РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А.И. ГЕРЦЕНА

ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ В ФИЛОЛОГИИ НАУЧНАЯ ШКОЛА «ПРИКЛАДНЫЕ ИССЛЕДОВАНИЯ ЯЗЫКА И РЕЧИ, ШКОЛА Р.Г. ПИОТРОВСКОГО»

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА МАТЕМАТИЧЕСКОЙ ЛИНГВИСТИКИ

**ПРИКЛАДНАЯ ЛИНГВИСТИКА В НАУКЕ И ОБРАЗОВАНИИ**

**ALPAC REPORT — ПОЛВЕКА ПОСЛЕ РАЗГРОМА**

**Труды VIII Международной научной конференции 24–26 ноября 2016 г.**



HERZEN STATE PEDAGOGICAL UNIVERSITY OF RUSSIA PHILOLOGICAL FACULTY

DEPARTMENT OF EDUCATIONAL TECHNOLOGIES IN PHILOLOGY SCIENTIFIC SCHOOL «APPLIED STUDIES OF LANGUAGE AND SPEECH. RAIMUND PIOTROWSKI SCHOOL»

SAINT-PETERSBURG STATE UNIVERSITY PHILOLOGICAL FACULTY

MATHEMATICAL LINGUISTICS DEPARTMENT

**APPLIED LINGUISTICS**

**IN RESEARCH AND EDUCATION ALPAC REPORT — 50 YEARS AFTER DEFEAT**

**Proceedings of the VIII th International Biannual Conference November 24–26, 2016**

### Санкт-Петербург 2016

**Saint-Petersburg 2016**

УДК 801

ББК 81-1

Т 78

*Печатается по решению организационного комитета*

*конференции «Прикладная лингвистика в науке и образовании» и УМО РГПУ им. А.И. Герцена*

**Организационный комитет VIII международной конференции**

**«Прикладная лингвистика в науке и образовании»**

*Л.Н. Беляева* (председатель), *А.С. Герд* , *В.З. Захаров*, *А.В. Зубов*, *Л.К. Гейхман*, *О.Н. Камшилова* (ответственный секретарь), *Л.В. Рычкова*, *К.Р. Пиотровская*, *В.А. Широков*, *Н.Л. Шубина*

**Рецензенты:**

д-р филологических наук, профессор *М.А. Марусенко* (СПбГУ); д-р филологических наук, профессор *В.Е. Чернявская*

(Санкт-Петербургский государственный политехнический университет)

Т 78 **Труды VIII Международной научной конференции «Прикладная лингвистика в науке и образовании. ALPAC REPORT — полвека после разгрома».** 24–26 ноября 2016. г., Санкт-Петербург. — СПб.: ООО «Книжный Дом», 2016. — 328 с.

ISBN 978-5-94777-406-1

Сборник содержит материалы докладов, представленных на VIII между- народной научной конференции «Прикладная лингвистика в науке и обра- зовании. ALPAC REPORT — полвека после разгрома».

В докладах обсуждаются актуальные вопросы прикладных исследова- ний специальных языков и текстов (терминология, терминография, словари и перевод), разработка и применение методов структурного анализа текста (математическое моделирование, стилеметрия и лингво-статистический анализ), концепции и инструменты создания и ведения корпусов текстов. Анализируются и предлагаются современные технологии для поддержки обучения языкам.

ББК 81-1

ISNB 978-5-94777-406-1

 Коллектив авторов, 2016

 Филологический факультет РГПУ им. А.И. Герцена, 2016

 РГПУ им. А.И. Герцена

***Посвящается юбилею профессора Л.Н. БЕЛЯЕВОЙ In honor of Dr. Sci. LARISA BELIAEVA***



Доктор филологических наук, профессор, академик РАЕН, Заслуженный деятель науки РФ

Лариса Николаевна Беляева Dr.Sci. Larisa Beliaeva,

professor, RANS member, RF Honored Scientist

## СОДЕРЖАНИЕ

***Беляева Л.Н.*** ALPAC REPORT: воспоминания о будущем

(*Beliaeva L. N.* ALPAC REPORT: future reminiscences) 11

[ТЕРМИНОГРАФИЯ, ОНТОЛОГИИ, СЛОВАРИ И ПЕРЕВОД](#_TOC_250006)

***Борисова Е.П.*** Специфика работы переводчика с художественными произведениями авторов-постмодернистов (***Borisova E.P.*** Speciﬁc features of the postmodern ﬁction translation) 17

*Бычкова Т.В.* Аббревиация как новшество терминотворчества (*Bychkova T.V.* Abbreviation as terminologial innovation) 23

***Брега О.Н.*** Обучающий тезаурус как компонент учебно- методического пособия по дисциплине «Перевод специализированного текста» (***Brega O.N.*** Training thesaurus as the component of

the handbook on discipline “Professional text translation”) 28

*Горбунов Ю.И., Горбунова О.Ю.* Тезаурусное моделирование отраслевой терминологии и дидактика перевода (*Gorbunov Y.I.,*

*Gorbunova O.Y.* Thesaurus modelling of specialized terminology

and translation didactics) 34

***Гриц М.А.*** Онтология предметной области как форма представления знаний (***Gritz M.A.*** Domain ontology as a form of knowledge representation) 40

*Джепа Т.Л.* Лексикографическая база данных как элемент

в управлении отраслевой терминологией (***Dzhepa T.L.*** Lexicographic database as an element in industry terminology management) 46

***Зорина М.А.*** Когнитивное моделирование англоязычной терминосистемы корпоративного права (***Zorina M.A.*** Cognitive model building of English terminological system of corporate law) 50

***Калинин С.С.*** Теория перевода и эколингвистика: процесс переводческой деятельности как процесс адаптации системы

к меняющимся факторам среды (***Kalinin S.S.*** Translation theory and ecolinguistics: the translation practice process as adaptation of the system

to the variable environment factors) 56

***Коряковцев М.А.*** Особенности построения онтологии электронного резюме при разработке кадровой информационной системы (***Koryakovtsev M.A.*** Characteristics of the ontology of digital resume

for the human resource information system) 61

***Михаловский П.*** Трехъязычный терминологический словарь для переводчиков — потребности и перспективы (***Michałowski Piotr***. Trilingual terminological dictionary for translators — needs and perspectives) 67

*Мочалова A В., Мочалов В.А.* Язык описания онтологически- семантических правил для вопросно-ответной системы (*Mochalova A.V., Mochalov V.A.* A language to describe ontological-

semantic rules for question-answer system) 73

*Невзорова О.А., Галиева А.М., Невзоров В.Н.* К разработке автоматических методов выявления соответствия грамматических моделей терминов в разноструктурных языках (*Nevzorova O.A.,*

*Galieva A.M., Nevzorov V.N.* On developing automatic methods

of detecting correspondence between grammatical models in structurally different languages) 80

***Прошина З.Г.*** Необходимость изучения опосредованного перевода специалистами западноевропейских языков (***Proshina Z.G.*** Necessity to study intermediary translation by students of West-European

languages) 86

*Рогачева В.Э.* Методы извлечения терминологических эквивалентов из двуязычного корпуса текстов (*Rogacheva V.E.*

[Extraction of equivalent terms from a bilingual text corpus) 91](#_TOC_250005)

*Табанакова В.Д., Левенкова А.Ю.* Семантические критерии отбора терминологических коллокаций в словарь комбинаторного типа (*Tabanakova V.D., Levenkova A.Y.* Semantic criteria of selecting

terminological collocations for a specialized dictionary of collocations) 97

*Трофимов А.Н.* Грамматическая и категориальная предикативность глагола в тексте русской научной аннотации (*Troﬁmov A.N.*

Grammatical and categorical predication of a verb in the text of

a Russian scientiﬁc abstract) 103

*Турганова К.Р.* Выбор вариантного переводного соответствия с помощью корпусного менеджера BNCWEB (*Turganova K.R.*

Variant translation equivalents choice using BNCWEB corpus manager) .109

***Чебанов С.В.*** Либерализм терминологии и стандартизация номенклатуры как средства обеспечения развития научного творчества (***Chebanov S.V.*** Liberalism of terminology and standardization of the nomenclature as means of ensuring of development of scientiﬁc creativity) 115

[КОРПУСНАЯ ЛИНГВИСТИКА](#_TOC_250004)

*Бакурадзе Л.Д., Беридзе М.М.* Моделирование лингвокультурного пространства диалектного острова в грузинском диалектном корпусе (ферейданский диалект грузинского языка) (*Bakuradze L.,*

***Beridze M.*** The linguo-cultural model of a “dialect island” in GDC (Fereydanian dialect)) 121

***Беридзе М.М.*** Концептуализация грузинской языковой действительности в грузинском языке (по корпусным

и лексикографическим данным) (***Beridze M.M.*** Conceptualization of the Georgian linguistic reality in the Georgian language (based on

corpus and lexicographic data)) 127

*Бертеме Е.* Использование корпусов при обучении прагматической составляющей речевых формул (*E. Berthemet.* The use of corpora

in learning of pragmatic part of formulae) 133

*Birzer S., Zinsmeister H.* The utility of colour markup in texts for learners of Russian. 139

*Богданова С.Ю.* Выявление коллокатов ключевого слова как метод корпусной лингвистики (*Bogdanova S.Yu.* Revealing collocates

of the keyword as a method of corpus linguistics) 146

*Гроховский П.Л., Михайлова М.О., Крамскова A.C.* Корпусно- ориентированные методы изучения научного функционального стиля в тибетском языке (*Grokhovskiy P.L., Mikhailov M.O.,*

*Kramskova A.C.* Corpus-oriented methods of studying scientiﬁc writing

in Tibetan) 152

*Захаров В.П., Бенко В., Хохлова М.В.* Кто ищет, тот найдет. Поисковая функциональность корпусов русского языка (*Zakharov V.P.,*

*Benko V., Khokhlova M.V.* Search functionality of Russian language corpora) 158

*Игнатов К.Ю.* «Оксфордский словарь цитат» как корпус: разметка по жанрам высказываний (*Ignatov K.Y. Oxford dictionary of quotations*

as a corpus: annotating the genres of quotations) 164

*Камшилова О.Н., Смульская Е.Д.* Планирование лонгитюдного исследования: большие проблемы маленького корпуса (*Kamshilova O.N.,* Smulskaya E.D. Planning a longitudinal study: big problems of a small corpus) 170

*Кедрова Г.Е., ПотемкинС.Б.* Разработка корпуса народных сказок (*Kedrova G.E., Potemkin S.B.* Development of the folk tales corpus) 176

*Коробейникова Н.Н.* Корпусно-ориентированный подход

к исследованию синонимического ряда темпоральных наречий (***Korobeinikova N.N.*** Corpus-oriented approach to the study

of synonymous temporal adverbs) 182

***Milojković M.*** Сorpus-derived subtext in Russian-English translation 188

***Митрофанова О.А.*** Возможности использования параллельных и сопоставимых текстов в построении тематических моделей корпусов (***Mitrofanova O.A.*** Possibilities of Parallel and Comparable

Texts in the Development of Topic Models for Corpora) 194

*Стриганов В.Н., Трефилова Г.С.* Программа обработки англоязычных текстов корпуса SPbEFL LC (*Stryganov V.N., Treﬁlov G.S.*

[SPbEFL LC texts processing program) 200](#_TOC_250003)

*Татаринова Л.В.* Корпусное исследование конструкции «как бы» (*Tatarinova L.V.* Corpus research of the construction *как* *бы*) 205

*Хохлова М.В.* К вопросу о «забытых» мерах связанности

для выявления словосочетаний (*Khokhlova M.V.* On rarely used

statistical tools for measuring collocability) 209

[АВТОМАТИЧЕСКИЙ И СТРУКТУРНЫЙ АНАЛИЗ ТЕКСТА](#_TOC_250002)

*Андреев В.С.* Соотношение элементов стихотворного текста (корреляционный анализ) (*Andreev V.S.* Interrelation of elements

of verse text (correlation analysis)) 215

*Гумовская Г.Н.* Исследование семантического потенциала текста методами теории информации (*Gumovskaya G.N.* Research

of the semantic potential of the text by methods of the theory of information) 220

*Данильчук М.В.* Маркетинговая фоносемантика в системе нейминга (*Danilchuk M.V.* Marketing phonosemantics in the system of naming) 226

*Десятириков Ф.А., Белоусов А.В., Белоусова А.Г., Десятирикова Е.Н.* Компьютерный анализ речевого психоинформационного воздействия (*Desyatirikov F.A., Belousov A.V., Belousova A.G., Desyatirikova E.N.* Computer analysis of unauthorized speech exposure) 231

