



ИмPLICITНЫЕ ЗНАНИЯ В ВОСПРИЯТИИ ТЕКСТА (на работах В. Набокова и И. Бродского): нейросетевой подход

Харламов А. А. Пильгун М. А.

Аннотация

В работе представлены результаты использования нейросетевого подхода в психолингвистических исследованиях восприятия текста. Были проведено тестирование возможностей технологии TextAnalyst для анализа восприятия и интерпретации текста, а также для выявления в тексте имплицитной информации. Материалом послужили произведения «Другие берега» В. Набокова и «Набережная неисцелимых» И. Бродского. Автобиографические произведения были выбраны для анализа как содержащие личную, важную для внутренней жизни авторов информацию. Нейросетевой подход, реализованный в технологии TextAnalyst позволил использовать потенциал ассоциативных связей в тексте на нескольких уровнях: для выявления эксплицитной информации на уровне семантической сети и для выявления имплицитной информации с помощью ассоциативного поиска. Данный тип автоматического анализа текста дал возможность определить объективную информацию, выраженную речевыми средствами: тематическую структуру текста — внешнюю канву повествования и семантическую структуру — смысловую основу текста, содержащую объективную информацию, сформулированную автором. Между тем, ассоциативный поиск выявил лексические ассоциации, которые свидетельствуют об оценочной парадигме, интенции автора, позволяют делать выводы об подтекстовой информации, коннотативных значениях, заложенных в тексте.

Ключевые слова: нейросетевой текстовый анализ, восприятие, Набоков, Бродский

Введение

Психолингвистические ресурсы активно адаптируются к новым технологиям: для структурирования и поиска информации, построения и расширения семантические сети, а также для анализа Больших данных и пр. (см. Lérique, Roth, 2018; Ekmekci, Sahin, Gulacar, Almus 2018; Vivas, Manoiloff, García, Lizarralde, Vivas 2019 и др.). Mark Steyvers and Joshua

В. Тепенбаум показали, что развернутая структура семантических сетей, возникающая из механизмов семантического приращения, имеет серьезные последствия для психолингвистических теорий [2010].

В различных практиках широко применяются парадигмы лексических ассоциаций (Word Association — WA). Универсальность тестов и экспериментов на базе лексических ассоциаций подтверждается их долгой успешной историей: от первых ассоциативных экспериментов (Galton, 1880) до использования для поисковых запросов, создания баз данных и классификаций изображений, нейросетевых технологий (Pejtersen 1991; Ornager 1997. Medford 1997; Kharlamov, Le 2017 и др.). Нормы свободной ассоциации являются одними из наиболее востребованных в когнитивных исследованиях.

Полученные данные об ассоциациях использовались для построения ассоциативной сети и изучения центральности узлов (De Deyne, Storms (2008). Исследование в свете многокомпонентных теорий семантической памяти показали, что ассоциативная сеть может кодировать однородные отношения между понятиями, а контекстные представления могут использоваться для обработки информации о более абстрактных понятиях (Gruenenfelder, Recchia, Rubin, Jones 2016). Использование лексических ассоциаций позволяет выявить лингвистические и концептуальные особенности грамматических форм (Guida, Lenci 2009).

Ассоциативные структуры фиксируют общий лексический опыт носителей языка и позволяют решать различные исследовательские проблемы (Nelson et al., 2000; Dennis, 2003; White & Abrams 2004 и др.) Например, было установлено, что анализ модульной организации ассоциативных сетей может служить инструментом для изучения общественного мнения в отношении различных социальных проблем. В частности, индуктивный характер модулей CoOp позволяет классифицировать динамику семантических характеристик, формулировать представление на базе данных локальной социальной ситуации в конкретный временной период (File, Keczer, Vancsó et al. 2018). Ассоциативный прайминг используется в рамках так называемого «эффекта семантического прайминга» (Ferrand & New 2003).

Поскольку способность представлять понятия и отношения между ними имеет решающее значение для человеческого познания, интерес к анализу ассоциативных процессов, протекающих в мозге человека, привлекает особое внимание ученых. Так, различные аспекты функционирования ассоциативной памяти описаны в работах Hopfield, J.J. 1982, Cheng WC., Liou CY. 2008; Madan, C.R., Lau, C.S., Caplan, J.B., Fujiwara E. 2009; Miyata, R., Ota, K. & Aonishi, T. 2013; Sinha, A., Davey, N., Adams, R. Steuber V. 2015; Харламова 2017 и др.

Анализ нейронных корреляций семантических ассоциаций представлен в работе Rebecca L. Jackson, Paul Hoffman, Gorana Pobric, Matthew A. And Lambon Ralph, исследовавших специфику кодирования мозгом отношений между предметами, которые имеют общие концептуальные свойства, с одновременным представлением ассоциативных связей между разнородными предметами, которые встречаются в определенных контекстах. Результаты исследований нейронных основ данных семантических компонентов, категориальные и ассоциативные отношения сформированы в единой семантической системе (Jackson, Hoffman, Pobric & Lambon 2015) и др.

Развитие автоматического анализа текста привело к расширению использования возможностей ассоциаций. В частности, нейросетевые технологии выводят тематическую структуру текста в виде тематического дерева с выделением связной иерархической структуры: коренная тема, подтемы, подподтемы и т.д. В противоположность тематическому моделированию, в процессе которого возможно только выделение множества отдельных тем. Кроме того, при нейросетевой интерпретации смыслового наполнения текста семантическая сеть формирует многочленную структуру взаимосвязанных понятий. Т.е. именно нейросетевой подход позволил в определенной степени преодолеть ограничения в представлении когезии и когерентности текста, которые (ограничения) присутствуют в векторных технологиях, тематическом анализе, основанном на статистическом подходе (например, LSA, pLSA, LDA), с монограммной моделью текста и индексами — тематическими перечнями сравниваемых структур. На основе нейросетевого подхода были созданы сервисы для автоматического формирования из текстового корпуса однородной семантической сети, которая определяет смысловой портрет текста. Однородная семантическая сеть, способная представить семантическую основу текста, появилась с развитием представления семантики текста в сторону использования объединения отдельных тематических элементов текста в единое представление, а, потому, использования вместо монограммной модели текста его n -граммной модели. (Hofmann 1999; Kharlamov, Le 2017; Kharlamov 2017 и др.).

Однородная семантическая (ассоциативная) сеть представляет собой граф, с вершинами обозначающими понятия, а дуги указывают на близость этих понятий в рамках анализируемой структуры (например, предложения текста). Подобная сеть, полученная с помощью технологии для автоматического смыслового анализа текстов TextAnalyst, в отличие от модели языка, где используется не более чем 3-граммная модель, учитывает зависимость между словами корпуса текстов, описывающего предметную область на n шагов (в случае использования программы TextAnalyst — на 10 шагов). В текстовой базе данных выделяются взаимосвязанные ключевые понятия, и строится семантическая сеть конкретной предметной области (описанной соответствующим корпусом текстов).

