

О применении компьютерных технологий для выявления ключевых концептов в рассказе В.Г. Короленко «Сон Макара»

On the use of computer technology to identify key concepts in V.G. Korolenko's story «Makar's Dream»

Майер Р.В.

Д-р пед. наук, звание доцент, Заслуженный деятель науки Удмуртской Республики, профессор, Глазовский государственный педагогический институт им. В.Г. Короленко
e-mail: robert_maier@mail.ru

Mayer R.V.

Doctor of Pedagogical Sciences, Associate Professor, Honored Scientist of Udmurt Republic, professor, Glazov Korolenko State Pedagogical Institute
e-mail: robert_maier@mail.ru

Аннотация

Проанализирована лексика рассказа В.Г. Короленко «Сон Макара», получен список идеоглосс с указанием числа использований, выявлены основные тематические группы: 1) человек, люди; 2) крестьянский быт, работа; 3) природа, окружающий мир; 4) здоровье, органы, части тела; 5) хронотоп; 6) религия, вера; 7) мысли, знания, чувства; 8) животные; 9) взаимодействие персонажей друг с другом; 10) пьянство; 11) населённый пункт.

Ключевые слова: В.Г. Короленко, идеоглосса, концепт, лексическое пространство, картина мира, языковая личность.

Abstract

The vocabulary of V.G. Korolenko's short story «Makar's Dream» is analyzed, a list of ideoglosses with an indication of the number of uses is obtained, the main thematic groups are identified: 1) man, people; 2) peasant life, work; 3) nature, the surrounding world; 4) health, organs, body parts; 5) chronotope; 6) religion, faith; 7) thoughts, knowledge, feelings; 8) animals; 9) interaction of characters with each other; 10) drunkenness; 11) locality.

Keywords: V.G. Korolenko, ideoglossa, concept, lexical space, worldview, linguistic personality.

Введение

Одной из важных проблем филологии является изучение индивидуальной языковой картины мира (КМ) писателя, его языковой личности. Языковая КМ включает в себя индивидуальные и творчески переработанные представления о мире, которые художник стремится воплотить в своих произведениях, а также словесные образы, лексику и синтаксис, отражающие индивидуальный стиль и особенности мышления автора.

По мнению Ю.Н. Караулова [6, с. 43] полное описание языковой личности предполагает: 1) изучение ее вербально-семантического уровня (полный лексикон, ассоциативный словарь и т.д.); 2) характеристику языковой модели мира писателя, его тезауруса (концепты и связи между ними); 3) исследование мотивационного уровня, означающее «выявление жизненных или ситуативных доминант, установок, мотивов, находящих отражение в процессах порождения текстов и их содержания, а также в особенностях восприятия чужих текстов». Лексическая составляющая языковой картины мира может быть охарактеризована совокупностью используемых автором ключевых концептов (идеоглосс), упорядоченных по

частоте употребления. Ключевые слова являются основными доминантами семантического пространства художественного произведения, выступают связующим звеном между языком художественного текста и текстообразующим концептом. Они притягивают внимание читателя, обладают наибольшей эстетической значимостью, представляя собой доминанты семантического пространства текста, несут высокий когнитивный потенциал, выступая в качестве маркеров идейно-тематического поля текста [13].

Цель статьи: обсудить методику изучения концептуального и лексического пространства художественных текстов путем их анализа на компьютере и применить ее для анализа рассказа В.Г. Короленко «Сон Макара». Полученные результаты будут способствовать более глубокому постижению языковой личности писателя. **Методологической основой** исследования являются работы Т.В. Бахваловой и А.Р. Поповой [1], Е.В. Туктангуловой [12], Е.А. Узенцовой [13], А.И. Хартиковой [14] (лексическое и концептуальное пространство текста), Ю.Н. Караулова [6], О.Д. Тихоненко [11] (понятия «идиолект», «идиостиль», «языковая личность»), С.Г. Воркачева [3], В.И. Карасика [5], О.И. Лыткиной [7] (лингвокультурология, языковая личность, концепт), И.Р. Гальперина [4], Р.В. Майера и Н.Н. Закировой [8] (методы исследования текста), Т.Б. Радбиля [10] (языковой менталитет). А.А. Бувеч [2], Е.А. Макарова [9] (религиозные концепты).