*Дубинина Е.Ю.* Разработка модели автоматического выделения информативных фрагментов текста (*Dubinina E.U.* Development

of the model for automatic extraction of informative text fragments) 238

*Мартиросян А.А.* Математическое моделирование как метод исследования текста (*Martirosyan A.A.* Mathematical modeling

as a method of text study) 243

*Нымм В.Р., Пиотровская К.Р., Еремеева Ю.П., Макогон И.С.* Статистические методы в исследовании проблемы компьютерного обучения языку (*Nymm V., Piotrowska X., Yeremeeva Y., Makogon I.* Statistical methods in CALL-problem research) 248

[ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ В ФИЛОЛОГИИ](#_TOC_250001)

*Балуева А.С.* Научный текст в новых условиях функционирования 254

***Баранчеева Е.И.*** Опыт разработки и внедрения электронных курсов по иностранному языку в аппаратно-программной языковой среде eLang (НГТУ, Новосибирск) (***Barancheeva E.I.*** Design and application of foreign language courses in Electronic language environment eLang

(NSTU, Novosibirsk)) 263

***Блох М.Я.*** Коммуникативная ответственность в условиях информационной революции (***Blokh M.Y.*** Communicative responsibility

in conditions of informational revolution) 269

*Гейхман Л.К., Ставцева И.В.* Педагогический потенциал массовых открытых онлайн-курсов (*Geykhman L.K., Stavtseva I.V.* MOOCs Educational Value) 275

*Знаменская Т.В., Пакудина А.А., Ровбо О.Н.* Создание дистанционного курса для обучения специальной лексике в условиях вуза: анализ и поиски решений (*Znamenskaia T.V., Pakudina A.A.,* Rovbo O.N. Developing a professional vocabulary distance learning course: analysis and ﬁnding solutions) 281

***Зубов А.В.*** От учебной программы — к компьютерной программе обучения языкам (***Zubov A.V.*** From educational program — to computer program of teaching languages) 287

*Корнеева Л.И.* Информационно-коммуникационные технологии

в обучении переводчиков **(*Korneeva L.I.*** Information and communication technologies for training would-be translators) 292

*Коханик А.А.* Анализ сайтов для изучения иностранного языка при развитии лексических навыков у обучаемых (*Kokhanik A.A.*

[Analysis of websites for learning a foreign language while developing learners’ lexical skills) 298](#_TOC_250000)

***Кучина С.А.*** Электронная художественная литература: лингводидактический аспект (***Kuchina S.A.*** Electronic literature: linguodidactical aspect) 304

*Носкова Т.Н., Павлова Т.Б., Яковлева О.В.* Социальные медиа: новые образовательные практики и компетенции (*Noskova T.N.,*

*Pavlova T.B., Yakovleva O.V.* Social media: new educational practices

and competences) 310

*Рябцева Н.К.* Современные информационные технологии

и прикладные проблемы преподавания иностранных языков (***Riabtseva N.K.*** New information technologies and their application

in foreign language learning) 316

***Шубина Н.Л., Виландеберк А.А.*** Образовательные программы

«альтернативных форматов» в высшем образовании: конкурентное преимущество или сомнительный риск? (***Shubina N.L.,***

***Vilandeberk A.A.*** Alternative educational programs in Higher

education: competitive advantage or doubtful risk?) 322

***Л.Н. Беляева***

**ALPAC REPORT: ВОСПОМИНАНИЯ О БУДУЩЕМ**

**Аннотация.** История разработки систем машинного перевода и ее перелома в результате отчета ALPAC отражает основные проблемы автоматической пере- работки текстов на естественных языках и развития информационных технологий в целом. Реализация практических систем машинного перевода в индустрии лока- лизации является условием согласованного научного и технического перевода.

**Ключевые слова.** Машинный перевод, отчет ALPAC, индустрия локализации.

***L.N. Beliaeva***

**ALPAC REPORT: FUTURE REMINISCENCES**

**Abstract.** History of machine translation systems and its turning point as a result of the ALPAC report reﬂects the main problems of natural language processing and information technologies development as a whole. Realization of practical machine translation systems in the localization industry is condition of consistent scientiﬁc and technical translation.

**Ключевые слова.** Machine translation, ALPAC report, localization industry.

В современном мире экономики знаний системы машинного пере- вода являются привычным сервисом, используемым как в технологиче- ской цепочке научного и технического перевода, так и самыми разными категориями пользователей. Однако история создания и развития этих систем, их взлеты и поражения отражают путь, который проходили раз- личные информационные технологии, и существование психологических и технических препон. Поэтому рассмотрение этой истории в момент ее перелома представляет особый интерес.

С меморандума Уоррена Уивера, разосланного в 1949 г. (полный текст меморандума см. в [Locke, Booth 1955]), ведет отсчет история практической реализации идеи машинного перевода (МП). Одним из

© БЕЛЯЕВА Лариса Николаевна. — Beliaeva Larisa. Российский государственный педаго- гический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [lauranbel@gmail.com](mailto:lauranbel@gmail.com)

тех, кто воспринял этот меморандум как призыв к действию, был Леон Достерт, человек, создавший современную систему синхронного перевода и основатель разработки системы машинного перевода в Джорджтаунском университете. Именно разработки, проводившиеся под руководством Л. Достерта, позволили провести в 1954 г. так называемый джорджта- унский эксперимент [Hutchins 2004], доказавший на примере перевода 60 предложений с русского языка на английский принципиальную реа- лизуемость подобных систем. Этот эксперимент послужил толчком для создания огромного количества центров МП во всем мире [Hutchins 2000]. Идеи создания систем машинного перевода были интересны как лингвистам, так и специалистам в области активно развивающейся вычис- лительной техники и программирования. За первые годы исследований, проведенных в этой области, были получены результаты, ценность которых осталась неизменной и сегодня. К ним относятся идеи автоматического мор- фологического анализа, разделения грамматики и словаря, формирования словарей слов и оборотов, модульности алгоритмов и словарей, иерархиза- ции задач и передачи параметров, выбора трансфера в качестве основной модели преобразований и уровневой структуры обработки предложения. Однако системы МП можно было реализовать только на существующих тогда «больших» компьютерах, доступ к которым был ограничен, а стои- мость машинного часа велика. Поэтому те центры МП, у которых такие

вычислительные возможности были, оказывались в особом положении.

Благодаря своему знакомству с высшими чиновниками администра- ции США, Л. Достерт обладал особыми возможностями при получении грантов на проведение исследований и в доступе к вычислительным машинам, в частности, к компьютерам Пентагона. Будучи человеком идеи, он использовал эти возможности для реализации собственного проекта машинного перевода, но тем самым ограничивал доступ других исследователей к грантам. Кроме того, с его именем связаны два важных просчета, влияние которых ощущается до сих пор. Во-первых, таким просчетом было растиражированное СМИ предсказание, что в бли- жайшем будущем системы МП заменят переводчиков. Несмотря на то что это будущее, по утверждению Л. Достерта, должно было наступить в конце 50-х гг. прошлого века, переводческое сообщество (особенно в нашей стране) и сегодня воспринимает компьютер как личного врага и конкурента.

«Мифологическая» составляющая этого утверждения состоит в том, что подобная замена возможна только там и только тогда, когда реально известен прототип, который подлежит замене. Только точно зная соответ- ствующее устройство, процесс или процедуру, понимая, как они функцио- нируют или протекают, можно создать адекватную модель на компьютере.

Во всех остальных случаях компьютерная модель представляет собой реализацию некоторой гипотезы об устройстве памяти человека, протека- нии ментальных процессов, способах обработки информации, восприятии и понимании [Беляева 2012].

Вторым важным просчетом было использование заранее подготовлен- ных и/или упрощенных текстов при проведении экспериментов, доказы- вающих высокое качество результатов разработанной в Джорджтаунском университете системы машинного перевода GAT. При таком подходе любая проверка, проведенная на новых текстах, вызывала сомнения в качестве перевода и возможностях системы в целом.

Возникшее на фоне доминирования Л. Достерта недовольство линг- вистического и программистского сообщества привело к созданию в Ака- демии Наук США специального Консультативного комитета по проблемам автоматической переработки текстов на естественных языках (Automatic Language Processing Advisory Committee — ALPAC). Этот комитет должен был объективно проанализировать состояние дел, соответствие уровня затрат и полученных практических результатов как в области машинного перевода, так и в других областях, связанных с автоматической переработкой текстов. Печально знаменитый отчет ALPAC был опубликован в ноябре 1966.

Конкретным результатом этого отчета было прекращение значи- тельного финансирования исследований в области машинного перевода в Соединенных Штатах Америки в течение приблизительно 20 лет. Госу- дарственное финансирование этих исследований на многие годы было прекращено и во всем мире. При отсутствии персональных компьютеров резкое сокращение или прекращение финансирования привело к разви- тию так называемого «машинного перевода без машин», т. е. к разработке алгоритмов, которые не получали программной реализации и верифи- кации. Более существенным результатом и более опасным в теоретиче- ском и практическом плане, возможно, был ясный сигнал, обращенный к широкой общественности и другим научным сообществам о том, что МП представляет собой безнадежную задачу. И по сей день о «провале» идеи МП часто говорят как о неком бесспорном факте.

Однако сегодня прочно забыто, что конкретно говорилось в этом отчете, хотя его тень продолжает «нависать» над работами в области при- кладной и компьютерной лингвистики.

Сам отчет был очень краток — всего 34 страницы, но они сопровожда- лись 20 приложениями, составляющими в общей сложности 90 страниц. Отчет был назван достаточно широко: *Languages and machines: computers in translation and linguistics* (Языки и машины: компьютеры в переводе и лингвистике) и охватывал разные проблемы автоматической переработки текстов.

Заключительные рекомендации ALPAC определяли те направления исследований, которые целесообразно поддерживать, к ним относятся:

1. практические методы определения качества переводов;
2. средства ускорения процесса ручного перевода;
3. методы определения качества и стоимости различных источников переводов;
4. исследование использования переводов для того, чтобы предотвра- щать создание ненужных переводов;
5. исследование причин задержек, возникающих при полном переводе, и средств их устранения;
6. определение относительной скорости и стоимости различных видов автоматизированного перевода;
7. адаптация существующих методов автоматизации редактирования и технологических процессов, используемых при переводе;
8. собственно процесс перевода в целом;
9. создание специальных справочных ресурсов для переводчиков, включая адаптацию глоссариев, созданных для просмотра автоматиче- ского словаря системы машинного перевода.

В то же время в отчете и сопроводительных документах подчеркива- лась необходимость поддержки исследований в области вычислительной лингвистики. Более того, определялась необходимость финансирования серьезных исследований, поскольку лабораторные эксперименты и работа с миниатюрными моделями языка уже доказали сомнительность полу- чаемых результатов. Реальные результаты можно получить только при условии определенных объемов грамматики, словаря и корпуса текстов. Все перечисленные выше направления доказали свою целесообразность и важность в современных информационных технологиях.

Однако воспринятый как доказательство невозможности разработки практических систем МП этот отчет нанес удар по всем исследованиям в области прикладной и компьютерной лингвистики, от которого не сразу, но удалось оправиться. Однако у этого отчета кроме явных следствий, связанных с закрытием большинства центров МП и закрытием государ- ственного финансирования соответствующих разработок, было еще одно имплицитное следствие: хотя в нем Л. Достерт явно не обвинялся, но его ошибки в истории машинного перевода повторялись. Завышенные ожи- дания и неоправданные обещания оказались тем камнем преткновения, о который разбился начинавшийся с большой помпой проект EUROTRA, обещавший разработать универсальную систему машинного перевода с 9 европейских языков, о котором сегодня не помнит даже вездесущий интернет. История машинного перевода — история несостоявшихся амби- ций и предательств, которая для ее участников, к сожалению, повторяется не как фарс, а как трагедия, и неверные шаги Л. Достерта повторяются на каждом новом этапе.

Изменение ситуации в области схемотехники и электронных устройств: распространение персональных компьютеров, переход сети интернет в сферу доступности для разных категорий пользователей создало совер- шенно новые условия для разработки и внедрения систем МП. Кроме того, даже на фоне отчета ALPAC крупные фирмы практически сразу оценили возможности использования систем МП для предварительной подготовки перевода с согласованной терминологией, разрабатывали собственные системы и обязывали переводчиков работать с их результатами.

Эти же изменения привели в начале нового века к созданию и посте- пенному расширению высокотехнологичных сред: промышленных, иссле- довательских, образовательных, что, в свою очередь, потребовало осмыс- ления новых форм передачи информации, извлечения знаний и т. д.