Компьютерная модель, интерпретирующая формирование семантической структуры и ее конечную форму, модель семантической обработки зависят от структурных особенностей, которые всегда влияют на функции, ими выполняемые. (Strogatz 2001, p. 268; Steyvers & Tenenbaum 2005). Collins, A. M., and M. R. Quillian еще в 1969 г. описали взаимодействие одного вида простой структуры семантической сети и дополняющих процессов и положили начало целой серии исследований общих процессов, например, распространения активности, действующей на произвольные структуры. В дальнейших исследованиях рассматривались различные аспекты семантической связности сети и пр. (Steyvers & Tenenbaum 2005).

Метод тематического моделирования, направленный на выявление латентных тем в корпусе, можно найти в работах: Mimno and Blei 2012; Blei, 2012, Daud, 2012 and al.

Уместно вспомнить, что в традиционной лингвистической литературе используются понятия «семантическое поле», «лексико-семантическое поле», которые имеют различные трактовки. Первоначальное теоретическое осмысление поля связано с именами Й. Трира и Г. Инсена. Так, Й. Трир, применив полевой подход к анализу лексической системы языка, на материале немецкого языка раскрыл внутреннюю форму языка. Именно Й. Трир ввел понятие «семантическое поле», сформированное на базе смыслового и понятийного родства слов и возникающее в сознании носителя языка при произнесении того или иного слова. Такие «понятийные ассоциации» образуют совокупность, представляющую собой расчлененное целое (Trier 1931).

К настоящему времени достаточно подробно описаны: структура лексического значения и семантическая структура полисеманта, типы полисемии, и особенности семантической деривации (см. работы Ю. Д. Апресяна 1974, Д. Н. Шмелева 1977, Лайонз 1978 и др.), системный характер лексики на микро- и макроуровнях (термины В. Г. Гака), семантические группировки (Ю. Н. Караулов 1981 и др.).

Семантическое поле в лингвистической традиции характеризуется связью отдельных компонентов, его составляющих, системным характером этих связей, их взаимозависимостью и взаимоопределяемостью, относительной автономностью семантико-коммуникационного поля, непрерывностью смыслового пространства, обозримостью и психологической реальностью для среднего пользователя. Поскольку в семантическое поле входят слова разных частей речи, его единицам свойственны не только синтагматические, парадигматические, но и ассоциативно-деривационные отношения.

Цель работы: определение возможности использования технологии TextAnalyst для анализа восприятия и интерпретации текста, выделения имплицитной текстовой информации.

Гипотеза: нейросетевой подход, использованный в технологии TextAnalyst позволяет использовать потенциал ассоциативных связей на нескольких уровнях: для выявления эксплицитной информации на уровне семантической сети и для выявления имплицитной информации с помощью ассоциативного поиска.

Материал: тексты «Другие берега» В. Набокова, «Набережная неисцелимых» И. Бродского. Были выбраны автобиографические произведения, содержащие личную, важную для внутренней жизни авторов информацию. В данном исследовании анализируются данные, которые отражаются в биографических материалах, фиксируются в различных исторических свидетельствах и могут быть верифицированы с той или иной степенью точности. Авторы объединяет творческое двуязычие, вынужденное изменение языкового выражения, вызванное объективными обстоятельствами — эмиграцией.

Метод: анализ данных проводился с помощью программы TextAnalyst 2.0, были сформированы семантическая сеть, извлеченная из сети тематическая структура анализируемого контента, проводился ассоциативный поиск по релевантным запросам.

Полученный из семантической сети перечень наиболее значимых понятий и концептов (имеющих наивысший ранг), несущих основную смысловую нагрузку, позволил определить ядро информации, наиболее важные для акторов семантические акценты.

Выделение тематической структуры выбранного контента позволило описать содержание в виде иерархии связанных тем и подтем, которые отражают основные концепты и соответствуют узлам сети понятий. Используемый инструментарий позволил создавать иерархическую тематическую структуру, выявить базовую тему и раскрывать связи с подтемами, которые формируют разноуровневые сематические сети.

Кроме того, контент был проанализирован с точки зрения связности тематического дерева. Были выделены тематические кусты, появившиеся в результате изменения порога по весу связей в сети понятий (разрыв более или менее сильных связей в зависимости от величины порога). Подобный анализ собранного датасета позволил проанализировать структуру текста в различных срезах и на разных уровнях семантической глубины.

Результаты и дискуссия

Интерпретация текста, в частности, интенция автора художественного текста — проблемы, которые имеют длительную историю изучения в различных научных парадигмах, но, тем не менее, еще далеки от решения, представляется трудновыполнимой.

Выявление имплицитной информация, заложенной в тексте, также неоднозначно решенная задача.

Имена В. Набокова и И. Бродского как двуязычных авторов занимают особое положение в истории мировой культуры. Творчество В. Набокова, владевшего несколькими языками и создавшего неповторимый, стилистически безукоризненный художественный мир на двух языках, не имеет аналогов в истории мировой литературы. И. Бродский переводил с английского, испанского, польского, сербохорватского, но его собственное творчество представлено только на русском и английском языках. Между тем следует заметить, что И. Бродский был более совершенен в русской поэтической форме, а по-английски предпочитал писать прозаические тексты.

Существуют многочисленные теории, объясняющие механизмы обеспечения билингвизма (см., напр., Mouthon, Khateb, Lazeyras, Pegna, Lee-Jahnke, Lehr & Annoni, 2018 и др.). Во второй половине XX века функционирование языковых структур в мозге билингвов изучалось очень активно, однако всё ещё нет единой точки зрения, объясняющей, идентичные или нет нейронные механизмы регулируют использование двух языков двуязычного человека (Hamers, & Blanc 2000). У билингва совмещаются два языковые кода, и способность пользоваться ими с наименьшей интерференцией в максимальной степени осложняется в процессе творчества двуязычного писателя.

В истории русской литературы сформировалось резкое противопоставление художественного перевода и художественного творчества на нескольких языках.

Показательна цитата из письма И. С. Тургенева Л. Пичу (21 ноября 1880 г.), написанного, между прочим, по-немецки: «Для меня человек, который считает себя писателем и пишет не только на одном — притом своем родном языке, — мошенник и жалкая, бездарная свинья» (цит. по Федоров 1971, 171). Напомним, что И. С. Тургенев, владевший французским, немецким, и английским, сочинил либретто нескольких комических опер, переводил русскую классику и оставил огромное эпистолярное наследие на нескольких языках.