Результаты исследования

Лексический состав художественного произведения зависит от: 1) языковой картины мира писателя; 2) описываемых событий. Изучение языковой личности писателя предусматривает лексический анализ его произведений, перечисление всех используемых слов и их группировку по тематическим группам. При этом можно перейти от изучения довольно большой совокупности слов, относящихся к различным тематическим группам, к изучению ключевых концептов – идеоглоссов, которые представляют собой тезаурусообразующие понятия и воспроизводят содержание текста в свернутом виде. Под идеоглоссами понимают «единицы когнитивного уровня в структуре языковой личности, лексические маркеры индивидуальной манеры создания художественного текста, важный концепт в языковой картине» [6]. В тексте они играют роль ключевых концептов, образуют точки концентрации смысла, своеобразные центры кристаллизации, вокруг которых формируются ассоциативные и концептуальные поля. Идеоглоссы – единицы индивидуального авторского лексикона, важные концепты, характеризующие идеолект автора; часто они выступают семантическими множителями, от которых образуются другие слова.

В качестве предмета изучения нами был выбран рассказ «Сон Макара» (В.Г. Короленко, 1883 г.), повествующий о крестьянине Макаре, который мечтал спасти свою душу, много трудился, но бедно жил и часто пил. Накануне Рождества Макар поехал в тайгу, чтобы проверить поставленные ранее ловушки. В лесу он встретил соседа Алёшку и подрался с ним. Затем Макар лег на снег и заснул (или умер). Он оказался в загробном мире, поп отвел Макара на суд к Тойону, который должен был определить, куда направить душу Макара. Помощники Тойона с помощью весов взвешивали грехи и добрые дела Макара. Когда он поведал о своём горьком существовании, Тойон и его окружающие, расплакались, чаша с грехами стала легче.

Для выявления ключевых концептов произведения его текст был проанализирован с помощью Интернет-сервиса advego.com. На выходе получился список слов в начальной форме с указанием количества и частоты использования (файл `spisok.txt`). При этом причастия и деепричастия преобразовывались в глаголы. Изучение этого списка позволило выявить часто используемые слова и разделить их на тематические группы, соответствующие смысловым компонентам лексического пространства произведения. Для этого применялась компьютерная программа, написанная в ABCPascal, которая из файла `slova.txt` берет подслово (например, -вес-, -хлеб-, -юрт- и т.д.), выбирает из файла `spisok.txt` все слова, содержащие данное подслово, записывает их в файл `vihod.txt`.

В результате анализа и классификации удалось выявить следующие тематические группы слов и соответствующие им ключевые концепты: 1) слова обозначающие людей (концепт – человек, люди), $x_1 = 450$; 2) крестьянский быт, работа, $x_2 = 308$; 3) природа, окружающий мир, $x_3 = 238$; 4) здоровье, органы, части тела, $x_4 = 200$; 5) хронотоп, $x_5 = 177$; 6) религия, вера, $x_6 = 159$; 7) мысли, знания, чувства, $x_7 = 153$; 8) животные, $x_8 = 100$; 9) взаимодействие персонажей, $x_9 = 71$; 10) пьянство, $x_{10} = 52$; 11) населённый пункт, $x_{11} = 23$.

Полученные результаты представлены ниже.