Немаловажным стимулом для реализации идей автоматической переработки текстов и, в частности, МП стало развитие периферии (сканеров, современных носителей информации, диджитайзеров и т. д.), что полностью изменило контекст обработки текста. Сегодня устранена необходимость трудоемкой подготовки текста для последующей обра- ботки в компьютере — в прошлое навсегда ушел этап перфорации текста и многократной проверки ее результатов, что во многом увеличивало стоимость практического использования систем МП.

Новые возможности, многократное увеличение скорости обработки информации, огромные объемы памяти позволили реализовать новые под- ходы в области машинного перевода, стимулировали использование вычис- лительных методов, что привело к созданию систем статистического МП, основывающегося на сочетании систем переводческой памяти и методов перевода на основе примеров, перевода на основе нейронных сетей и т. п. Однако лингвистическая составляющая этого процесса в последнее время развивается мало, вероятно, на этом уровне достигнут некоторый предел.

Новое качество современного процесса перевода определяется тем, что он включен в так называемую индустрию локализации (Localization Industry) [Ачкасов 2016]. Сегодня можно утверждать, что все перечис- ленные выше направления развиваются как в области индустрии лока- лизации, к которой принадлежит и технический перевод, так и в области разработки информационных технологий и сетевых ресурсов. Индустрия локализации вовлекает в свой круг не только собственно перевод, но и адаптацию результата перевода к культуре принимающего языка, решение маркетинговых и технологических задач, задач информационного поиска и извлечения информации. В основе процесса локализации лежит решение задачи управления операциями с терминологией (Terminology Manage- ment), которое представляет собой набор действий, выполняемых для того, чтобы гарантировать, что во всех материалах и документах конкретной организации последовательно используются согласованные термины для

поддержания полного цикла разработки и совершенствования продукции, коммуникации, перевода/локализации и распространения.

Использование лингвистических технологий (электронных баз данных и знаний, систем машинного перевода, тезаурусов, онтологий, систем проверки орфографии, систем доступа к информации по различным сетям передачи данных) стало элементом профессиональной работы перевод- чика и терминолога, а для специалиста — средством извлечения знаний из иноязычного текста.

**Литература**

1. *Ачкасов А.В*. (2016), Индустрия локализации и подготовка переводчиков // Индустрия перевода: материалы VIII Международной научной конференции (Пермь, 6–8.06.2016). Пермь: Пермский национальный исследовательский по- литехнический университет, с. 13–19.
2. *Беляева Л.Н*. (2012), Практикум по работе с системами машинного перево- да. Материалы для учебно-методического сопровождения дисциплины: учебное пособие. СПб.: Книжный дом, 44 с.
3. *Hutchins W.J*. (ed.) (2000), Early years in machine translation. Memoirs and biographies of pioneers // Studies in the history of the language sciences. Amsterdam; Philadelphia, 400 p.
4. *Hutchins W.J*. (2004), The Georgetown-IBM experiment demonstrated in January 1954. // Machine translation: from real users to research: 6th conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA. Proceedings, ed. Robert E.Frederking and Kathryn B. Taylor. Berlin: Springer, pp. 102–114.
5. *Locke W.N., Booth A.D*. (1955), Machine translation of languages: fourteen essays. Cambridge, Mass.: Technology Press of the Massachusetts Institute of Techno- logy, pp. 15–23.

**Literature**

1. *Achkasov A.V.* (2016), Industriya lokalizatsii i podgotovka perevodchikov // Industriya perevoda: materialy VIII Mezhdunarodnoy nauchnoy konferentsii (Perm', 6–8.06.2016). Perm': Permskiy natsional’nyi issledovatel'skiy politekhnicheskiy universitet, s. 13–19.
2. *Belyaeva L.N.* (2012), Praktikum po rabote s sistemami mashinnogo perevoda. Materialy dlya uchebno-metodicheskogo soprovozhdeniya distsipliny: uchebnoe posobie. SPb.: Knizhnyi dom, 44 s.
3. *Hutchins W.J*. (ed.) (2000), Early years in machine translation. Memoirs and biographies of pioneers // Studies in the history of the language sciences. Amsterdam; Philadelphia, 400 p.
4. *Hutchins W.J*. (2004), The Georgetown-IBM experiment demonstrated in January 1954. // Machine translation: from real users to research: 6th conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA. Proceed- ings, ed. Robert E.Frederking and Kathryn B. Taylor. Berlin: Springer, pp. 102–114.
5. *Locke W.N., Booth A.D*. (1955), Machine translation of languages: fourteen essays. Cambridge, Mass.: Technology Press of the Massachusetts Institute of Techno- logy, pp. 15–23.

# ТЕРМИНОГРАФИЯ, ОНТОЛОГИИ, СЛОВАРИ И ПЕРЕВОД

***Е.П. Борисова***

**СПЕЦИФИКА РАБОТЫ ПЕРЕВОДЧИКА**

**С ХУДОЖЕСТВЕННЫМИ ПРОИЗВЕДЕНИЯМИ АВТОРОВ-ПОСТМОДЕРНИСТОВ**

**Аннотация.** Данная статья посвящена рассмотрению специфики работы переводчика с художественными произведениями, выполненными в постмодер- нистском жанре. В работе обосновывается важность понимания философии произ- ведения как основы качественного выполнения переводческой задачи. Результаты данного исследования могут быть интересны переводчикам художественной литературы, поскольку произведения постмодернистов являются своего рода

«квалификационным тестом», проверяющим общий уровень культурологических знаний и эрудиции, а также умение работать с разными литературными жанрами и стилями.

**Ключевые слова.** Перевод, постмодернизм, постмодернистская техника создания текста, поэтическая структура текста, постмодернисты, Барнс, Кинг, Паланик, Кизи.

***E.P. Borisova***

**SPECIFIC FEATURES**

**OF THE POSTMODERN FICTION TRANSLATION**

**Abstract.** The article reviews peculiar competences and skills required for solving various problems occurring within the process of the postmodern ﬁction translation. The author emphasizes that a good translator should understand the philosophic basis of the postmodern literary works that, in their turn, can be regarded as a “qualiﬁcation test” checking the translator’s expertise as well as the knowledge of language and culture.

**Keywords.** Translation, postmodernism, postmodern technique, poetic structure, postmodernists, Barnes, King, Palahniuk, Kesey.

В настоящее время научное сообщество пришло к пониманию того, что перевод — это не просто механическая замена слов из текста ориги- нала на их эквиваленты в языке перевода. Переводчик должен учитывать стиль и сохранять поэтическую структуру текстов. Следовательно, работа с литературными произведениями эпохи постмодернизма, представляю-

© БОРИСОВА Екатерина Павловна. — Borisova Ekaterina. Национальный исследователь- ский университет «Высшая школа экономики», Россия. — National Research University “Higher School of Economics”, Russia. E-mail: [borisova\_ekaterina73@mail.ru](mailto:borisova_ekaterina73@mail.ru)

щими собой комплексные и многоплановые произведения, которые часто выполнены на стыке разных стилей и жанров, безусловно, является серьез- ной и нелегкой задачей для переводчика. Стоит отметить, что характерной особенностью работ авторов-постмодернистов является их особая поэти- ческая структура, ставшая следствием философии данной эпохи. Отсюда следует, что переводчик в первую очередь должен осознать произведение во всей его глубине, понять основные идеи и мысли, выражением которых стало использование автором тех или иных синтаксических и стилистиче- ских средств. Для этого переводчику необходимо иметь обширные куль- турологические знания, высокий уровень эрудированности и, безусловно, профессиональные навыки и компетенции.

Произведения авторов-постмодернистов часто представлены в форме пастиша, что предполагает создание организации текста в виде эклек- тичной конструкции, объединяющей тексты разных стилей и жанров. Каждый стиль, в свою очередь, характеризуется специфической лексикой, синтаксическими конструкциями и стилистическими приемами, и, сле- довательно, переводчик должен обладать навыками и художественного, и технического, и других видов перевода. Более того, для переводчика особенно важно в этом случае уметь быстро «переключаться» от текстов одного жанра к текстам другого жанра с целью сохранения замысла и философии автора.

Например, в романе Дж. Барнса «История мира в 10 ½ главах» в основе поэтической структуры текста лежит именно стилистическое различие глав. Так, первая глава произведения предстает в форме мемуаров жука-древоточца, третья глава — это выдержки из судебного разбирательства эпохи Средне- вековья с большим количеством слов из французского языка и латыни, а в четвертой главе автор включает в текст романа отрывки из поэзии. Таким образом, каждая глава романа представляет собой отдельное законченное произведение, выполненное в своем собственном стиле, что требует от пере- водчика умения работать с текстами разных литературных жанров.

Важно понимать, что авторы-постмодернисты выступают за свободу и спонтанность в интерпретации деятельности человека и ставят перед собой задачу переосмыслить узнаваемые метанарративы. Этим можно объяснить то, что интертекстуальность является главной характеристи- кой литературных произведений рассматриваемой эпохи. Особенное внимание стоит уделить тому факту, что интертекстуальность в данном случае является не просто стилистическим приемом, а художественным принципом, закономерным следствием философии эпохи. Этим можно объяснить наличие в произведениях постмодернистов большого числа аллюзий — ссылок на ранее созданные произведения, мифологические и религиозные тексты, а также на культурные реалии. Из этого следует, что для того, чтобы суметь распознать аллюзивные отсылки, переводчик

должен обладать определенным уровнем знаний в области истории, лите- ратурного творчества и культуры.

К примеру, анализ рассказа Ст. Кинга «Вещи, которые остались после них» показал, что в тексте присутствуют аллюзии на большое разнообра- зие источников. Центральное место в тексте рассказа занимает отсылка к историческому факту — серии терактов, произошедших 11 сентября 2001 года в Нью-Йорке: “*And I know these things happened in late August of 2002, not quite a year after* ***a piece of the sky fell down*** *and everything changed for all of us*” [King 2008]. Ст. Кинг также ссылается на тексты древнегрече- ских мифов: “*I felt like* ***Helen of Troy*** — *stupid but absolutely true…*” [King 2008]. В данном случае речь идет о мифе, рассказывающем о Елене Тро- янской, из-за которой, как гласит предание, разгорелась Троянская война. Кроме того, автор также включает в текст рассказа аллюзии на библейские сюжеты, в частности на притчу Иисуса, изложенную в 10 главе Евангелия от Луки: “*On an afternoon about a week after Sir Scott Staley donned his* ***Good Samaritan armor*** *and successfully battled the fearsome air conditioner..*.” [King 2008]. В данном отрывке главный герой рассказа сравнивает себя с «добрым самаритянином» — так в притче назвали человека, который пришел на помощь несмотря на то, что от него этого ждали меньше всего.

Дж. Барнс, в свою очередь, включает в текст романа «История мира в 10 ½ главах» строки из автоэпитафии Дж. Гея, автора комедии «Опера нищих»: “*In the autumn of that year I was to wedge into the mirror of my college bedsitting-room a piece of paper bearing the following lines:* ***Life’s a cheat and all things shew it/I thought so once and now I know it***” [Barnes 2005: 173–174]. В подобном случае переводчик может либо дать свой собственный перевод, либо использовать перевод другого автора. При переводе данного отрывка можно использовать работу С.Я. Маршака: «*Какая шутка* — *наша жизнь земная! Так думал я. Теперь я это знаю*» [Маршак 1969: 141].

Интертекстуальный подход к созданию текстов также находит свое отражение в том, что в произведениях авторов-постмодернистов меняется синтаксический статус цитаты, что связано с идеей о «цитатности» мышле- ния современного человека. Так, автор может не оформлять цитату как кон- струкцию с прямой или косвенной речью, может не использовать авторский ввод, не оформлять цитату пунктуационно и графически, что, безусловно, переводчик должен сохранить при выполнении перевода. Примером может стать отрывок из романа К. Кизи «Полет над гнездом кукушки»:

*“… mean old Monday morning, you know, boys…” “Yeah, Miz Ratched…*

*“… and we have quite a number of appointments this morning, so perhaps, if your standing here in a group talking isn’t too urgent …”*

*“Yeah, Miz Ratched…”* [Kesey 1962: 2].