Достаточно долго существовало устойчивое мнение о том, что творческое владение несколькими языками практически невозможно, мнение, которое было распространено не только в рамках русской культуры. Например, в немецкой литературной традиции по-французски пытались писать Гейне и Р.-М. Рильке; в английской — по-французски писали О.Уайльд и Суинберн, у которого есть стихи и на древнегреческом и латинском; во французской — Ж. Верн использовал английский язык. Известны сдержанные, или нелестные отзывы об этих экспериментах.

В русской классической литературе известно небольшое число примеров создания писателями авторских произведений (или частей их) на иностранных языках. (В. К. Тредиаковский, М. Ю. Лермонтов, Ф. И. Тютчев, Л. Н. Толстой, А. К. Толстой). А.С. Пушкин по-французски написал только несколько заметок. Показательно, что письма к невесте он пишет исключительно по-французски, а письма к жене — по-русски, с использованием просторечия. Как известно, иностранные языки (в первую очередь французский) использовались в XIX веке преимущественно для светского общения, в обиходной речи, в письмах, в редких случаях — в публицистике (А. И. Герцен, Ф. И. Тютчев).

В XX веке мотивация писателей, обращающихся к творчеству на нескольких языках принципиально иная. Как правило, смена творческого языка вызвана сменой языковой среды обитания и, следовательно, читательской аудитории. Среди писателей, использовавших несколько языков, можно назвать В. Аксенова (русский, английский), И. Бурихина (русский, английский, немецкий), В. Ерошенко (русский, японский, эсперанто), В. Линденберга (русский, немецкий), Ю. Мамлеева (русский, английский), С. Шестаковского (русский, еврейский) и некоторых других. Разумеется, все эти писатели занимают различное положение в истории литературы.

В. Набоков продолжает традицию разграничения переводческой деятельности и творчества на нескольких языках, но на принципиально новом уровне. Можно заметить различие в подходе В. Набокова к переводу чужих текстов и собственных. Показательно, что Набоков благосклонно относился к переводам своих произведений, ограниченным кругом людей (в первую очередь женой и сыном), между тем свои стихотворные тексты он не доверял никому и переводил только сам, даже поэтические вставки в «Даре» (включая записанную в строчку рифмованную концовку). Уместно вспомнить, что уже в первом опубликованном Набоковым переводе с английского Lewis Carroll «Alice's Adventures in Wonderland» — «Аня в стране чудес» (Берлин: Гамаюн, 1923) одна из неоспоримых удач — перевод стихотворных каламбуров Кэрролла, который пародирует известные тексты Р. Саути, Дж. Тейлор, И. Уотса, английские пословицы. Набоков использует для своих каламбуров строчки А. Пушкина, М. Лермонтова, Н. Гоголя и др.

В. Набоков, неоднократно декларировавший свою приверженность буквальному переводу, в действительности в собственно переводных текстах и тем более

в автопереводах придерживается традиции транспонирования иноязычного текста в русский контекст. Подобный тип приспособления исходного текста к культурным ценностям и стереотипам переводного языка известен достаточно давно и нашел теоретическое обоснование еще в 1791 г. в работе А. Тайлера («Essay on the Principles of Translation»). В дальнейшем этот метод разрабатывался в работах И. Кашкина, К. Чуковского, Г. Гачечиладзе, Ю. Найды и др. Уже в «Ане в стране чудес» В. Набоков стремится сократить дистанцию между первичным и вторичным текстом, устраняя языковые и культурные различия, использует прием компенсации, позволяющий не разрушать структуру текста (восполняющий непонятный для носителя другой культуры компонент, известный любому представителю первичного текста), отказывается от приема отчуждения (буквальный перевод + сноска), который в значительной степени затрудняет чтение, выбирает творческие функциональные эквиваленты, которые способны сделать переводной текст наиболее понятным и естественным для русскоязычного читателя. Это отражено в переводе названия, номинаций, имен персонажей и т. д. Например, замена Англии и Вильгельма Завоевателя Киевской Русью и Владимиром Мономахом в отрывке, который зачитывает мышь; подружка Алисы (Ани), Мейбл становится Асей и т.д. Данный тип перевода оставляет большую творческую свободу переводчику и позволяет создать полноценное художественное произведение. Таким образом, В. Набоков предложил изящное решение главной проблемы художественного перевода: компромисс между адекватностью перевода и уровнем художественной ценности вторичного текста. Именно этноцентрическая адаптация лингвокультурного контекста станет характерной чертой для автопереводов В. Набокова.

В отношении создания собственных произведений ситуация со сменой языковой стихии для писателя чрезвычайно болезненная. В. Набоков приоткрыл для читателя сложность перехода на другой творческий язык в предисловии к русскому изданию «Других берегов»: «Переходя на другой язык, я отказывался, таким образом, не от языка Аввакума, Пушкина, Толстого — или Иванова, няни, русской публицистики — словом, не от общего языка, а от индивидуального, кровного наречия. Долголетняя привычка выражаться по-своему не позволяла довольствоваться на новоизбранном языке трафаретами, — и чудовищные трудности предстоящего перевоплощения, и ужас расставания с живым, ручным существом ввергли меня сначала в состояние, о котором нет надобности распространяться; скажу только, что ни один стоящий на определенном уровне писатель его не испытывал до меня» (СС, т.4, 133). Вспомним, что В. Набоков вырос в семье, следовавшей английскому укладу в противоположность большинству русских дворянских семей.

Тематическая структура (полученная с помощью программы TextAnalyst, см. Рис. 1) автобиографического романа «Другие берега» (1954) представляет структуру текста и описывает содержание в виде иерархии связанных тем и подтем, которые отражают основные понятия и соответствуют узлам сети понятий.