1. ЧЕЛОВЕК, ЛЮДИ ($x_1 = 450$): а) по именам (208): Макар – 167, Иван – 26, Алешка – 14, Макарушко – 1; б) по национальности (60): татарин (татарский) – 27, чалганец (чалганский) – 16, якут (якутский, якутка) – 16, поселенец-татарин – 1; в) по роду деятельности (113): поп (попик) – 64, работник – 13, всадник – 6, исправник – 6, хозяин – 5, убийца – 3, воин – 2, вор – 2, комночит (работник) – 2, крестьянин – 2, купец – 2, по одному разу: заседатель, матушка-попадья, прихожанин, путник, солдат, старшина; г) другие слова (95): человек – 21, старуха (старухин) – 20, старик – 16, сын – 4, трапезник – 4, брат – 3, отец – 3, приятель – 3, староста – 3, бедняга – 2, жена – 2, ленивец – 2, народ – 2, обманщик – 2, по одному разу: житель, нарушитель, нечестивец, посетитель, сосед, старший, супруга, суруксут (писарь).

2) РАБОТА, БЫТ ($x_2 = 308$): чашка – 14, вес – 12, камелек – 12, плаха – 12, огонь (огненный) – 12, воз (возить) – 10, изба – 10, ловушка (для лис) – 10, шапка – 9, деньги (деньга) – 8, дорога – 8, дрова – 8, работа (работать) – 8, рубить (срубить) – 8, дверь – 7, жердь – 6, одежда – 6, постель – 6, руга – 6, юрта (юртенка) – 6, книга – 5, махорка – 5, стол – 5, табак – 5, чай – 5, дровни – 4, дым – 4, мыть (домыть, обмыть) – 4, рукав – 4, торбас (торбасишка – обувь) – 4, ворота – 3, дорожка – 3, хлеб – 3, седло – 3, след – 3, угол – 3, шерсть – 3, бергес (шапка) – 2, вожжа – 2, воротник – 2, жареный – 2, жерло – 2, жернов – 2, кусок – 2, лепешка – 2, молот – 2, мясо – 2, облако – 2, обманщик – 2, овес – 2, пол – 2, полоз – 2, птица – 2, пыль – 2, рубаха – 2, рукавица – 2, сапог – 2, солома – 2, стул – 2, топор – 2, тпру-у – 2, шкура – 2, шуба – 2, по одному разу: алас (плоская котловина), весить, весы, взвесить, дымный, изгородь, плис (мех), полено, сквозняк, столб, уголь, шиворот.

3) ПРИРОДА, ОКРУЖАЮЩИЙ МИР ($x_3 = 238$): снег (снежный) – 31, тайга – 26, тьма – 15, дерево (деревянный) – 12, свет (светлый, просвет, рассвет, светло, светать) – 12, земля – 10, равнина – 11, гора – 9, звезда (звездочка, звездный) – 9, луна (лунный) – 8, холод (холодное) – 7, воздух – 6, небо – 6, сияние – 6, туман – 6, листок (листочек) – 5, мороз – 5, холм (холмик) – 5, горизонт – 4, лес (лесной) – 4, лиственница – 4, луч – 4, ночь – 4, ветвь – 3, солнце – 3, таять – 3, вершина – 2, иней – 2, сугроб – 2, снежинка – 2, эхо – 2, по одному разу: морозный, лучистый, возвышенность, полынья, поляна, река, озеро, метель, земляца, месяц.

4) ЗДОРОВЬЕ, ЧАСТИ ТЕЛА, ОРГАНЫ ($x_4 = 200$): а) части тела, органы (146): нога – 24, сердце – 21, лицо – 19, голова – 18, глаз – 15, рука – 13, ухо – 6, спина – 5, слеза – 4, тело – 4, борода – 3, горло – 3, волос (волосяной) – 3, жила (жилистый) – 2, колено – 2, по одному разу: нос, кожаный, живот, грудь; б) здоровье, жизнь, смерть (54): жизнь (жить, живой, оживляться) – 22, сон – 9, смерть – 8, глухой – 4, дыхание – 3, здорово – 3, больной – 2, умереть – 2, выживать – 1.