Автор также может включать в текст «псевдоссылки», тем самым создавая ощущение наличия еще одного пространства текста:

“…*And the Mommy says, “Try to at least stand up straight.”*

*The kid rolls his shoulders back and imagines the headlights are a ﬁring squad. He deserves pneumonia. He deserves tuberculosis.*

***See also: Hypothermia***…” [Palahnik 2001: 3].

Кроме того, в своих произведениях авторы-постмодернисты часто используют прием синтаксического разъединения и слияния. Это необхо- димо для имитации устной речи и передачи потока сознания автора. Автор как будто фиксирует на бумаге все мысли, приходящие к нему в голову, при этом не придавая им четкой структуры и не пытаясь преподнести текст в соответствии с языковыми нормами: *“The kid stands, shivering now in the glare, trying to not move, and the Mommy keeps working, telling the huge shadow how someday it will teach people everything that she’s taught it. Someday it will be a doctor saving people. Returning them to happiness. Or something better than happiness: peace.*

*It’ll be respected.*

*Someday”* [Palahnik 2001: 2].

Писатели-постмодернисты часто намеренно включают в текст редкие, архаичных и малоизвестные слова, а также прибегают к включениям из дру- гих иностранных языков, что требует от переводчика умения осуществлять лексикографический поиск. В таком случае задача переводчика — сохранить в тексте перевода данные слова, не давать понятных синонимов, а заставить читателя думать. В своем романе Дж. Барнс часто прибегает к включениям из турецкого языка: *“Miss Logan lost count of the times their luggage was detained, or they were told that a* ***buyurulda*** *or special permit would be neces- sary in addition to the* ***tezkare*** *they had already procured…”* [Barnes 2005: 152]. Тем не менее при осуществлении перевода не всегда удается сохранить стилистику тех слов, которые автор намеренно использует. Примером может стать следующий отрывок из романа Дж. Барнса: “*Ham and his brothers happened to be passing his* ***`tent’*** *(they still used the old sentimental desert word to describe their palaces) and called in to check that their alco- holic father hadn’t done himself any harm*” [Barnes 2005: 16-17]. В данном отрывке для обозначения места, где жил Ной, автор использует слово

«**tent**», которое в этом значении является достаточно редким для англо- язычного читателя. По этой причине автор вводит в текст дополнитель- ный комментарий, объясняющий значение данного слова в конкретном контексте. По нашему мнению, наиболее приемлемым переводом «**tent**» на русский язык является слово «**шатёр**». В данном случае для русско- язычной аудитории теряется архаичность термина, что можно объяснить историей русского населения, на протяжении многих веков бок о бок про-

живающего с народами, для которых шатер является традиционным видом жилища. Тем не менее данный перевод позволяет сохранить библейскую стилистику текста, что, на наш взгляд, имеет наиболее важное значение в тексте романа.

Кроме того, авторы могут намеренно включать в текст малоизвестные термины с целью заставить читателя думать над их значением или прибег- нуть к помощи словарей и справочников: “*You know the social moral of the story: the* ***meritocrat*** *becomes infected with snobbery*” [Barnes 2005: 171].

Мы проанализировали перевод романа Дж. Барнса «История мира в 10 ½ главах», выполненный В. О. Бабковым. Переводчик поясняет значение слова меритократ, говоря, что это «человек, выдвинувшийся благодаря своим способностям» [Барнс 2000: 186]. По нашему мне- нию, в данном случае переводческий комментарий является ошибкой, поскольку данный прием облегчает работу читателя и противоречит цели автора романа.

Еще одной характерной особенностью текстов, выполненных в пост- модернистской технике, является наличие ряда ключевых слов, в которые авторы вкладывают наиболее важные идеи произведения. Переводчику необходимо распознать эти ключевые слова, пронизывающие весь текст, для того, чтобы сохранить декоративный и смысловой стержень произведения и соответственно его игровое устройство. В рассказе Ст. Кинга концепт неба *‘sky’*, входящий также в аллюзивную отсылку к теракту 11 сентября 2001 года, является ключевым элементом, красной линией проходящим через весь текст рассказа, создающим внутреннюю поэтическую систему текста и таким образом выполняющим функцию текстообразующей аллю- зии. Читатель сталкивается тем или иным образом с упоминанием «неба» на протяжении всего рассказа: “*…I had thought when he told me what he’d seen. Over drinks in the Blarney Stone Pub on Third Avenue, this had been, about six weeks* ***after the sky fell down***” или “*I told her about looking out my apart- ment window and seeing the seven A.M.* ***sky was perfectly cloudless***…” [King 2008]. Переводчику необходимо обнаружить данный ключевой элемент и сохранить его, поскольку он объединяет текст в единое целое и в то же время проводит параллели с ранее созданными сюжетами и известными ситуациями, т. е. участвует в моделировании всего текстового пространства. Еще одной характерной чертой произведений, выполненных в рассма- триваемой нами технике, является создание авторами и внедрение в текст окказиональных неологизмов, выполняющих игровую функцию: “*I was a normal eighteen-year-old: shuttered, self-conscious,* ***untravelled*** *and sneer- ing…*” [Barnes 2005: 171]. Следовательно, задача переводчика — передать придуманный автором неологизм на языке перевода и при этом сохранить

его игровой характер.

Определенную трудность может представлять передача фонетических игр, аллитерации, имитации звукозаписи и других фонетических приемов: *“****Yessir****, that’s what I came to this establishment for, to bring you birds fun* ***an’*** *entertainment around the* ***gamin’*** *table. Nobody left in that Pendleton Work Farm to make my days interesting any more, so I requested a transfer,* ***ya see****…*” [Kesey 1962: 5].

Переводчику необходимо сохранять и намеренные нарушения в пунк- туации, и графическом оформлении текста. К примеру, Ч. Паланик в своем произведении выделяет наиболее важные с точки зрения смыслового содержания слова с помощью изменения шрифта: *“****TODAY, A MAN CALLED*** *from Long Beach. He left a long message on the answering machine, mumbling and shouting, talking fast and slow, swearing and threatening to call the police, to have you arrested…”* [Palahnik 2004: 14].

Таким образом, качественный перевод постмодернистских текстов — это не просто подбор эквивалентов и придание им красивой формы, а передача текста на языке перевода с учетом стилистики произведений, их поэтической структуры, отличающейся от других литературных жанров. Следовательно, переводчику для того, чтобы качественно выполнить работу, необходимо прежде всего понять дух и суть документа, а в данном случае — философию постмодернизма, породившую свою собственную технику создания текстов. Данная задача переводчика, в свою очередь, допускает отступление от установленных академических правил и техник перевода.

**Литература**

1. *Барнс Дж.* (2000), История мира в 10 ½ главах. М.
2. *Маршак С.Я.* (1969), Собрание сочинений: в 8 т. М. Т. 4.
3. *Barnes J.* (2005), A History of the World in 10 ½ Chapters. London.
4. *Kesey K.* (1962), One Flew Over the Cuckoo’s Nest. New York.
5. *King St.* Just After Sunset (2008), available at: <http://www.twirpx.com/> ﬁle/1450061/ (дата обращения 17.09.2016).
6. *Palahniuk Ch*. (2001), Choke. New York.
7. *Palahniuk Ch.* (2004), Diary. New York.

**Literature**

1. *Barns Dzh*. (2000), Istoriya mira v 10 ½ glavakh. M.
2. *Marshak S.Ya*. (1969), Sobranie sochineniy: v 8 t. M. T. 4.
3. *Barnes J.* (2005), A History of the World in 10 ½ Chapters. London.
4. *Kesey K.* (1962), One Flew Over the Cuckoo’s Nest. New York.
5. *King St.* Just After Sunset (2008), available at: <http://www.twirpx.com/> ﬁle/1450061/ (data obrashcheniya 17.09.2016).
6. *Palahniuk Ch*. (2001), Choke. New York.
7. *Palahniuk Ch.* (2004), Diary. New York.

***Т.В. Бычкова***

**АББРЕВИАЦИЯ КАК НОВШЕСТВО ТЕРМИНОТВОРЧЕСТВА**

**Аннотация.** В статье рассматривается роль аббревиации в образовании новых терминов. Аббревиация — один из продуктивных способов словообразования. При данном безаффиксальном способе словообразования наблюдается линейное усечение слов. Сокращенные слова образуют новую номинативную единицу. Аббревиатурный способ номинации — новшество в сфере терминотворчества. В статье рассматриваются три основных типа аббревиатур и способы их пере- дачи на русский язык.

**Ключевые слова.** Аббревиация, терминология, номинативная единица.

***T.V. Bychkova***

**ABBREVIATION AS TERMINOLOGIAL INNOVATION**

**Abstract.**The paper covers the role of abbreviation in formation of new terminol- ogy. Abbreviation is a productive means of word formation. Abbreviated words become new nominative entities. Abbreviation is a manifestation of language evolution. Three basic types of abbreviated terms are discussed. Ways of their transfer into Russian are illustrated.

**Keywords.** Abbreviation, terminology, nominative entity.

В эпоху научно-технической революции, характеризующуюся интен- сивным развитием науки и техники, наблюдается дальнейшее усложнение процессов и технологий, для номинации которых потребовалось создать многосоставные термины. Они состоят из нескольких слов и имеют линейно-протяженную структуру. Эта тенденция подтверждается кван- титативными исследованиями научно-технических текстов, в частности, данные, полученные в области химического машиностроения, демон- стрируют весьма красноречивые результаты — 71% терминов состоит из более одного слова, 41% слов превышают в длину два слова и, наконец, 20% терминов превышают в длину три словесных позиции [цит. по: Авербух 2005: 67].

Более того, в любой конкретной предметной области термины выра- жают понятия с высоким уровнем обобщения. В плане выражения они представлены либо словами, либо многословными единицами с различной степенью компактности или распространенности. Они образованы по моделям и характеризуются высокой воспроизводимостью. Рост числа сокращений — явление закономерное. Многие ученые склонны объяснять его широкое распространение все убыстряющимся темпом жизни.

© БЫЧКОВА Татьяна Васильевна. — Bychkova Tatiana. Российский государственный соци- альный университет, Россия. Russian State Social University, Russia. — E:mail: tatuana-3V@ yandex.ru

Рассматривая тенденции, действующие в терминологических систе- мах, остановимся на проблеме наличия и активного функционирования закона или, как его еще называют, «принципа экономии». Как мы уже отмечали, термины в терминологических системах, обслуживающих раз- личные научные и технические области деятельности человека, обладают линейной протяженностью, что противоречит тенденции коммуникации к компактности, когда нужно в условиях нехватки времени и роста интен- сивности образа жизни и деятельности человека передать максимальный объем информации за минимальную единицу времени, затратив минимум усилий.

Огромное значение для объяснения процессов, воздействующих на изменения в языке, а именно его эволюцию, имели работы Е.Д. Полива- нова и затем А. Мартине и его ученика М. Мамудяна [Мамудян 1985: 8]. Уже в 1932 г. в рецензии на книгу Р. Якобсона о фонетической револю- ции Е.Д. Поливанов писал о принципе экономии на примере фонетической

системы.

Принцип экономии в языке объяснялся действием эволюционного процесса, что, по мнению Е.Д. Поливанова, есть не что иное «как трудо- вая деятельность по обеспечению коммуникации между членами данного коллектива» [Поливанов 1968: 57].

Мотивирующий фактор языковой эволюции Поливанов понимал как «стремление уменьшить (экономить) расход трудовой энергии <…> именно до тех пор, пока сокращение энергии не угрожает бесплодностью данного процесса». «Экономия трудовой деятельности или проявление

«лени человеческой» сказывается на нескольких уровнях и влияет на изме- нения языка. Эта экономия физиологической, психической деятельности также включает экономию мыслительных процессов» [Поливанов 1968: 81, 83].

В современной лингвистике проблема языковой эволюции получила дальнейшее развитие в области изучения процессов развития профес- сиональных семантических полей. Ю.Н. Сложенкина пишет: «Поскольку терминологические системы являются живыми ветвями общеязыкового древа. Они постоянно эволюционируют. Одним из проявлений языкового развития профессиональных полей следует признать аббревиацию — вто- ричное сокращенное именование некоторых специальных реалий. Новая усеченная лексическая единица называет в форме одного знака исходное полное словосочетание». Ю.Н. Сложенкина разделяет коллективное мне- ние лингвистов в том, что аббревиация — это проявление закона языко- вой экономии. Коммуникант стремится выразить некоторое содержание

«минимумом материальных форм выражения» [Сложенкина 2013: 84].

Под аббревиацией принято понимать процесс образования новых слов. По классическому определению Е.С. Кубряковой, аббревиация —

это «процесс создания единиц вторичной номинации со статусом слова, который состоит в усечении любых линейных частей источника моти- вации и который приводит в результате к появлению такого слова, кото- рое в своей форме отражает какую-либо часть или части компонентов исходной единицы» [Кубрякова 1981: 200].