Рис. 1. Тематическая структура текста романа В. Набокова «Другие берега»

Семантическая сеть романа, показывает смысловую канву текста, свидетельствует об особенностях повествовательной техники В. Набокова, сложных смысловых вариациях, объединяющих варьирующие лейтмотивы, складывающихся в изящные тематические узоры (Рис. 2):

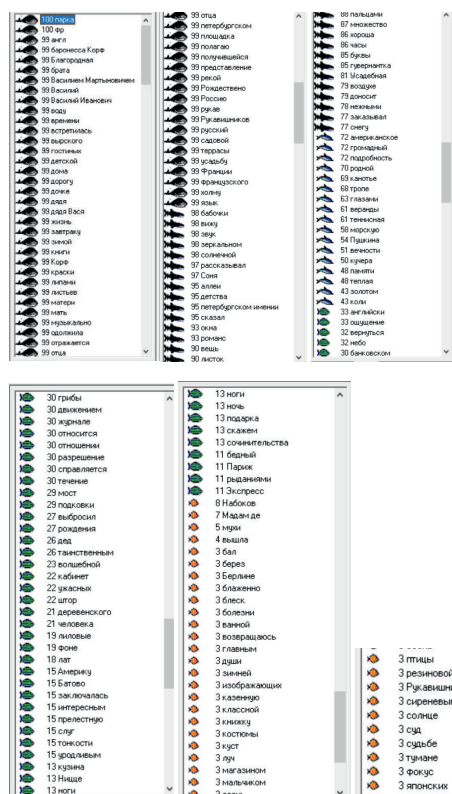


Рис. 2. Семантическая сеть текста романа В. Набокова «Другие берега»

Стиль В. Набокова известен мистификациями, непредсказуемыми поворотами сюжета, непредсказуемыми развязками. Автор создает полифонический текст, в котором соединяются субъективные противоположные точки зрения на описываемые события, взаимоисключающие трактовки происходящего, или вводится необъективный, неадекватный повествователь, т.е. сознательно создается атмосфера неопределенности, недосказанности.

Мистификации, травестирирование достигали предельного значения в интервью писателя, цитаты из некоторых часто приводятся как важные характеристики его творчества, окончательно запутывая некоторых читателей и даже исследователей.

Например, часто цитируются фразы В. Набокова из следующих интервью, без учета характера изданий и пародийной манеры автора при общении с журналистами:

1. «Я американский писатель, рожденный в России, получивший образование в Англии, где я изучал французскую литературу перед тем, как на пятнадцать лет переселиться в Германию» (Playboy 1964 г.).
2. «Моя голова разговаривает по-английски, мое сердце — по-русски, и мое ухо — по-французски» (интервью Джейн Ховард (Jane Howard), журнал Life от 20 ноября 1964 г.

Напомним, что Life — развлекательный журнал, в период данного интервью с сильным уклоном в фотожурналистику. Playboy — скандально известный эротический журнал для мужчин.

В свете данных ассоциативного поиска (см. ниже), очевидно, что не вполне корректно буквально интерпретировать первую цитату, а следует расценивать её в духе травестирированной манеры В. Набокова общения с журналистами. Кроме того, следует вспомнить, что В. Набоков мог даже разыграть журналиста, серьезно рассуждая в интервью о некоем своем литературном произведении, которого никогда не существовало.

В творчестве В. Набокова стремление спрятать истинные чувства под маской иронии проявлялось в столкновении в рамках определенного контекста стилистически, оценочно, коннотационно и семантически противопоставленных элементов. Подобный прием будет в дальнейшем с большим энтузиазмом заимствован писателями различных литературных направлений. Таким образом, важные для В. Набокова семантические акценты оказываются в подтексте, роль имплицитной информации становится ведущей. Анализ лексических ассоциаций, выделенных с помощью технологий TextAnalyst 2.0, даёт материал для анализа имплицитной информации, позволяет в определенной степени судить об аксиологической парадигме автора, которую он пытался тщательно скрыть от внешних наблюдателей.

Очевидно, что проблема вынужденного перехода на английский язык в художественном творчестве — серьезная проблема, которая стояла перед В. Набоковым.

Результаты ассоциативного поиска подтверждают, насколько болезненным был данный процесс для писателя, что английский язык, Англия имели для Набокова ценность, прежде всего, как отражение детства, как элемент счастливого утраченного мира (см. Рис. 3):

Англия - [10/33]

Совершенно владею с младенчества и английским и французским, я перешел бы для нужд сочинительства с русского на иностранный язык без труда, будь я, скажем, Джозеф Конрад, который, до того как начал писать по-английски, никакого следа в родной [польской] литературе не оставил, а на избранном языке [английском] искусно пользовался готовыми формулами.

Предлагаемая русская книга относится к английскому тексту, как прописные буквы к курсиву, или как относится к стилизованному профилю в упор глядящее лицо: «Позвольте представиться, – сказала попутчик мой без улыбки, – моя фамилия N.»

Книга «Conclusive Evidence» писалась долго (1946–1950), с особенно мучительным трудом, ибо память была настроена на один лад – музыкально недоговоренный русский, – а навязывалась ей другой лад, английский и обстоятельный.

Тут я мог бы невероятными подробностями взбесить самого покладистого читателя, но ограничусь только несколькими словами о русском алфавите: латинский был мною разобран в английском оригинале этой книги.

Когда, в 1940 году, я решил перейти на английский язык, беда моя заключалась в том, что перед тем, в течение пятнадцати с лишком лет, я писал по-русски и за эти годы наложил собственный отпечаток на свое орудие, на своего посредника.

У отца сильная прямая подача в классическом стиле английских игроков того времени, и, сверясь с упомянутой книгой, он все справляется у меня и у брата, сошла ли на нас благодать – отзывается ли драйв у нас от кисти до самого плеча, как полагаются.

Неосторожный гувернер поторопился объяснить мне – в восемь лет – логарифмы, а в одном из детских моих английских журналов мне попалась статья про феноменального индуса, который ровно в две секунды мог извлечь корень семнадцатой степени из такого, скажем, приятного числа, как 3529471145760275132301897342055866171392 [кажется, 212, но это неважно].

По-русски Василий Иванович выражался с нарочитым трудом, предпочитая для разговора замысловатую смесь французского, английского и итальянского.

J. – замечала M-1e Golay, чопорная старая пессимистка, бывшая гувернантка матери, продолжавшая жить у нас в доме, всегда кислая, всегда в ужасных отношениях с детскими англичанками и французженками.

По английскому обычаю, гувернантка привязывала к нашим кроваткам в рождественскую ночь, пока мы спали, по чулку, набитому подарками, а будила нас по случаю праздника сама мать и, деля радость не только с детьми, но и с памятью собственного детства, наслаждалась нашими восторгам при шуршащем развертывании всяких волшебных мелочей от Пето.

Рис. 3. Результаты ассоциативного поиска по запросу «Англия»: тематическое дерево – в левом окне – окне результатов поиска и тематический реферат – в правом

Французский язык предстает у В. Набокова как один из иностранных языков, которыми владел писатель, ассоциируется у него с романсом, музыкальными произведениями.

французский - [10/16]

Совершенно владею с младенчества и английским и французским, я перешел бы для нужд сочинительства с русского на иностранный язык без труда, будь я, скажем, Джозеф Конрад, который, до того как начал писать по-английски, никакого следа в родной [польской] литературе не оставил, а на избранном языке [английском] искусно пользовался готовыми формулами.