5) ХРОНОТОП ($x_5 = 177$): а) время, возраст (91): старый (состариться, старость, престарый) – 39, долгий (задолго) – 9, год – 8, молодой – 7, завтра – 6, после – 6, тотчас – 6, секунда – 3, сейчас – 2, сегодня – 2, вечер (вечерний) – 2, час – 1; б) пространство, направление (86): далёкий (далеко, даль, дальний, недалёкий, далее, удаляться) – 29, длинный – 9, сзади – 7, вперед – 6, назад – 6, правый – 5, край – 4, восток – 3, направление – 3, близко (близкий) – 3, впереди – 2, левый – 2, север – 2, по одному разу: направо, верста, влево, запад, налево.

б) РЕЛИГИЯ, ВЕРА ($x_6 = 159$): поп (попик) – 59, тойон (тойон-исправник) – 47, грех – 7, праздник – 7, церковь (церковный) – 7, бог (божий, божиться) – 5, колокол (колокольная) – 5, вера (верить) – 3, черт – 3, крестины – 2, ряса (ряска) – 2, по одному разу: бесноваться, гос-

подь, дух, дьявол, заколдовать, иконописец, матушка-попадья, молебен, молиться, нечестивец, рождество, святочный.

7) МЫСЛИ, ЗНАНИЯ, ЧУВСТВА ($x_7 = 153$): знать (сознавать, узнавать) – 33, подумать (дума, вдумчивый, вздуматься, задумчивый) – 21, чувствовать – 11, любить – 10, сердитый (сердиться) – 10, желать (желание, пожелать) – 8, мысль – 8, помнить (вспомнить, припомнить) – 8, жалеть (сожаление, жалко) – 7, плакать (заплакать) – 7, забыть – 6, проклясть – 5, торжественный – 5, радость (порадоваться, радостный) – 3, отчаяние – 3, бояться – 2, гордиться – 2, знакомый – 2, изумиться – 1, смеяться – 1.

8) ЖИВОТНЫЕ ($x_8 = 100$): конь (конек, конный) – 25, лисица (лисий) – 12, лысанка – 8, крыло – 6, хвост – 6, бык – 5, заяц – 4, зверь – 4, лошадь – 4, мерин – 4, коготь – 2, копыто – 3, лапка – 2, морда – 2, птица – 2, скотина – 2, теленок (телячий) – 2, хлестнуть (подхлестывать) – 2, по одному разу: животное, корова, лапа, пегашка (пегая лошадь), собака.

9) ВЗАИМОДЕЙСТВИЕ ПЕРСОНАЖЕЙ ($x_9 = 71$): говорить – 22, ответить – 16, песня (петь, запеть) – 9, кричать (вскричать) – 6, обратиться (с вопросом) – 6, рассказ (рассказать) – 3, колотить – 2, ругать – 2, толкнуть – 2, ударить – 2, распорядиться – 1.

10) ПЬЯНСТВО ($x_{10} = 52$): пить (выпить, напиться, пропить) – 18, водка – 15, пьяный (пьяница, опьянеть) – 10, бутылка – 9.

11) НАСЕЛЁННЫЙ ПУНКТ ($x_{11} = 23$): Чалган – 8, слобода – 5, село – 4, Ямалаха – 2, по одному разу: город, поселок, слободка, сельный.

Кроме того, писатель использует абстрактные существительные (всего 184 употреблений): а) числительные (86): один – 37, тысяча – 9, сотня – 8, пять (пятый, впятеро) – 6, два – 4, три – 4, четыре – 4, десять – 2, тринадцать – 2, четыреста – 2; по одному разу: двадцать, два-три, девятьсот, тридцать, триста, шестнадцать, шестнадцать, шесть; б) другие абстракции (98): работа – 5, нужда – 4, положение – 4, внимание – 3, надежда – 3, направление – 3, отчаяние – 3, множество – 2, молчание – 2, намерение – 2, негодование – 2, обстоятельство – 2, решение – 2, система – 2, удивление – 2, желание – 2, забота – 2, зло – 2, по одному разу: беспокойство, воспоминание, горе, готовность, движение, действие, дерзость, дряхлость, задумчивость, звание, излияние, исполнение, мгновение, надобность, невозможность, неудобольствие, обаяние, обращение, обязанность, ожидание, отпущение, ощущение, помышление, попечение, почесть, предложение, пространство, прошение, радость, рождество, собрание, сожаление, сознание, состояние, старость, стеснение, страх, существование, терпение, течение, убеждение, удовольствие, упорство, устройство, хозяйство, ярость.