Наблюдается линейное усечение компонентов исходного слова или словосочетания и образуется новая лексическая единица, состоящая из сокращенных элементов исходного слова или словосочетания. Аббревиа- туры являются характерной особенностью терминологических систем. Они отличаются от исходных терминов в семантическом отношении, так как происходит своеобразное обобщение и сжатие или конденсация информации [Алексеев 1979: 212–224].

В английской литературе для специальных целей: технической, эко- номической и др. встречается большое количество разного рода сокра- щений, причем сокращаются как отдельные слова, так и словосочетания. Насыщенность терминами-аббревиатурами отражает одну из основных тенденций по информационной оптимизации сообщения, так как сложные слова и словосочетания, как правило, громоздки и неудобны как при пись- менной, так и устной коммуникации. Хотя аббревиатуры неотъемлемая и характерная составляющая деловой терминологии и делового дискурса, и за последнее время тенденция образования новых слов путем сокра- щения существующих слов или словосочетаний значительно возросла, многообразие форм создания сокращений в английском языке нередко приводит к затруднениям при их интерпретации в русском языке, особенно в случаях, когда авторы не расшифровывают вводимые в своих публика- циях сокращения, предполагая, что они хорошо известны читателю. При работе с ними важно помнить, что в связи с отсутствием международных и национальных стандартов на сокращения многие английские сокраще- ния не имеют эквивалентов в русском языке. Аббревиатуры терминологи- ческих систем принято классифицировать. Выделяют от трех до восьми классов. Три основных типа следующие: инициальные аббревиатуры, слоговые аббревиатуры или частичносокращенные слова, смешанные аббревиатуры.

Инициальные аббревиатуры составлены из начальных букв компонен-

тов словосочетания. Они представлены двумя типами:

а) инициальные аббревиатуры звукового типа, т. е. читаемые как про- стые слова, ударение падает на первый слог, например: *IBM (International business machine) — компания «Ай-Би-Эм»*;

б) инициальные аббревиатуры буквенного типа, т. е. читаемые как ряд букв, вернее, алфавитных названий, входящих в их состав букв, например *OECD (Organisation for Economic Cooperation and Development) ОЭСР (Организация экономического сотрудничества и развития.*

1. Аббревиатуры слогового типа, т. е. составленные из начальных сло- гов слов, составляющих словосочетание, например: *internet (International network).*
2. Аббревиатуры смешанного типа, например: *X-mas demand (Christ- mas demand — рождественский спрос).*

По мнению специалистов по переводу, сокращение должно на общем основании подчиняться правилам перевода лексики, в основе этого подхода лежит семантическое соответствие между соотносительными единицами данной пары языков. Близость к терминам, т. е. отсутствие коннотаций, еще больше выдвигает на передний план необходимость передачи смысловой стороны аббревиатур. Аббревиатуры передаются на переводящем языке аббревиатурой, транскрипцией, транслитерацией, развернутой формой. В ряде случаев аббревиатуры переводятся, тогда содержание передается соответствующими средствами переводящего языка: эквивалентами, аналогами переносится в текст переводимого языка как есть, сохраняя отчасти или полностью форму (заимствование). Обладая почти исключительно назывной функцией, подобно термину или имени собственному, аббревиатура переводится эквивалентом — названием того же референта — на переводящий язык, а при отсутствии такого — нередко аналогом, названием близкого понятия. По существу, переводом это можно назвать лишь условно, поскольку аббревиатура, как правило, собственного значения не имеет, а является уменьшенным отражением значения исходной единицы. Например, в русском языке есть утвердившийся перевод следующих аббревиатур: *ETUC (The European Trade Union Confederation) — Европейская конфедерация профсоюзов (ЕКП).* В следующих примерах в русском языке отсутствуют утвердив- шиеся соответствия переводов, поэтому возможно создание аббревиатуры на переводящем языке: *The European Association of Craft and Small to Medium-sized Enterprises (UEAPME) — Европейская профессиональная ассоциация предприятий малого и среднего бизнеса (ЕПФПМСБ).* В этом случае перевод осуществляется при помощи других специальных приемов передачи значения того или иного термина, но отсутствие «прямого» эквивалента приводит к частичной потере значения переводимого слова, утрачивается его семантическая составляющая и стилистическая окраска. Перевод развернутой формой применим в тех случаях, когда отсутствует сокращение-эквивалент на переводящем языке, например: *ATM (automated teller machine) — банкомат.*

**Литература**

1. *Авербух К.Я.* (2005), Общая теория термина. Иваново: ИГУ, 2005.
2. *Алексеев Д.А.* (1979), Сокращенные слова в русском языке. Саратов: СГУ, 1979. 362 с.
3. *Кубрякова Е.С.* (1981), Типы языковых значений. Семантика производного слова М.: Наука, 1981. 200 с.
4. *Мамудян М.* (1985), Лингвистика. М.: Прогресс, 1985.
5. *Поливанов Е.Д.* (1968), Статьи по общему языкознанию. М.: Наука, 1968. 368 с.
6. *Сложенкина Ю.В.* (2013), Основы терминологии: лингвистические аспекты теории термина. М.: Либроком, 2013. 120 с.

**Literature**

1. *Averbukh K.Ya*. (2005), Obshchaya teoriya termina. Ivanovo: IGU, 2005.
2. *Alekseyev D.A.* (1979), Sokrashchennye slova v russkom yazyke. Saratov: SGU, 1979. 362 s.
3. *Kubryakova E.S.* (1981), Tipy yazykovykh znacheniy. Semantika proizvodnogo slova M.: Nauka, 1981. 200 s.
4. *Mamudyan M.* (1985), Lingvistika. M.: Progress, 1985.
5. *Polivanov E.D.* (1968), Stat'i po obshchemu yazykoznaniyu. M.: Nauka, 1968. 368 s.
6. *Slozhenkina Yu.V*. (2013), Osnovy terminologii: lingvisticheskie aspekty teorii termina. M.: Librokom, 2013. 120 s.

***О.Н. Брега***

**ОБУЧАЮЩИЙ ТЕЗАУРУС КАК КОМПОНЕНТ УЧЕБНО-МЕТОДИЧЕСКОГО ПОСОБИЯ ПО ДИСЦИПЛИНЕ**

**«ПЕРЕВОД СПЕЦИАЛИЗИРОВАННОГО ТЕКСТА»**

**Аннотация.** Информационный взрыв и глобализация диктуют потребность в способности обновлять знания, не прибегая к помощи технического переводчика. Процесс разработки инструментов, обучающих техническому переводу, довольно сложный и требует значительной подготовки. В этой связи учет требований ученых к качеству перевода научно-технической литературы, содержания компетенции в сфере профессионального общения как комплекса ЗУН и компетенций (с мета- языковой компетенцией как важнейшей ее частью), вытекающих из государствен- ного образовательного стандарта, позволил разработать учебно-методическое посо- бие по дисциплине «Перевод специализированного текста» на основе обучающего тезауруса для эффективной подготовки студентов — будущих химиков в вузе.

**Ключевые слова.** Тезаурусная лексикография, лексические единицы, формирование профессиональной компетенции, металингвистическая ком- петенция, метаязык, интенсионал термина, семантическая сеть, асимметрия, информационно-семиотическая природа, семиотико-синергетический подход, компонентный анализ.

***O.N. Brega***

**TRAINING THESAURUS AS THE COMPONENT OF THE HANDBOOK ON DISCIPLINE “PROFESSIONAL TEXT TRANSLATION”**

**Abstract.** Information glut and globalization causes need for ability to update knowledge without technical translator’s help. The process of translation training tools development is quite difﬁcult. It requires signiﬁcant preparation. Therefore it is neces- sary to consider scientists’ requirements to the quality of the translation of scientiﬁc and technical literature, to the set of knowledge, skills, and competences arising from the state educational standards. As a result of the research a metalinguistic competence was deﬁned as a prominent part of the student’s traning. The study allowed to develop a handbook on the subject “Professional text translation” with the training thesaurus as a component for the effective university preparation of students — the future chemists. **Keywords.** Thesaurus lexicography, lexical units, formation of professional com- petence, metalinguistic competence, metalanguage, the intension of the term, semantic network, asymmetry, information-semiotics nature, the semiotic-synergetic approach,

componential analysis.

Целесообразность использования тезауруса для формирования про- фессиональной компетенции обосновывается тем, что между учебным процессом и совокупностью семантических процессов, происходящих

**©** БРЕГА Ольга Николаевна. — Brega Olga. Тольяттинский государственный университет, Россия. — Togliatti State University, Russia. E-mail: [onmatveeva@tltsu.ru](mailto:onmatveeva@tltsu.ru)

в языковой среде предметной области, проявляется аналогия. Система языка химии, образованная множеством элементарных понятий, выра- жающихся в терминах, дающих название теориям, законам, методам, способам, понимаются как хранилище лингвистических знаков, отражаю- щих в сознании людей обобщенные понятия об объектах окружающей действительности и отношениях между ними. Модельным представлением такой системы (языка в целом или подъязыка) является тезаурусная сеть, учитывающая отношения, в которые вступают понятия и реализующие их лексические единицы (иерархические — гиперонимии, гипонимии; неиерархические — антонимии, синонимии, вариативности, корреляции) [Пиотровский 1999: 24].

В процессе традиционного обучения эта система языка первоначально формируется у преподавателя, трансформируется в процессе творческого мышления в некоторую взаимосвязанную систему терминов (тезаурус) и лишь в такой форме впоследствии может быть передана студентам.

Обращаясь к работам, освещающим проблему обучения техническому переводу, можно перечислить разные нестандартные методы обучения. Ю.В. Сапожников предложил искать уже имеющиеся конструкции в предложенном списке шаблонов и применять их при переводе с учетом дискурса [Сапожников 2003: 158]. Н.В. Владимов устранял перевод- ческие трудности, применяя электронные корпуса текстов [Владимов 2005]. Но часто именно правильный подбор перевода термина является настоящей преградой для студента, что подчеркивается Н.Н. Гавриленко в монографии «Теория и методика обучения переводу в сфере про- фессиональной коммуникации» [Гавриленко 2009]. В ее интегративной модели формирования профессиональной компетенции переводчика среди многочисленных задач переводчика находим и терминологический поиск. Другими словами, по мнению автора, переводчик должен уметь сопостав- лять на двух языках общенаучную и специальную лексику и правильно подбирать эквиваленты, знать способы терминообразования и уметь нахо- дить значение терминов-неологизмов, знать средства терминологического поиска, использовать внешние ресурсы для терминологического поиска и создавать собственную терминологическую картотеку. Учебное пособие А.А. Стрельцова «Научно-технические тексты. От понимания к переводу» состоит из шести модулей, каждый из которых формирует определен- ные компетенции. Соглашаясь со словами автора методики, отметим, что важно способствовать формированию у обучающихся лексического минимума и научной картины мира с учетом достижений XX в. в точных и естественных науках, обучать применять экстралингвистические знания для перевода естественнонаучного текста [Стрельцов 2012].

В соответствии с государственным образовательным стандартом сту-

денты указанной специальности на выходе должны «обладать знаниями

базовой терминологической лексики, базовых лексико-грамматических конструкций и форм; показать понимание прочитанного и прослушанного материала; проявить навыки поиска профессиональной информации, реферирования и аннотирования текстов профессиональной направлен- ности…» [Приказ Минобрнауки России]. Отсюда, по-нашему мнению, составляющей коммуникативной компетенции в устной и письменной формах на иностранном языке для решения задач межличностного и меж- культурного взаимодействия (ОК-5) является метаязыковая компетенция, а одной из целей обучения оказывается актуализация способности инди- вида к профессиональному общению и способности к творческой само- реализации личности в профессиональном плане.

Эффективность формирования профессиональной межкультурной метаязыковой компетенции у студента — будущего химика во многом определяется наличием терминологического словаря по химии, который выполняет не только роль справочника, но интегрирован в учебный про- цесс. Учебно-методическое пособие «Перевод специализированного текста (химия)», разработанное на кафедре «Теория и практика пере- вода» Тольяттинского государственного университета, ориентировано на обучение бакалавров (04.03.01 Химия, 18.03.01 Химическая техно- логия, 18.03.02 Энерго- и ресурсосберегающие процессы в химической технологии, нефтехимии и биотехнологии) основам чтения и письменного перевода специального текста с английского на русский язык.

Разработка пособия подразумевала постановку тесно взаимосвязанных между собой вопросов о целях обучения, содержании образования и фор- мах, методах, приемах и способах передачи этого содержания учащимся. Далее следует представить эту модель и показать, как она реализуется в пособии (см. рис.).

*Цель —* формирование профессиональной компетентности студен- тов посредством использования английского языка в профессиональной деятельности, направленного на правильное понимание и перевод специ- ального текста с английского на русский язык.