Он писал романсы – неалкоголически-журчащую музыку и французские стихи, причем хладнокровно игнорировал все правила насчет учета немого «е».

И уже с совершенной обидой вспоминаю, как наш швейцарец гувернер, коренастый и обычно добродушный Нуазье, брызгал ядовитым сарказмом, разбирая однажды французские стихи и музыку дяди – «October» – лучший его романс.

По-русски Василий Иванович выражался с нарочитым трудом, предпочитая для разговора замысловатую смесь французского, английского и итальянского.

В день же пятнадцатых моих именин он отвел меня в сторону и довольно хмуро, на своем порывистом, точном, старомодном французском языке, объявил меня своим наследником.

Книга «Conclusive Evidence» писалась долго (1946–1950), с особенно мучительным трудом, ибо память была настроена на один лад – музыкально недоговоренный русский, – а навязывалась ей другой лад, английский и обстоятельный.

Когда, в 1940 году, я решил перейти на английский язык, беда моя заключалась в том, что перед тем, в течение пятнадцати с лишком лет, я писал по-русски и за эти годы наложил собственный отпечаток на свое орудие, на своего посредника.

Переходя на другой язык, я отказывался таким образом не от языка Авакума, Пушкина, Толстого – или Иванова, няни, русской публицистики – словом, не от общего языка, а от индивидуального, кровного наречия.

Недостатки объяснялись такие, так отшатнувшись тараясь иная фраза, так много было и пробелов и лишнего пояснений, что точный перевод на русский язык был бы карикатурой Мисозинны.

Предлагаемая русская книга относится к английскому тексту, как прописные буквы к курсиву, или как относится к стилизованному профилю в упор глядящее лицо: «Позвольте представиться, – сказала попутчик мой без улыбки, – моя фамилия N.»

Рис. 4. То же, что и на рис. 3, только для запроса «французский»

Ср.: «Совершенно владею с младенчества и английским и французским, я перешел бы для нужд сочинительства с русского на иностранный язык без труда, будь я, скажем, Джозеф Конрад, который, до того как начал писать по-английски, никакого

следа в родной (польской) литературе не оставил, а на избранном языке (английском) искусно пользовался готовыми формулами»

Между тем, минимальный результат дает ассоциативный поиск по запросу *американский* — только лексемы *русский* (67) и *книги* (67).

Дом — ключевое понятие, вокруг которого строится повествование (максимальный вес связи в тематической структуре и семантической сети — 99). Дом для В. Набокова — это петербургский (а после — деревенский) дом, гостиная, парк, детская, завтрак в кругу родных и пр. (см. рис. ниже). Очевидно, иного наполнения для понятия дом для В. Набокова не существовало до конца жизни. Уместно вспомнить, что во время обучения в Кембриджском университете В. Набоков основал Славянское общество, которое в дальнейшем стало Русским Обществом Кембриджского университета. В Швейцарии писатель жил в отеле, не покупая собственного дома, несмотря на финансовую свободу, которую ему подарил гонорар за «Лолиту» и киносценарий в начале 1960-х гг., когда писатель смог оставить преподавание в Америке и вернуться в Старый Свет.

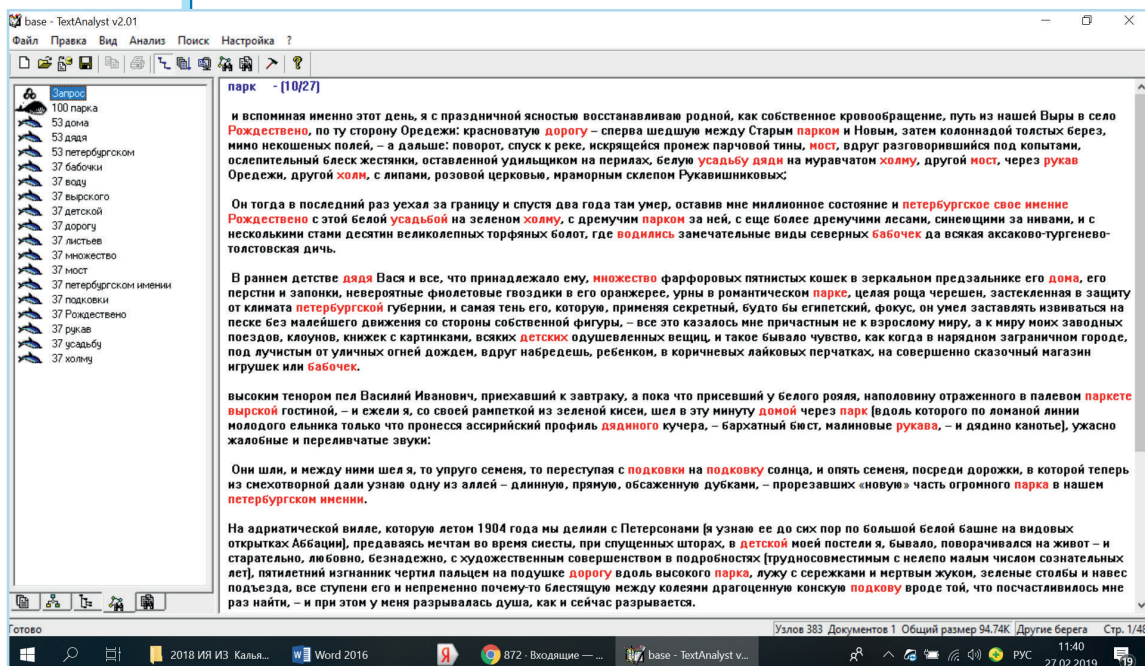


Рис. 5. То же, что и на рис. 3, только для запроса «дом»

Любопытные результаты предоставляет анализ данных с понятием *парк*. В тематической структуре данная лексема вообще не представлена, т.е. во внешнем тематическом представлении текста лексема отсутствует, зато в семантической сети данное понятие доминирует, имеет вес связи 100. Наивысший ранг, который имеет слово *парк* характеризует его как наиболее значимое, несущее основную смысловую нагрузку, составляющее ядро информации, наиболее важные семантические акценты для автора.

