926 |
| Россия | 0,215 | 3 515 | империя | 0,159 | 1 156 |
| право | 0,212 | 4 908 | орган | 0,153 | 2 038 |
| партия | 0,21 | 2 145 | собственность | 0,15 | 1 654 |
| республика | 0,203 | 2 265 | революция | 0,149 | 1 564 |
| жизнь | 0,189 | 3 405 | развитие | 0,148 | 2 623 |
| город | 0,184 | 3 356 | русь | 0,144 | 1 869 |
| мир | 0,182 | 3 056 | строй | 0,143 | 950 |
| сила | 0,182 | 3 132 | хозяйство | 0,141 | 1 517 |
| союз | 0,181 | 1 580 | нация | 0,141 | 942 |

*Таблица 2.* **Дистрибутивный тезаурус для слова «государство» в китайском языке**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Lemma** | **Score** | **Freq** | **Lemma** | **Score** | **Freq** |
| 社会(общество) | 0,422 | 11 991 | 革命(революция) | 0,222 | 3 729 |
| 政治(политика) | 0,369 | 11 034 | 王朝(династия) | 0,215 | 2 737 |
| 中国(Китай) | 0,362 | 21721 | 人民(народ) | 0,213 | 3 266 |
| 政府(правительство) | 0,361 | 6 329 | 力量(сила) | 0,212 | 1 938 |
| 人(человек) | 0,327 | 10 611 | 地区(регион) | 0,212 | 2 559 |
| 民族(нация) | 0,318 | 8 993 | 结构(структура) | 0,205 | 2 202 |
| 制度(строй) | 0,316 | 4 994 | 文明цивилизация) | 0,203 | 2 387 |
| 国(страна) | 0,313 | 9 764 | 日本(Япония) | 0,192 | 4 701 |
| 经济(хозяйство) | 0,295 | 6 738 | 政党(партия) | 0,192 | 1 800 |
| 权力(власть) | 0,272 | 3 560 | 体系(система) | 0,192 | 1 601 |
| 帝国(империя) | 0,264 | 2 840 | 政策(политика) | 0,188 | 2 501 |
| 关系(отношение) | 0,262 | 4 040 | 政权(власть) | 0,181 | 1 567 |
| 世界(мир) | 0,256 | 4 895 | 利益(интерес) | 0,178 | 2 035 |
| 组织(организация) | 0,23 | 2 751 | 统治(господство) | 0,176 | 2 176 |
| 发展(развитие) | 0,23 | 3 144 | 个人(индивид) | 0,172 | 1 888 |

Из сравнения элементов из двух семантических полей видно, что совпадают 17 из 30, а именно: *общество, власть, народ, страна, человек, система, партия, мир, сила, организация, политика, империя, революция, развитие, строй, хозяйство, нация.* Таким образом, в семантических полях «государство» на китайском и русском языках характеризуются значительным объемом совпадающих элементов.

Можно отметить и различия между китайским и русским полями. В русском тезаурусе по коэффициенту семантической близости (score) занимают первые 5 места эти слова – *общество, власть, народ, страна, человек*, а в китайском тезаурусе – *社会(общество), 政治(политика), 中国(Китай), 政府(правительство), 人(человек)*. Самые частотные слова в дистрибутивном тезаурусе для слова «государство» в русском языке – *человек, власть, право, народ, страна*, а в китайском дистрибутивном тезаурусе они – *社会(общество), 中国(Китай), 政治(политика), 人(человек), 国(страна)*. Кроме того, в ядре русского поля существует такие уникальные элементы, как *Россия, князь, Русь, церковь.* В ядре китайского поля следующие лексемы, которых нет в русском: *中国(Китай), 朝代(династия), 日本(Япония), 世界(мир) и др.*

Рассматриваемые различия в русском и китайском полях показывает национально-культурную специфику в русской и китайкой картинах мире.

**4. Переводческие эквиваленты семантического поля «государство»**

В этой части мы исследуем переводные эквиваленты для элементов из семантического поля «государство», мы рассматриваем русский язык как исходный язык, а китайский – язык перевода.

В исследовании используется параллельный корпус OPUS в системе Sketch Engine. С помощью инструмента «Тезаурус» в Sketch Engine мы построили дистрибутивный тезаурус для слова «государство».



*Рис 1.* **Дистрибутивный тезаурус для слова «государство» по корпусу OPUS**

Инструмент «Параллельный корпус» в Sketch Engine предоставляет нам возможность составить параллельные конкордансы для заданного слова.



*Рис 2.* П**араллельные конкордансы в системе Sketch Engine**

Для каждой лексемы из тезауруса можно найти возможные переводные эквиваленты. С помощью инструмента «Параллельный корпус» мы получили из корпуса OPUS случайную выборку с 200 параллельными конкордансами для первых 10 элементов тезауруса «государство». В таблице показаны переводческие эквиваленты для первых 10 элементов тезауруса «государство» и их процентное соотношение.