Проанализировав содержание компетенции по ФГОС ВО, цель и задачи изучаемой дисциплины, потребности рынка, составители пособия пришли к выводу, что при формировании профессиональной компетент-

Цель → Задачи → Принципы → Комплекс дидактических условий

→ Содержание обучения



→ Технология обучения

→ Средство — ТЕЗАУРУС

**Прогнозируемый результат**

*Рис.* Модель формирования профессиональной межкультурной компетенции специалиста-химика на основе тезауруса

ности у студентов в рамках этой дисциплины решаются задачи, тесно связанные с умением пользоваться терминологическими словарями для решения переводческих задач; знаниями терминологии в области про- фессиональной деятельности на английском языке; готовностью работать в коллективе. Реализовывать такую подготовку возможно при включении заданий, основанных на тезаурусе, решая задачи:

Образовательные: формировать с помощью тезауруса вторичную языковую личность, владеющую определенной суммой знаний о картине мира другой языковой общности, а именно в сфере химии, и способную использовать полученные в процессе обучения метаязыковые знания и умения в своей будущей профессиональной деятельности.

Воспитательные: стимулировать познавательные интересы студентов, создавать положительную мотивацию за счет необычной формы через презентации профессиональной лексики.

Развивающие: накапливать у студентов опыт построения зритель- ных блок-схем, что способствует созданию предпосылок для развития абстрактного, иерархического, визуального мышления. Увеличивать твор- ческие способности вторичной языковой личности при создании и даль- нейшем использовании блок-схем, генерировании идей, что улучшает запоминание содержащейся в блок-схеме информации. Развивать память и интеллект за счет эффективного хранения информации, что означает ее понимание и усвоение.

Комплекс дидактических условий: содержание, технология обучения и средство обучения. При формировании профессиональной межкультур- ной компетенции важным оказывается содержательная сторона, т. е. отбор языкового материала: терминов; базовых словосочетаний, включающихся в контекст дискурса в области химии; отбор текстов, способствующих развитию профессиональной межкультурной компетенции.

Технология формирования профессиональной межкультурной ком- петенции подразумевает активное не только усвоение англоязычных метаединиц или формирование лексического навыка особых единиц языка, но и становление вторичной языковой личности в профессиональном переводческом аспекте, которое в данном исследовании представляется как процесс, состоящий из уровней:

Уровень 1. Восприятие и описание чужой культуры через изучение профессиональной информации: ознакомление студентов с основопола- гающими концептами химии в англоязычных странах.

Уровень 2. Выявление сходств и различий между двумя культурами русскоязычной и англоязычной предметными областями химии. При- чем успешность межкультурной профессиональной коммуникации спе- циалиста гарантируется: 1) когда студент четко ориентируется в системе понятий предметной области на русском языке и видит, как соотносится

уже сформированная «картина мира» с той, что имеется у носителей иноязычной культуры; 2) обучающийся осознает необходимость и спо- собен освоить инокультурный профессиональный опыт; 3) изучение инокультуры предполагает встроенность в нее культуры «англоязычной химии»; 4) определяются значимые отличия между инокультурой как тако- вой и встроенной в нее культурой обеспечения «англоязычной химии»;

1. носителями культуры определяются способности носителей инокуль- туры к адаптации к чужому влиянию, внедрению в собственную культуру, принятию «чужих» правил игры в пространстве подъязыка химии и т. п.;
2. выявляются точки соприкосновения культуры и инокультуры, схожести в моделях внутри разных культур и способности к построению общей модели химии [Павлова 2004].

Уровень 3. Формирование знания об инокультурной профессиональ- ной терминосистеме химии. На данном уровне происходит накопление информации об инокультурной концептосфере, выявляется специфичность составляющих ее концептов и осуществляется отбор знаний, необходимых для встраивания в инокультуру. Носители внедряющейся культуры ведут поиск и накопление ресурсов, которые окажутся жизненно необходимыми для выживания в рамках чужеродного культурного слоя и, более того, слияния с ним и растворения в нем.

Уровень 4. Активация накопленных профессиональных знаний в пере- водческом пространстве (апробация системы межкультурных навыков по применению сформированного профессионально-ориентированного тезауруса на уровне медиатора).

Четвертый уровень межкультурной метаязыковой компетенции представляет собой практическое развитие эффективности культурного знания, накопленного и структурированного внедряющейся культурой. Переводчик для успешной реализации своей деятельности в условиях инокультуры, вооруженный знаниями об инокультуре предметной обла- сти химии, вырабатывает систему таких навыков, которые помогают ему в этой деятельности. Свой профессиональный опыт он адаптирует к чужой для него среде и использует такой новый опыт, как обновленный старый.

Развивая все четыре уровня межкультурной профессиональной компе- тенции, личность приобретает знания и способности интегрировать куль- турные профессиональные различия в новые и более успешные способы творческой профессиональной деятельности.

**Литература**

* 1. *Владимов Н.В.* (2005), Корпусный подход к решению переводческих про- блем (на материале переводов с русского языка на английский) : дис. канд.

филол. наук. М. 198 с.

* 1. *Гавриленко Н.Н*. (2009), Теория и методика обучения переводу в сфере про- фессиональной коммуникации. Кн. 1. М.: Научно-техническое общество имени академика С.И. Вавилова.
  2. *Павлова Л.П*. (2004), Исследование влияния игровой учебной деятельности на формирование межкультурной компетенции студентов вуза (на примере ино- странного языка): дис. на соиск. уч. степ. канд. пед. наук. Ставрополь.
  3. «Об утверждении федерального государственного образовательного стан- дарта высшего образования по направлению подготовки 04.03.01 Химия (уровень бакалавриата)» / Приказ Минобрнауки РФ от 12.03.2015 № 210: КонсультантПлюс [www.consultant.ru](http://www.consultant.ru/) (дата обращения 15.08.2016).
  4. *Сапожников Ю.В.* (2003), Язык французской научно-технической литера- туры и особенности его перевода (На материале подъязыка холодильной техники): дис. канд. филол. наук. М.
  5. *Стрельцов А.А*. (2012), Научно-технические тексты. От понимания к пере- воду. М.

**Literature**

1. *Vladimov N.V*. (2005), Korpusnyi podkhod k resheniyu perevodcheskikh problem (na materiale perevodov s russkogo yazyka na angliyskiy) : dis. kand. ﬁlol.

nauk. M. 198 s.

1. *Gavrilenko N.N*. (2009), Teoriya i metodika obucheniya perevodu v sfere professional’noy kommunikatsii. Kn. 1. M.: Nauchno-tekhnicheskoe obshchestvo imeni akademika S.I. Vavilova.
2. *Pavlova L.P*. (2004), Issledovanie vliyaniya igrovoy uchebnoy deyatel’nosti na formirovanie mezhkul’turnoy kompetentsii studentov vuza (na primere inostrannogo yazyka): dis. na soisk. uch. step. kand. ped. nauk. Stavropol’.
3. «Ob utverzhdenii federal’nogo gosudarstvennogo obrazovatel’nogo standarta vysshego obrazovaniya po napravleniyu podgotovki 04.03.01 Khimiya (uroven’ baka- lavriata)» / Prikaz Minobrnauki RF ot 12.03.2015 № 210: Konsul’tantPlyus [www.](http://www/) consultant.ru (data obrashcheniya 15.08.2016).
4. *Sapozhnikov Yu.V.* (2003), Yazyk frantsuzskoy nauchno-tekhnicheskoy literatury i osobennosti ego perevoda (Na materiale pod'yazyka kholodil'noy tekhniki): dis. ...

kand. ﬁlol. nauk. M.

1. *Strel'tsov A.A.* (2012), Nauchno-tekhnicheskie teksty. Ot ponimaniya k pere- vodu. M.

***Ю.И. Горбунов, О.Ю. Горбунова***

**ТЕЗАУРУСНОЕ МОДЕЛИРОВАНИЕ ОТРАСЛЕВОЙ ТЕРМИНОЛОГИИ И ДИДАКТИКА ПЕРЕВОДА**

**Аннотация.** В статье рассматриваются результаты исследования, направ- ленного на создание лингводидактического инструмента формирования профес- сиональной переводческой компетентности в сфере общественно-политической коммуникации на тезаурусной основе. Таким инструментом является «Обучающий французско-русский общественно-политический тезаурус», который используется в составе учебно-методического комплекса, включающего, помимо самого тезау- руса, учебное пособие по общественно-политическому переводу с французского языка на русский.

**Ключевые слова.** Обучающий французско-русский тезаурус, общественно- политическая терминология, профессиональная переводческая компетентность, дидактика перевода.

***Y.I. Gorbunov, O.Y. Gorbunova***

**THESAURUS MODELLING OF SPECIALIZED TERMINOLOGY AND TRANSLATION DIDACTICS**

**Abstract.** The given article deals with the results of research oriented on linguo- didactic instrument of formation of professional translator’s competence in the sphere of social political communication based on thesaurus approach. Such an instrument is “Educational French-Russian Social Politics Thesaurus” used as a part of educational and methodological complex including textbook on Social Politics translation from French to Russian.

**Keywords.** Educational French-Russian thesaurus, social and political terminology, competence necessary for a professional translator, translation didactics.

Тезаурусный подход к моделированию семантики отраслевой тер- минологии различных предметных областей способствует внедрению в учебный процесс в высшей школе новых словарей-тезаурусов, которые выполняют функцию лингводидактического инструмента, формирующего профессиональную компетентность специалистов различных направлений подготовки, в том числе и лингвистов-переводчиков. Особое место в фор- мировании профессиональной компетентности лингвиста-переводчика занимает подготовка к работе в сфере общественно-политической ком- муникации. В этом направлении накоплен немалый учебно-методический опыт. Тезаурусный подход к формированию профессиональной перевод-

© ГОРБУНОВ Юрий Иванович. — Gorbunov Yuri. Тольяттинский государственный универ- ситет, Россия. — Togliatti State University, Russia. E-mail: [yourigorbounov@tltsu.ru](mailto:yourigorbounov@tltsu.ru)

© ГОРБУНОВА Ольга Юрьевна. — Gorbunova Olga. Тольяттинский государственный университет, Россия. — Togliatti State University, Russia. E-mail: [ole4ka@hotmail.fr](mailto:ole4ka@hotmail.fr)

ческой компетентности разработан и внедрен на кафедре теории и прак- тики перевода Тольяттинского государственного университета [Горбунов 2003; Матвеева 2010; Горбунова 2012]. В настоящей статье предлагается рассмотреть результаты исследования, направленного на создание лингво- дидактического инструмента формирования профессиональной перевод- ческой компетентности в сфере общественно-политической коммуникации на тезаурусной основе.

Обучающий французско-русский общественно-политический тезаурус построен с учетом лексикографических традиций, сложившихся в между- народной научной группе «Статистика Речи», основателем и научным руководителем которой являлся на протяжении нескольких десятилетий доктор филологических наук, профессор Р.Г. Пиотровский (РГПУ им. А.И. Герцена, г. Санкт-Петербург) [Пиотровский 1999]. При этом учиты- вался также значительный опыт развития общей теории лингвистических словарей-тезаурусов [Караулов 1981], достижения прикладной лингви- стики в науке и образовании [Беляева 2007], новые информационные технологии в лингвистике [Зубов и др. 2004], а также результаты исследо- ваний в области компьютерной лингводидактики [Пиотровская 2002].

Основное назначение настоящего тезауруса состоит в его использова- нии в качестве лингводидактического инструмента формирования профес- сиональной компетентности у студентов — будущих переводчиков в сфере специальной общественно-политической коммуникации. Предлагаемый

«Обучающий французско-русский общественно-политический тезаурус» используется в учебном процессе в работе со студентами направления подготовки бакалавров 45.03.02 «Лингвистика» профиля «Перевод и пере- водоведение» и ориентирован на выработку переводческих компетенций в области общественно-политического перевода с французского языка на русский и, наоборот, на практических занятиях по дисциплине «Прак- тический курс перевода второго иностранного языка (французского)».