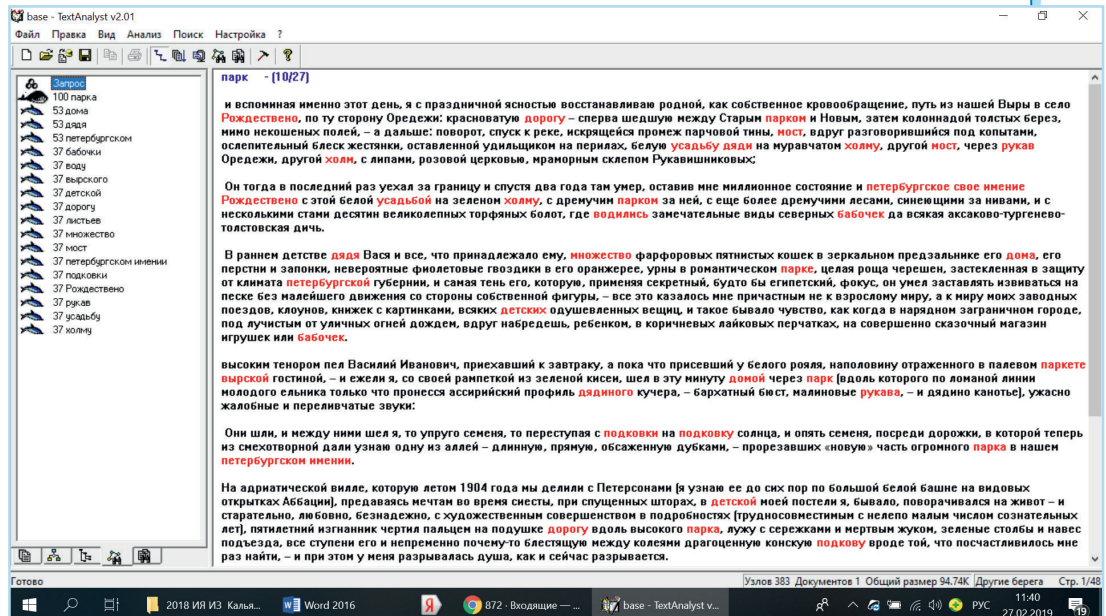


Рис. 6. То же, что и на рис. 3, только для запроса «парк»

Ср.: «Тогда-то я вдруг понял, что двадцатисемилетнее, в чем-то бело-розовом и мягком, создание, владеющее моей левой рукой, — моя **мать**, а создание тридцатитрех-летнее, в бело-золотом и твердом, держащее меня за правую руку, — отец. Они шли, и между ними шел я, то упруго семеня, то переступая с подковки на подковку солнца, и опять семеня, посреди дорожки, в которой теперь из смехотворной дали узнаю одну из аллей — длинную, прямую, обсаженную дубками, — прорезавших «новую» часть огромного **парка** в нашем **петербургском** имени» (...) «Пикники, спектакли, бурные игры, наш таинственный вырский **парк**, прелестное бабушкино Батово, великолепные витгенштейновские имения — Дружноселье за Сиверской и Каменка в Подольской губернии — все это осталось идилически гравюрным фоном в памяти, находящей теперь схожий рисунок только в совсем старой русской литературе» (В. Набоков Другие берега).

Аналогичным образом в автобиографическое эссе И. Бродского «Набережная неисцелимых» можно с помощью нейросетевых технологий провести анализ на разных уровнях семантической глубины и выделить несколько уровней в информации, выраженной эксплицитными средствами (тематическая структура и семантическая сеть) и получить данные для анализа подтекстовой информации с помощью ассоциативного поиска.

Тематическая структура текста представляет собой 5 лексем: городе (100), книгами (99), Сан (99) мрамор (96), мост (61) и отражает внешнюю канву эссе, посвященного Венеции, городу с величественными мостами, обилием мрамора:

«Затем небо на мгновение затмила гигантская *мраморная* скобка *моста*, и вдруг все залил свет».

Семантическая сеть (см. Рис. 7) позволяет выявить ядро информации, определить важные для автора нюансы.

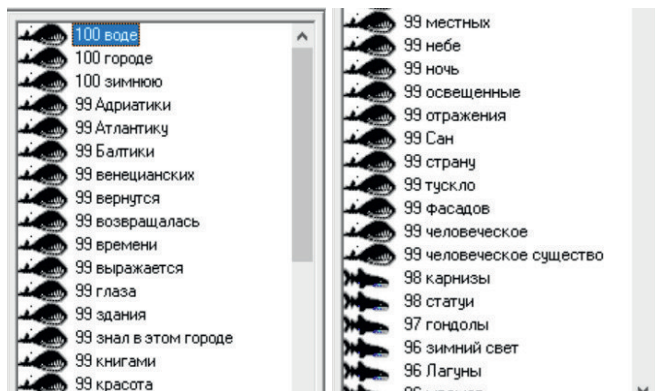


Рис. 7. Тематическая структура текста романа И. Бродского «Набережная неисцелимых»

Водная стихия в зимнем городе — важнейшие семантические акценты для И. Бродского (вес связей — 100), вокруг которых строится смысловой рисунок текста. Между тем, результаты ассоциативного поиска показывают, что волновало автора: прежде всего сравнение прекрасных мест на Адриатике и Атлантике с Балтикой, отражение в прекрасной Италии родных для автора мест:

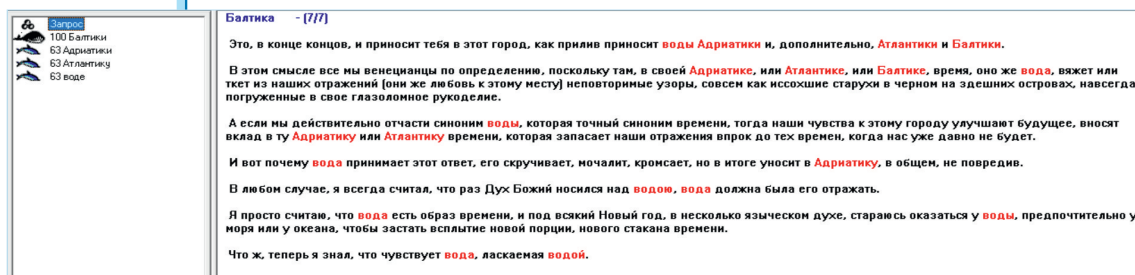


Рис. 8. Результаты ассоциативного поиска по запросу «Балтика»: тематическое дерево — в левом окне — результатов поиска и тематический реферат — в правом