*Таблица 3*. **Переводческие эквиваленты для первых 10 элементов тезауруса «государство»**

|  |  |  |
| --- | --- | --- |
| **лексемы** | **эквиваленты в китайском языке** | **процентное отношение** |
| правительство | 政府(правительство) | 82% |
| 国(государство) | 2% |
| 国家(государство) | 2% |
| страна | 国家(государство) | 57% |
| 国(государство) | 19% |
| 境(территория) | 2% |
| орган | 机构(орган) | 32% |
| 机关(орган) | 10% |
| 当局 (администрация) | 11% |
| 部门 (отделение) | 4% |
| 实体 (субъект) | 3% |
| 组织 (организация) | 3% |
| 局(управление) | 2% |
| 政府(правительство) | 1% |
| 机制(механизм) | 1% |
| организация | 组织(организация) | 58% |
| 机构(орган) | 4% |
| 举行(проведение) | 2% |
| 开展(проведение) | 2% |
| сторона | 方(сторона) | 50% |
| 方面(сторона) | 5% |
| 者(человек) | 5% |
| 国(страна) | 2% |
| учреждение | 机构(учреждение) | 65 |
| 组织(организация) | 2% |
| сообщество | 社会(общество) | 63% |
| 共同体(сообщество) | 12% |
| 界(круг) | 5% |
| 社区(община) | 1% |
| группа | 组(группа) | 31% |
| 小组(группа) | 24% |
| 集团(группа) | 7% |
| 团体(коллектив) | 5% |
| 群体(коллектив) | 5% |
| 股(секция) | 2% |
| 队(отряд) | 2% |
| 界(круг) | 1% |
| 组织(организация) | 1% |
| совет | 理事会(совет) | 74% |
| 委员会(комитет) | 8% |
| 会议(собрание) | 2% |
| 董事会(правление) | 1% |
| 局(управление) | 1% |
| лицо | 人(человек) | 78% |
| 者(человек) | 10% |
| 个人(отдельный человек) | 7% |
| 人员(человек) | 5% |

Эти 10 русских слов из дистрибутивного тезауруса «государство» имеют более одного китайского переводного эквивалента. Иногда эти эквиваленты довольно близки по значению русским словам, но следует отметить наличие в китайских переводах оттенков понятий, выражающихся разными словами (иероглифами). Обращают на себя также внимание различия в сочетаемости слов данного семантического поля в русском и китайском языках.

**5. Заключение**

В статье представлен анализ концепт «государство/国家» в русской и китайской языковых картин мира. В русской и китайской языковых картин мира соотносимые концепты «государство/国家» чрезвычайно важны. В семантических полях «государство/国家» имеют значительное число одинаковых элементов: *общество, власть, народ, страна, человек, система, партия, мир, сила, организация, политика, империя, революция, развитие, строй, хозяйство, нация*, это свидетельствует о принципиальном сходстве соответствующих концептов в сопоставляемых языковых картинах мира. Однако нельзя игнорировать и существующие различия лексем, например, лексемы *Россия, князь, Русь, церковь* в русском поле, лексемы *中国(Китай), 朝代(династия), 日本(Япония), 世界(мир)* из китайского поля, они отражают национально-культурную специфику в языковых картинах мира русской и китайской.

**Литература**

1. *Аскольдов С.А.* (1980), Концепт и слово // Русская словесность. От теории словесности к структуре текста. Антология. М.

2. *Захаров В.П*. (2019) Методы автоматизированного формирования семантических полей // Структурная и прикладная лингвистика. Выпуск 13. СПб., Изд-во С.-Петерб. ун-та, c. 56-79.

3. *Кобозева И.М.* (2000), Лингвистическая семантика. Эдито-риал УРСС. М.

4. *Стернин И.А.* (2001), Методика исследования структуры концепта// Методологические проблемы современной лингвистики – Воронеж

5. *Zakharov, V., Pivovarova, S., Gvozdyova, E., & Semenova, N.* (2020). Corpus methods and semantic fields: The concept of empire in English, Russian and Czech. In A. Ronzhin, T. Noskova, & A. Karpov (Eds.), R. Piotrowski's Readings in Language Engineering and Applied Linguistics. PRLEAL-2019: Proceedings of the III International Conference, pp. 233-244

**References**

1. Askol'dov S.A. (1980), Kontsept i slovo // Russkaya slovesnost'. Ot teorii slovesnosti k strukture teksta. Antologiya. – M.

2. Zakharov V.P. (2019) Metody avtomatizirovannogo formirovaniya semanticheskikh poley // Strukturnaya i prikladnaya lingvistika. Vypusk 13. SPb., Izd-vo S.-Peterb. un-ta, c. 56-79.

3. Kobozeva I.M. (2000), Lingvisticheskaya semantika. Edito-rial URSS. M.

4. Sternin I.A. (2001), Metodika issledovaniya struktury kontsepta// Metodologicheskiye problemy sovremennoy lingvistiki – Voronezh, 2001.

5. Zakharov, V., Pivovarova, S., Gvozdyova, E., & Semenova, N. (2020). Corpus methods and semantic fields: The concept of empire in English, Russian and Czech. In A. Ronzhin, T. Noskova, & A. Karpov (Eds.), R. Piotrowski's Readings in Language Engineering and Applied Linguistics. PRLEAL-2019: Proceedings of the III International Conference pp. 233-244

**A corpus-based lexical analysis of the concept "state" in the Russian and Chinese language pictures of the world Title of the paper**

P. Zhang

**Abstract.** The concept "state" is an important component of the conceptual sphere of the Russian and Chinese people. The article is devoted to the study of the lexical content of the concepts "state" on the basis of comparable corpora of the Russian and Chinese languages, identifying the similarities and differences in the content of the distribution thesauri of Russian and Chinese words.

**Keywords.** сoncept, distributional thesaurus, comparative research, the language picture of the world
**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Чжан Пэйлинь**

Санкт-Петербургский государственный университет (Россия).

**Zhang Peilin**

Saint Petersburg State University (Russia).

***E-mail: zhangpl@yandex.ru***