**«**Обучающий французско-русский общественно-политический тезау- рус» построен в результате предварительного анализа корпуса француз- ских общественно-политических текстов, который включает актуальные материалы французской печатной прессы, а также электронных сайтов крупных французских периодических изданий. Выявленные в ходе анализа французских общественно-политических текстов термины и их определения в текстах сопоставляются с дефинициями этих терминов, приводимых в лингвистических словарях и тезаурусах современного французского языка. В результате компонентного анализа дефиниций французских общественно-политических терминов была выявлена синоптическая схема, которая отражает систему ключевых понятий, вер- тикально и горизонтально связанных между собой сетью семантических отношений. Данная синоптическая сеть состоит из отдельных словарных

статей тезауруса, каждая из которых отражает отдельное семантическое поле ключевого термина предметной области «Общество и политика». Словарная статья «Обучающего французско-русского общественно- политического тезауруса» строится по следующей схеме:

* 1. заглавный термин, отражающий одно из ключевых понятий «тер- минополя» и его русский эквивалент;
  2. варианты термина во французском языке и их русские соответствия;
  3. синонимы французского термина и их русские эквиваленты;
  4. антонимы французского термина и их русские соответствия;
  5. гипероним французского термина с переводом на русский язык;
  6. гипонимы заглавного термина и их русские эквиваленты;
  7. голоним французского термина и его русский эквивалент;
  8. меронимы французского термина и их русские эквиваленты и т. д.

Термины связаны между собой семантическими отношениями, среди которых преобладают:

1. гиперонимия (hyperonymie): от hyperonyme — гипероним: понятие в отношении к другому понятию, выражающее более общую сущность. Так, в рассматриваемой предметной области термин “la politique” («поли- тика») является родовым термином, или гиперонимом, по отношению к терминам “la politique intérieure” («внутренняя политика»), “la politique extérieure” («внешняя политика»), “la politique colonialiste” («колониаль- ная политика»), “la politique conservatrice” («консервативная политика»), “la politique expansionniste” («экспансионистская политика»), “la politique libérale” («либеральная политика»);
2. гипонимия (hyponymie): от hyponyme — гипоним: понятие, выра- жающее частную сущность по отношению к другому, более общему поня- тию. Так термины “la police judiciaire” («судебная полиция»), “la police militaire” («военная полиция»), “la police politique” («политический сыск»), “la police privée” («частный сыск»), “la police secrète” («тайная полиция»), “la police de l’air” («воздушная полиция»), “la police des moeurs” («поли- ция нравов»), “la police de la route” («дорожная полиция»), называющие отдельные типы полиции, связаны отношением гипонимии с термином “police” («полиция»), который является их гиперонимом;
3. меронимия (méronymie): это классификация явлений, основанная на отношениях меронимии и голонимии. Меронимия и голонимия как семантические отношения являются взаимно обратными друг другу. Мероним выражает понятие, которое является составной частью другого, а голоним выражает понятие, которое является целым над другими поня- тиями, т. е. другое понятие является частью первого. Например: термин “corps de police” («состав полиции») является голонимом по отношению к терминам: “agent de police” («полицейский»), “commissaire de police” («комиссар полиции»), “ministre de la police” («министр по делам полиции»),

“préfet de la police” («префект (начальник) полиции»), “indicateur de la police” («полицейский осведомитель»), которые являются его меронимами;

1. cинонимия (synonymie): синонимы — слова одной части речи, раз- личные по звучанию и написанию. Например, термины “société” (обще- ство), “compagnie” (компания) и “enterprise” (предприятие) являются синонимами и нередко замещают друг друга в подобных ситуациях: “la société d’assurance” / “la compagnie d’assurance” (страховое общество / страховая компания), “la société de chemin de fer” / “la compagnie de che- min de fer” (железнодорожное общество / железнодорожная компания), “la société commerciale” / “l’entreprise commerciale” (торговое общество / торговое предприятие);
2. корреляция (сorrélation): от *лат*. correlatio — соотношение, взаимо- связь. Так, например, термин “la classe ouvrière” («рабочий класс») связан отношением корреляции с терминами “lutte de classe” («классовая борьба») и “rapports de classe” («классовые отношения»). Таким же образом про- слеживается корреляция в цепочке терминов “prolétariat” («пролетариат») и “la dictature du prolétariat” («диктатура пролетариата»).

Отличительная черта словарной статьи «Обучающего французско- русского общественно-политического тезауруса» состоит в том, что она имеет открытый характер, т. е. с учетом специальных целей словарная статья может быть расширена за счет новых терминов с помощью особых пометок, которые вводят в тезаурус новую терминологию, отражающую актуальные тенденции эволюции рассматриваемой предметной области. В процессе самостоятельной работы по общественно-политическому переводу с использованием настоящего словаря-тезауруса студент должен следовать алгоритму переводческих действий, который включает следу-

ющие операции:

1. Поиск французского общественно-политического термина в алфа- витном указателе к тезаурусу.
2. Определение адресов семантических полей, в которые входит фран- цузский общественно-политический термин.
3. Последовательное обращение к семантическим полям французского термина по указанным адресам.
4. Выбор русского эквивалента французского термина в результате анализа его семантических полей с учетом реального контекста.

В результате многократного обращения к различным семантическим полям того или иного французского общественно-политического термина в поисках соответствующего русского эквивалента у студента формиру- ется профессиональный тезаурус, отражающий соответствующую пред- метную область. С течением времени в условиях практического обучения общественно-политическому переводу французско-русский общественно- политический тезаурус закрепляется в профессиональном сознании пере-

водчика и в конечном итоге становится показателем его профессиональной переводческой компетентности как специалиста в сфере специальной общественно-политической коммуникации.

Обучающий французско-русский общественно-политический тезаурус предназначен для практического использования в учебном процессе вуза в составе учебно-методического комплекса, который включает, помимо самого тезауруса, учебное пособие по общественно-политическому пере- воду с французского языка на русский и наоборот. Данное учебное пособие содержит актуальный учебный материал, отражающий основные тенденции общественно-политической жизни в мире, а также во Франции и России. Значительное внимание в данном пособии отводится организации само- стоятельной работы студентов при подготовке к практическим занятиям по переводу с целью глубокой проработки и освоения терминологического материала изучаемой отрасли знания с опорой на ее тезаурусную модель.

Предлагаемая методика с использованием специального термино- логического словаря-тезауруса не только способствует эффективному изучению содержания предметной области «Общество и политика», но и содействует развитию у студентов специального гуманитарного мыш- ления, которое необходимо при решении практических задач в будущей профессиональной деятельности лингвиста-переводчика в предметной области специального общественно-политического перевода.

**Литература**

* 1. *Беляева Л.Н.* (2007), Лингвистические автоматы в современных гуманитар- ных технологиях: учебное пособие. СПб., 192 с.
  2. *Горбунов Ю.И.* (2003), Французско-русский грамматический тезаурус: учебное пособие. Тольятти, 152 с.
  3. *Горбунова О.Ю.* (2012), Тезаурусный подход к формированию профессио- нальной компетенции у студентов — будущих лингвистов-переводчиков: Актуаль- ные проблемы теоретической и прикладной лингвистики и оптимизация препо- давания иностранных языков. Памяти профессора Р. Г. Пиотровского. К 90-летию со дня рождения: материалы III международной научной заочной конференции, 3–4 октября 2012 года, г. Тольятти, с. 153–163.
  4. *Зубов А.В., Зубова И.И.* (2004), Информационные технологии в лингвистике: учебное пособие. М., 208 с.
  5. *Караулов Ю.Н.* (1981), Лингвистическое конструирование и тезаурус лите- ратурного языка. М., 368 с.
  6. *Матвеева О.Н.* (2010), Обучающий тезаурус по английскому переводо- ведению. Тольятти, 84 с.
  7. *Пиотровская К.Р.* (2002), Обучающий лингвистический автомат: учебное пособие по курсу «Основы квантитативной лингводидактики». СПб., 40 c.
  8. *Пиотровский Р.Г.* (1999), Лингвистический автомат (в исследовании и непрерывном обучении): учебное пособие. СПб., 256 с.

**Literature**

1. *Belyaeva L.N.* (2007), Lingvisticheskie avtomaty v sovremennykh gumanitarnykh tekhnologiyakh: uchebnoe posobie. SPb., 192 s.
2. *Gorbunov Yu.I*. (2003), Frantsuzsko-russkiy grammaticheskiy tezaurus: uchebnoe posobie. Tol'yatti, 152 s.
3. *Gorbunova O.Yu*. (2012), Tezaurusnyi podkhod k formirovaniyu professional'noy kompetentsii u studentov — budushchikh lingvistov-perevodchikov: Aktual'nye problemy teoreticheskoy i prikladnoy lingvistiki i optimizatsiya prepodavaniya inostrannykh yazykov. Pamyati professora R. G. Piotrovskogo. K 90-letiyu so dnya rozhdeniya: materialy III mezhdunarodnoy nauchnoy zaochnoy konferentsii, 3–4 oktyabrya 2012 goda, g. Tol'yatti, s. 153–163.
4. *Zubov A.V., Zubova I.I*. (2004), Informatsionnye tekhnologii v lingvistike: uchebnoe posobie. M., 208 s.
5. *Karaulov Yu.N.* (1981), Lingvisticheskoe konstruirovanie i tezaurus literaturnogo yazyka. M., 368 s.
6. *Matveyeva O.N.* (2010), Obuchayushchiy tezaurus po angliyskomu perevodo- vedeniyu. Tol’yatti, 84 s.
7. *Piotrovskaya K.R.* (2002), Obuchayushchiy lingvisticheskiy avtomat: uchebnoe posobie po kursu «Osnovy kvantitativnoy lingvodidaktiki». SPb., 40 s.
8. *Piotrovskiy R.G.* (1999), Lingvisticheskiy avtomat (v issledovanii i nepreryvnom obuchenii): uchebnoe posobie. SPb., 256 s.

***М.А. Гриц***

**ОНТОЛОГИЯ ПРЕДМЕТНОЙ ОБЛАСТИ КАК ФОРМА ПРЕДСТАВЛЕНИЯ ЗНАНИЙ**

**Аннотация.** В докладе основные формы представления знаний: семантиче- ские сети, фреймовые системы и онтологии сопоставлены как различные способы моделирования предметных областей. Онтология предметной области представ- ляет собой модель предметной области как множества объектов, на котором зада- ется множество бинарных отношений. Неотъемлемой частью модели предметной области является таксономия классов объектов и таксономия классов бинарных отношений. Единицы таксономий рассматриваются как именованные множества, над которыми осуществляются операции, допустимые языком онтологии.

**Ключевые слова.** Онтология предметной области, семантическая сеть, фреймовая система, дискрипционные логики, моделирование предметной области

***M.A. Gritz***

**DOMAIN ONTOLOGY AS A FORM OF KNOWLEDGE REPRESENTATION**

**Abstract.** The basic forms of knowledge representation: semantic networks, frame systems and domain ontologies, are compared in the paper as different approaches to domain modelling. Ontology as a domain model represents a set of objects with a set of binary relations deﬁned on it and forms taxonomies of classes of objects and binary relations. Taxonomy entities are viewed as named sets exposed to operations determined by an ontology language.

**Keywords.** Domain ontology, semantic network, frame system, description logics, domain modelling.

Основными формами представления знаний являются семантические сети, фреймовые системы и онтологии предметной области. Все они представляют собой определенные способы моделирования предметных областей как совокупностей классов объектов, экземпляров классов и отношений, связывающих классы объектов и экземпляры классов. Соот- ветственно, базовыми единицами представления знаний являются струк- туры, отражающие представления о классах, экземплярах и отношениях моделируемой предметной области.

В семантических сетях классы и экземпляры классов заданной пред- метной области представлены именованными узлами, а отношения, свя- зывающие классы и экземпляры классов, представлены именованными дугами, соединяющими соответствующие узлы. Узлы и соединяющие

© ГРИЦ Мария Александровна. — Gritz Maria. Российский государственный педагогиче- ский университет им. А. И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [maria.gritz@yandex.ru](mailto:maria.gritz@yandex.ru)

их дуги составляют ориентированный граф, а взаимосвязанные графы формируют семантическую сеть. Один из узлов ориентированного графа называет определяемый класс, в то время как отходящие от него дуги и связанные ими узлы выражают наличие у определяемого класса атри- бута — отношения к другому классу предметной области. Атрибут класса наследуется всеми его подклассами и экземплярами, если для них в семан- тической сети не было установлено наличие соответствующего отношения к другому классу предметной области [Люгер 2003; Grimm et al. 2007].

Во фреймовых системах классы и экземпляры классов представ- лены именованными фреймами: фреймами-прототипами и фреймами- экземплярами соответственно, а отношения представлены именованными слотами, связывающими фреймы. Заполненный слот фрейма представляет атрибут соответствующего класса, который наследуется всеми его под- классами и экземплярами, если в обозначающих их фреймах не заполнен соответствующий слот [Люгер 2003; Brachmanetal. 2004]. Во всех семанти- ческих сетях и фреймовых системах выделяют дуги и слоты, представляю- щие отношения «класс-подкласс» и «класс-экземпляр» и формирующие таксономию классов. Прочие дуги и слоты задают отношения классов и экземпляров на созданной таксономии [Brachmanetal. 2004].