В рассказе также встречается большое количество слов-универсалий; среди них глаголы: был – 90, казаться – 13, лежать – 12, гонять – 11, деть – 11, подняться – 10, становиться – 10, увидеть – 10, делать – 9, сделать – 8, слышать – 8, давать – 7, ехать – 7, остановиться – 7, попадать – 7, пропадать – 7, ходить – 7, остаться – 6, посмотреть – 6, продолжать – 6, рубить – 6, слушать – 6 и др. В.Г. Короленко также использует редко употребляемые слова, относящиеся к региональной лексике: бергес (шапка), городьба, комночит (работник), лысанка (лошадь), падь (ущелье), руга, суруксут (писарь), Тойон, торбас (обувь), тытыма! (не тронь!), ушкан (заяц), Чалган, Ямалаха и т.д. Это означает, что писатель проявлял интерес к якутской культуре и языку.

Какие выводы можно сделать, анализируя лексическое пространство произведения? Часто используемое слово был (была, было) означает, что повествование ведется в прошедшем времени. Описанные в рассказе «Сон Макара» события произошли в Чалгане (Якутия), где живут татары, чалганцы и якуты. Кроме главного героя Макара упоминаются другие персонажи: поп Иван, Алешка, старик, старуха, работники, всадники, бог Тойон; многие из них не молоды. Кругом тайга, зима, снег, холод, тьма или свет, деревья, лес, равнины, холмы, горы, на небе звезды, солнце и луна. Описывая своих героев, автор упоминает части тела: нога, сердце, голова, глаз, рука, ухо и др. Персонажи живут, умирают, засыпают, оживляются; некоторые из них пьют водку и ходят пьяными. При этом они что-то знают, думают, чувствуют, любят, сердятся, желают, мыслят, вспоминают, жалеют, плачут, что-то забывают, проклинаят, изредка радуются и смеются. В рассказе упоминаются животные: конь, лошадь,

лисица и т.д. Присутствуют маркеры религиозного мировоззрения: поп (попик), тойон (тойон-исправник), церковь, бог, колокол и т.д., причём чёрт, дьявол и сатана упоминаются гораздо реже. Это означает, что автор и его герои уважительно относятся к религии.

Анализ лексики конечно же не позволяет восстановить сюжет рассказа, но помогает понять его общую направленность, в частности утверждать, что оно скорее относится к «печальным» текстам, чем к «оптимистичным» или «радостным». Об этом свидетельствуют маркеры «бедняга», «убийца», «вор», «старый» (в 5,6 раз чаще, чем «молодой»), «сердиться», «жалеть», «плакать», «проклясть», «отчаяние», «бояться», а также идеоглоссы, относящиеся к тематической группе «пьянство». Текстобразующий концепт рассказа, – тяжелая жизнь бедняка в Сибири; именно с ним связано большинство идеоглоссов.

Если количества слов, относящиеся к различным тематическим группам, записать в виде матрицы (x_1, x_2, \dots, x_{11}), то получится вектор (450, 308, 238, 200, 177, 159, 153, 100, 71, 52, 23). Аналогичный анализ очерка «На затмении» дает: $x_1 = 318$, $x_2 = 73$, $x_3 = 247$, $x_5 = 116$, $x_6 = 59$; $x_7 = 83$. Видно, что в обоих случаях в центре внимания – простые люди (x_1 велико), однако в первом случае автор уделяет больше внимания работе и быту простых людей ($x_2 = 308$), их мыслям о боге ($x_6 = 159$), а во втором – описанию природных явлений ($x_3 = 247$).