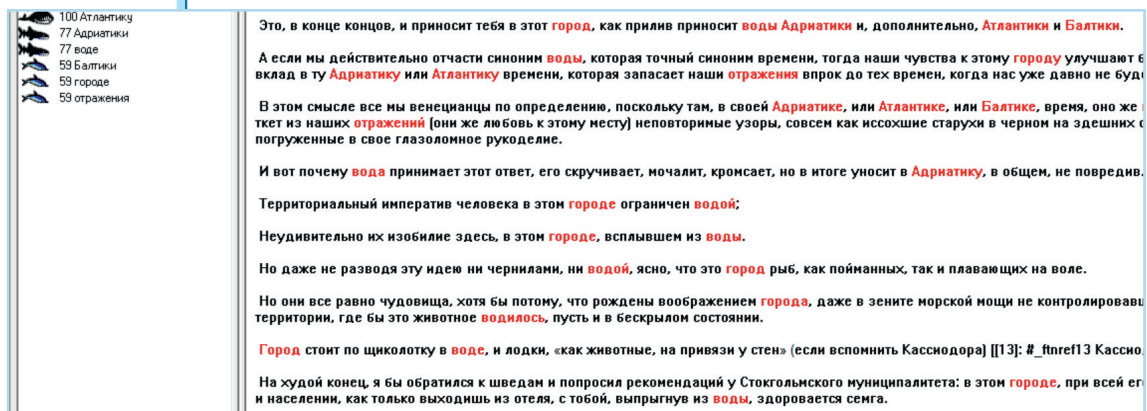


Рис. 9. То же по запросу «Атлантика»

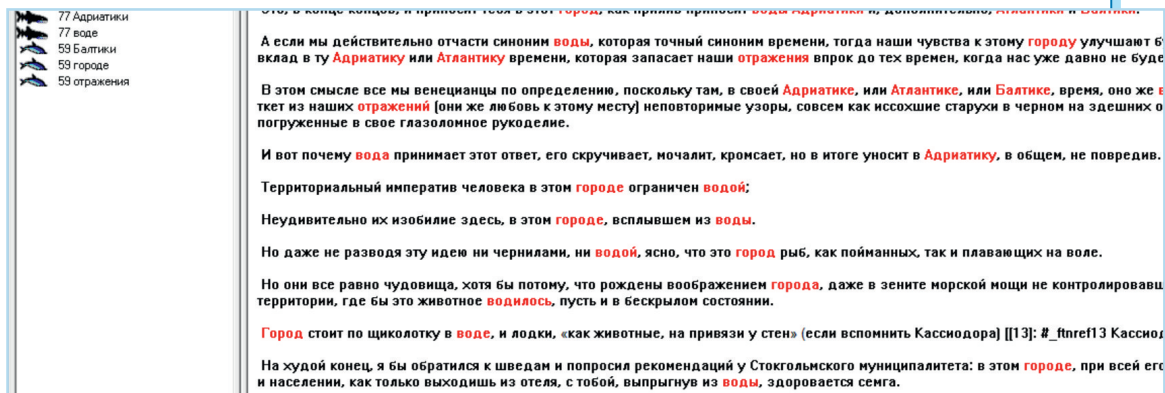


Рис. 10. И наконец, по запросу «Адриатика»

Заключение

На материале произведений «Другие берега» В. Набокова, «Набережная неисцелимых» И. Бродского были протестированы возможности нейросетевого подхода TextAnalyst для анализа восприятия и интерпретации текста, выявления и изучения имплицитной текстовой информации.

Гипотеза, выдвинутая в исследовании, подтвердилась: нейросетевой подход, использованный в технологии TextAnalyst выявляет потенциал ассоциативных связей на нескольких уровнях. Эксплицитная информация выявляется на уровне семантической сети, и имплицитная информация выявляется с помощью ассоциативного поиска.

Автоматический анализ текста, выполненный с помощью этой технологии позволяет проводить исследование на нескольких уровнях. Позволяет определить объективную информацию, выраженную эксплицитно (тематическую структуру текста — внешнюю канву повествования, семантическую структуру — смысловую основу текста, содержащую объективную информацию, заложенную автором). Между тем, ассоциативный поиск выявляет лексические ассоциации, которые свидетельствуют об оценочной парадигме, интенции автора, позволяют делать выводы об имплицитной информации, коннотативных значениях.

Таким образом, нейросетевой подход при использовании технологии TextAnalyst позволяет определить фактографическую основу сюжета, которая формирует тематическую структуру, семантический портрет текста, семантические акценты, главную для автора информацию, представленную в семантической сети, а подтекстовую имплицитную информацию, оценочность и коннотативные нюансы позволяют выделить операции, выполненные с помощью ассоциативного поиска.

References

1. Blei. D. Probabilistic topic models // Communications of the ACM. 2012. Vol. 55. No. 4. P. 77–84.
2. Cheng WC., Liou CY. (2008) Binary Kernel in Morphological Associative Memory. In: Wang R., Shen E., Gu F. (eds) Advances in Cognitive Neurodynamics ICCN 2007. Springer, Dordrech

3. Collins, A. M., and M. R. Quillian. «Retrieval Time from Semantic Memory.» *Journal of Verbal Learning and Verbal Behavior* 8 (1969): 240–248.
4. Daud A. Using Time Topic Modeling for Semantics-Based Dynamic Research Interest Finding // *Knowledge-Based*. 2012. Vol. 26, P. 154–163.
5. De Deyne, S., & Storms, G. (2008). Word associations: Network and semantic properties. *Behavior Research Methods*, 40(1), 213–231. <https://doi.org/10.3758/BRM.40.1.213>.
6. Dennis, S. (2003). A comparison of statistical models for the extraction of lexical information from text corpora. In *Proceedings of the Twenty-Fifth Conference of the Cognitive Science Society*. <http://www.cognitivesciencesociety.org>.
7. Ekmekci, A., Sahin, A., Gulacar, O., Almus, K. (2018). High School Students' Semantic Networks of Scientific Method in an International Science Olympiad Context. *Eurasia Journal of Mathematics, Science and Technology Education*, 14(10), em1604. <https://doi.org/10.29333/ejmste/93677>
8. Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *the mental lexicon* (pp. 25–43). New York: Nova Science Publishers.
9. File, B., Keczer, Z., Vancsó, A. et al. *Behav Res* (2018). Emergence of polarized opinions from free association networks *Behavior Research Methods*. <https://doi.org/10.3758/s13428-018-1090-z>
10. Gruenenfelder, T. M., Recchia, G., Rubin, T., & Jones, M. (2016). Graph-Theoretic properties of networks based on word association norms: Implications for models of lexical semantic memory. *Cognitive Science*, 40, 1460–1495. <https://doi.org/10.1111/cogs.12299>.
11. Guida, A., & Lenci, A. (2009). Semantic properties of word associations to Italian verbs. *Italian Journal of Linguistics*, 19(2), 293–326.
12. Hamers, J., & Blanc, M. (2000). Neuropsychological foundations of bilinguality. In *Bilinguality and Bilingualism* (pp. 135–161). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511605796.009.
13. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* 79, 2554 — 2558 (1982).
14. Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. (2015). The nature and neural correlates of semantic association versus conceptual similarity. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhv003>. CrossRefPubMedGoogle Scholar) A. A. Харламова (2019).
15. *Journal of Psycholinguistic Research* (2019) 48: 243. <https://doi.org/10.1007/s10936-018-9601-8>.
16. Kharlamov AA, Le M. K. (2017) Neyroseteviyе podkhody k classificatsii textov na osnove morphologicheskogo analiza. In: *Trudy MFTI*. 9(2):143–150. (In Russian) A. A. Харламов, Ле Мань Ха Нейросетевые подходы к классификации текстов на основе морфологического анализа ТРУДЫ МФТИ. 2017. Том 9, № 2, с. 143–150.
17. Lerique S. & Roth C. (2018) The Semantic Drift of Quotations in Blogspace: A Case Study in Short Term Cultural Evolution. *Cognitive Science* 42 (1):188–219 (2018).
18. Madan, C.R., Lau, C.S., Caplan, J.B., Fujiwara E. Emotion selectively impairs associative memory. *BMC Neurosci* (2009) 10 (Suppl 1): P341. <https://doi.org/10.1186/1471-2202-10-S1-P341>.
19. Mimno D., Blei D. Bayesian Checking for Topic Models // *Empirical Methods in Natural Language Processing*, 2011. P. 227–237.
20. Miyata, R., Ota, K. & Aonishi, T. Optimal pair of hippocampal CA1 phase response curve and spike-timing-dependent plasticity for hetero-associative memor. *BMC Neurosci* (2013) 14(Suppl 1): P9. <https://doi.org/10.1186/1471-2202-14-S1-P9>.