Полученные результаты позволяют очертить лексическое пространство рассказа «Сон Макара», выявить основные лексико-тематические группы слов, изучить ключевые слова и их лексическое окружение. Перечисленные ключевые концепты и соответствующие им тематические группы входят в состав языковой личности писателя и поэтому проявляются в других произведениях В.Г. Короленко.

Заключение

В статье обсуждается использование компьютера для выявления и изучения ключевых концептов художественных произведений, что необходимо для анализа языковой личности автора и особенностей его языковой картины мира. В качестве примера определены идеоглоссы рассказа В.Г.Короленко «Сон Макара», установлено, что они составляют следующие тематические группы: 1) человек, люди; 2) крестьянский быт, работа; 3) природа, окружающий мир; 4) здоровье, органы, части тела; 5) хронотоп; 6) религия, вера; 7) мысли, знания, чувства; 8) животные; 9) взаимодействие персонажей; 10) пьянство; 11) населённый пункт. Сосчитано количество слов, относящихся к каждой тематической группе, которые образуют лексическое пространство произведения и выступают компонентами языковой личности автора.

Литература

1. Бахвалова Т. В., Попова А. Р. Лексическое пространство классических произведений и его роль в лингвокультурологической подготовке будущих учителей (на материале текстов И.С. Тургенева) // Ученые записки Орловского государственного университета. – № 1 (90), 2021. – С. 106–112.
2. Бувевич А. А. Религиозный концепт «спасение» в русской культуре (на материале религиозного, философского и художественного дискурсов) // Вестник РУДН, серия Теория языка. Семиотика. Семантика. – 2014, № 2. – С. 45–61.
3. Воркачев С. Г. Лингвокультурология, языковая личность, концепт: Становление антропоцентрической парадигмы в языкознании // Филологические науки. – 2001. №. 1. – С. 64–72.
4. Гальперин И. Р. Текст как объект лингвистического исследования. – М.: Наука, 1981. – 138 с.
5. Карасик В. И. Языковой круг: личность, концепты, дискурс: – Волгоград: Перемена, 2002. – 477 с.
6. Караулов Ю. Н. Русский язык и языковая личность: Моногр. – М.: Издательство ЛКИ, 2010. – 264 с.

7. Лыткина О. И. Концептосоставляющие «АМЕРИКИ» в рассказе В.Г. Короленко «Без языка» // Вестник Нижегородского университета им. Н.И. Лобачевского. – 2011. № 6 (2). – С. 382–385.
8. Майер Р. В., Закирова Н. Н. Применение математических методов для анализа прозаических произведений (на примере очерка В.Г. Короленко «На Волге») // Вестник филологических наук. – 2023. Том 3. № 1. – С. 71 – 78.
9. Макарова Е. А. Формирование религиозно-философской позиции В.Г. Короленко (на материале дневниково-эпистолярной прозы) // Сиб. филол. журн. – 2007. № 4. – С. 38 – 49.
10. Радбиль Т. Б. Основы изучения языкового менталитета: Учеб. пособие. – М.: Флинта: Наука, 2009. – 328 с.
11. Тихоненко О. Д. Категории «идиолект», «идиостиль», «языковая личность» и методики их лингвистического описания в контексте изучения языка художественной литературы // Вестник КРСУ. – 2013. Том 13. № 9. – С. 159 – 164.
12. Туктангулова Е. В. Ключевое слово – словообраз – художественный концепт как универсальные текстообразующие константы художественного текста // Вестник Челябинского государственного ун-та. Научный журнал. Филология. Искусствоведение. – 2007. № 8. – С. 105–109.
13. Узенцова Е. А. Доминанты семантического пространства художественного текста: ключевое слово, концепт, символ // Интерпретация текста: лингвистический, литературоведческий и методический аспекты. – 2014. № 7. – С. 33–35.
14. Хартикова А. И. Концептуальное пространство текста в художественной картине мира автора // Международный журнал «Символ науки». – 2017. № 02-2. – С.70–172.