21. Mouthon, M., Khateb, A., Lazeyras, F., Pegna, A., Lee-Jahnke, H., Lehr, C., & Annoni, J. (n.d.). (2018). Second-language proficiency modulates the brain language control network in bilingual translators: An event-related fMRI study. *Bilingualism: Language and Cognition*, 1–14. doi:10.1017/S1366728918001141.
22. Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, 28 887–899.
23. Ornager, S. (1997) Image retrieval: theoretical analysis and empirical user studies on accessing information in images. In: In: Proceedings of the 60th ASIS annual meeting. Washington DC, November 1–6, 1997. Medford, NJ, 1997. pp. 202–2014/
24. Pejtersen, A.M. (1991). Interfaces based on associative semantics for browsing in information retrieval. Roskilde, Riso Laboratory, 1991.
25. Sinha, A., Davey, N., Adams, R. Steuber V. Structural plasticity and associative memory in balanced neural networks with spike-time dependent inhibitory plasticity. *BMC Neurosci* (2015) 16(Suppl 1): P235. <https://doi.org/10.1186/1471-2202-16-S1-P235>.
26. Steyvers M.k & Tenenbaum J. B. (2005) The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth *Cognitive Science* 29 (2005) 41–78 https://doi.org/10.1207/s15516709cog2901_3.
27. Steyvers M.k & Tenenbaum J. B. (2005) The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth *Cognitive Science* 29 (2005) 41–78 https://doi.org/10.1207/s15516709cog2901_3.
28. Strogatz, S. H. (2001, March 8). Exploring complex networks. *Nature*, 410,268–276.
29. Trier J. Der deutsche Wortschatz im Sinnebezirk des Verstandes: Die Geschichte eines sprachlichen Feldes. Bd.1. Von den Anfängen bis zum Beginn des 13. Jahrhunderts. — Heidelberg: Winter, 1931. — 20 S.
30. Vivas, L., Manoiloff, L., García, A.M., Lizarralde, F., Vivas, J. (2019) Core Semantic Links or Lexical Associations: Assessing the Nature of Responses in Word Association Tasks.
31. White, K.K. & Abrams, L. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers* (2004) 36: 408. <https://doi.org/10.3758/BF03195589>.
32. Апресян 1974 — Апресян Ю. Д. Лексическая семантика: Синонимические средства языка. М.: Наука. 1974. (Переизд.: Апресян Ю. Д. Избранные труды. Т. 1. М., 1995.
33. Апресян Ю. Д. Лексическая семантика: Синонимические средства языка. М., 1974. С. 175–315.
34. Бродского. И. Сочинения Иосифа Бродского. Том VII. Набережная неисцелимых. Санкт-Петербург: Издательство «Пушкинского Фонда». 2001. — 344 с.
35. Черниговская Т. В. Мозг билингва // Университетское переводоведение. Вып. 3. Материалы III Международной научной конференции по переводоведению «Федоровские чтения» 26–28 октября 2001 г. Спб., 2002. — С 538–543.
36. Д. Н. Шмелева, 1977. Современный русский язык. Лексика. Учеб. пособие. — М.: Просвещение, 1977. — 335 с.
37. Караулов Ю. Н. 1981 Лингвистическое конструирование и тезаурус литературного языка. М.: Наука, 1981. — 367 с.
38. Караулов Ю. Н. Лингвистическое конструирование и тезаурус литературного языка. М., 1981. С. 148–218.
39. Лайонз Дж. Введение в теоретическую лингвистику. М., 1978. С. 467–507. Лайонз Дж. Введение в теоретическую: монография. М.: Издательство: Прогресс: 1978. — 544 с.
40. Набоков В. Собрание сочинений. Т.4. Другие берега. М.: Правда, 1990. — 164 с. Харламов А. А. Ассоциативная память — среда для формирования пространства знаний. От биологии к приложениям. Дюссельдорф: Palmarium Academic Publishing, 2017.
41. Шмелев Д. Н. Современный русский язык: Лексика. М., 1977. С. 65–130, 183–232.

IMPLICIT KNOWLEDGE IN THE PERCEPTION OF THE TEXT (ON WORKS OF V. NABOKOV AND I. BRODSKY): NEURAL NETWORK APPROACH

Alexander A. Kharlamov, Maria Pilgun

Abstract.

The paper presents the results of the neural network approach applied in psycholinguistic studies of text perception. The TextAnalyst technology was tested to analyze the text perception and interpretation and to reveal the implicit information in the text. The material included works «Other Shores» by V. Nabokov and «Embankment of the Incurables» by J. Brodsky. The autobiographical works were selected for the analysis, since they contain personal information important for the authors' inner life. The neural network approach implemented in the TextAnalyst technology made it possible to use the potential of associative links in the text at several levels: to reveal explicit information at the level of the semantic network and to reveal implicit information using associative search. This type of automatic text analysis made it possible to determine objective information expressed by verbal means: the topic structure of the text as the external outline of the narration, and the semantic structure as the semantic basis of the text containing the objective information worded by the author. Meanwhile, the associative search revealed lexical associations that indicate the author's evaluation paradigm and intentions and lead to some conclusions about the subtextual information and connotative meanings embedded in the text.

Key words: neuronetwork text analysis, perception, Nabokov, Brodsky