Онтология может быть описана как модель предметной области, пред-

ставленной в виде множества объектов Δ*I* , связанных множеством бинар-

ных отношений Δ*I*  Δ*I* , описанная посредством формального языка,

принадлежащего к группе языков, именуемых дискрипционными логи- ками [Guarinoetal. 1995; Guizzardi 2005]. В дискрипционных логиках экземпляры классов представлены посредством констант, классы объектов представлены переменными, а отношения между классами и экземпля- рами классов представлены посредством атомарных высказываний, состоящих из предикатов арности 1 или 2, связывающих соответствующие переменные и константы [Baaderetal. 2003; Grimm et al. 2007].

Переменные обозначают классы объектов как неупорядоченные мно-

жества элементов, которые составляют универсальное множество Δ*I* .

Неупорядоченные множества элементов вступают в отношения включе-

ния: *AI*  *BI*  ( *A* m *B*)*I* (1).На множествах также задаются бинарные

отношения: *RI*  *AI*  *BI* (2), формирующие множества упорядоченных

пар элементов, связанных бинарными отношениями, которые также всту-

пают в отношения включения *RI*  *S I*  (*R* m *S* )*I* (3)*.* Отношения вклю-

чения неупорядоченных множеств элементов и множеств упорядоченных пар элементов, связанных бинарным отношением, описываются атомар- ными высказываниями вида: *S(R), B(A)*, состоящими из предиката арно- сти 1, обозначающего множество, и связываемого им аргумента, обозна-

чающего подмножество данного множества. Цепочки атомарных выска- зываний данного типа формуют таксономию классов объектов как неупо- рядоченных множеств элементов и таксономию связывающих данные множества бинарных отношений. Константы обозначают экземпляры

классов как элементы множеств: *А(a)* при *a*  *AI* , *B(b)* при *b*  *BI* (4),

связанные заданными на множествах бинарными отношениями: *R(a,b)*

при *RI*  *AI*  *BI* , *a*  *AI* , *b*  *BI* (5). Заданное на множествах *AI* и *BI*

бинарное отношение описывается атомарным высказыванием вида: *R(A, B),* состоящим из предиката *R* арности 2, обозначающего множество упорядоченных пар элементов, связанных отношением *RI*, и связываемых им аргументов. Первый аргумент обозначает область определения бинар-

ного отношения *RI*: *DomainR*  *a* Δ*I* |*b*.(*a*, *b*)  *RI*  (6), а второй аргу-

мент обозначает область значений бинарного отношения *RI* :

*RangeR*  *b* Δ*I* |*a*.(*a*, *b*)  *RI*  >> (7) [Baaderetal. 2004; Grimm et al. 2007].

Дискрипционные логики вводят синтаксические правила (см. табл. 1), которые допускают осуществление ряда операций над неупорядоченными множествами элементов, обозначенными переменными E и F, и множе- ствами упорядоченных пар элементов, обозначенными предикатами R и

S. Поскольку множества вступают в отношения равенства при условии тождества их состава, предложенная дискрипционной логикой система синтаксических правил позволяет определить границы некоторого мно- жества, входящего в универсальное множество, посредством осуществле- ния операций над другими множествами, включенными в заданное уни- версальное множество. В число допустимых дискрипционными логиками операций входит пересечение множеств (см. правила 1 и 8), объединение множеств (см. правила 2 и 9), разность множеств или дополнение множе- ства (см. правила 3 и 10) и произведение множеств (см. правило 11). Рядом дискрипционных логик допускается определение границ некоторого мно- жества как области определения бинарного отношения *RI* посредством связывания переменной, обозначающей область значений, квантором всеобщности (см. правило 4) или квантором существования (см. правило 5); определение границ некоторого множества как области определения бинарного отношения *RI* посредством наложения количественных огра- ничений на область значений в виде максимальных/минимальных значе- ний мощности множества (см. правила 6, 7); определение границ некото- рого множества упорядоченных пар элементов, связанных бинарным отношением *RI*, посредством операции транзитивного замыкания над

*R*

1

отношением

*I* (см. правило 12); определение границ некоторого мно-

жества упорядоченных пар элементов, связанных бинарным отношением

*RI*, посредством инверсии бинарного отношения [Baaderetal. 2003].

*R*

1

*I* (см. правило 13)

*Таблица 1.* Синтаксические правила дискрипционных логик

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Правило** | **Интерпретация** | | | **Пример применения синтаксического правила** |
| *E* H *F* (1) | *a* Δ*I* |*a*  *EI*  *F I*  | | | *Woman*  *Adult* H *Female* |
| *E* H *F* (2) | *a* Δ*I* |*a*  *EI*  *F I*  | | | *Person*  *Male* H *Female* |
| *E* (3) | *a* Δ*I* |*a*  *EI*  | | | *Male*  *Female* |
| *R*.*E* (4) | *a* Δ*I* |*b*.(*a*,*b*)  *RI*  *b*  *EI*  | | | *MotherofSons*  *Mother*  H*hasChild*.*Male* |
| *R*.*E* (5) | *a* Δ*I* |*b*.*a*,*b*  *RI*  *b*  *EI*  | | | *Mother*  *Woman*  H*hasChild* .*Person* |
|  *nR* (6)   *nR* (7) | *a* Δ*I* |  *a* Δ*I* | | *b*|*a*,*b*  *RI*   *b*|*a*,*b*  *RI*  |  *n*   *n* | *Motherwithmany* -  *Children*  *Mother*  H 3*hasChild*.*Person* |
| *R* H *S* (8) | (*a*,*b*) Δ*I*  Δ*I* |(*a*,*b*)  *RI*  (*a*,*b*) *SI*  | | | *hasSiblingChild*  *hasDaughter*  H *hasSon* |
| *R* H *S* (9) | (*a*,*b*) Δ*I*  Δ*I* |(*a*,*b*)  *RI*  (*a*,*b*) *SI*  | | | *hasGrandchild*  *hasGrandson*  H *hasGranddaughter* |
| *R* (10) | (*a*,*b*) Δ*I*  Δ*I* |(*a*,*b*)  *RI*  | | | *Isunmarriedto*  *hasHusband*  H*hasWife* |
| *R* ○ *S* (11) | ⎧⎪(*a*, *c*) Δ*I*  Δ*I* |*b*.*a*,*b*  *RI* ⎫⎪  ⎪⎨ (*b*, *c*) *SI* ⎬⎪  ⎩ ⎭ | | | *hasMotherinLaw*  (*hasHusband* H *hasWife*) ○ *hasMother* |
| *R* (12) | (*a*, *c*) Δ*I*  Δ*I* |(*a*,*b*)  *RI*  (*b*, *c*)  *RI*  | | | *hasAncestor*  *hasParent* |
| *R*1 (13) | (*a*,*b*) Δ*I*  Δ*I* |(*b*, *a*)  *RI*  | | | *hasHusband*  *hasWife*1 |

Границы определяемого множества могут быть описаны с помощью равенства, левая часть которого содержит переменную или предикат, обо- значающий определяемое множество, а правая часть содержит логические символы, обозначающие проводимые над множествами операции, а также переменные и предикаты, данные множества обозначающие. Примеры равенств приведены в правых столбцах табл. 1 и табл. 2.

*Таблица 2.* Факты о классах объектов и отношениях классов моделируемой предметной области в различных формах представления знаний

|  |  |
| --- | --- |
| **Семантические сети/ Фреймовые системы** | **Онтологии** |
| *Wife**Spouse**Man*(*Spouse**Wife*  *isMarriedto**Wife*, *Man*) | *Wife*  *Spouse* \ *Husband* ;  *Wife*  *Spouse* H *Woman* ;  *Wife*  *isMarriedto*.*Man* |
| *Husband**Spouse**Woman*  (*Spouse**Husband*   *isMarriedto* *Husband* ,*Woman*) | *Husband*  *Spouse* \ *Wife* ; *Husband*  *Spouse* H *Man* ; *Husband*  *isMarriedto*.*Woman* |
| *Mother**Parent**Person*  (*Parent*(*Mother*)   *hasChild* (*Mother*, *Person*)) | *Mother*  *Parent* \ *Fathe* ;  *Mother*  *Parent* H *Woman* ;  *Mother*  *hasChild*.*Person* H*Woman* |
| *Son**Child* *Person*(*Child* (*Son*)   *hasParent*(*Son*, *Person*)) | *Son*  *Child* \ *Daughter* ;  *Son*  *Child* H *Male* ;  *Son*  *hasParent*.*Person* H *Male* |
| – | *hasSibling*  *hasSister* H *hasBrother*  *hasSibling*  *hasParent* ○ *hasChild* |
| – | *hasChild*  *hasSon* H *hasDaughter*  *hasChild*  *hasParent* 1 |

Факты предметной области, фиксирующие достоверные данные о классах объектов и отношениях классов предметной области, в онто- логиях представлены равенствами, определяющими границы класса объектов или отношения классов предметной области как множества (см. табл. 2). Равенства могут быть также использованы для форма- лизации дефиниций специальных терминов, обозначающих классы и отношения.

В семантических сетях и фреймовых системах факты о классах объ- ектов предметной области представлены триплетами, состоящими из име- нованных узлов, связанных дугами, и именованных фреймов, связанных слотами. Представленные фреймовыми системами и семантическими сетями факты о классах объектов предметной области могут быть описаны формулами, в которых переменные, именующие определяемые классы, связаны квантором всеобщности.

**Литература**

1. *Люгер Д. Ф.* (2003), Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е изд.: пер. с англ. М.: Издательский дом «Вильямс».
2. *Baader F., Nutt W.* (2003), Basic Description Logics / The Description Logic Handbook: Theory, Implementation and Applications, USA, New York: Cambridge University Press, pp. 43–95.
3. *Brachman R. J., Levesque H. J.* (2004), Knowledge representation and reasoning, USA, San Francisco: Morgan Kaufmann Publishers.
4. *Grimm S., Hitzler P., Abecker A.* (2007), Knowledge Representation and Ontolo- gies: Logic, Ontologies and Semantic Web Languages / Semantic Web Services, USA, New York: Springer, pp. 51–105.
5. *Guarino N., Giaretta P.* (1995), Ontologies and Knowledge Bases: Towards a Terminological Clariﬁcation / Towards very large Knowledge Bases: Knowledge Build- ing and Knowledge Sharing, The Netherlands, Enschede, pp. 25–32.
6. *Guizzardi G.* (2005), Ontological Foundations for Structural Conceptual Models, The Netherlands, Enschede.

**Literature**

1. *Lyuger D. F.* (2003), Iskusstvennyi intellekt: strategii i metody resheniya slozh- nykh problem, 4-e izd.: per. s angl. M.: Izdatel’skiy dom «Vil’yams».
2. *Baader F., Nutt W.* (2003), Basic Description Logics / The Description Logic Handbook: Theory, Implementation and Applications, USA, New York: Cambridge University Press, pp. 43–95.
3. *Brachman R. J., Levesque H. J.* (2004), Knowledge representation and reasoning, USA, San Francisco: Morgan Kaufmann Publishers.
4. *Grimm S., Hitzler P., Abecker A.* (2007), Knowledge Representation and Ontolo- gies: Logic, Ontologies and Semantic Web Languages / Semantic Web Services, USA, New York: Springer, pp. 51–105.
5. *Guarino N., Giaretta P.* (1995), Ontologies and Knowledge Bases: Towards a Terminological Clariﬁcation / Towards very large Knowledge Bases: Knowledge Build- ing and Knowledge Sharing, The Netherlands, Enschede, pp. 25–32.
6. *Guizzardi G.* (2005), Ontological Foundations for Structural Conceptual Models, The Netherlands, Enschede.

***Т.Л. Джепа***

**ЛЕКСИКОГРАФИЧЕСКАЯ БАЗА ДАННЫХ КАК ЭЛЕМЕНТ В УПРАВЛЕНИИ ОТРАСЛЕВОЙ ТЕРМИНОЛОГИЕЙ**

**Аннотация.** В статье рассматривается возможность применения методов создания систем управления операциями с терминологией, а именно термино- логических банков данных, к созданию отраслевой лексикографической базы данных предметной области ‘авиационное двигателестроение’. Особое внимание уделяется особенностям построения двуязычного корпуса текстов для последую- щего извлечения терминов. Выявляются частные проблемы работы с корпусом.

**Ключевые слова.** Лексикографическая база данных, корпус текстов, авиаци- онное двигателестроение, система управления терминологией, терминологические банки данных.

***T.L. Dzhepa***

**LEXICOGRAPHIC DATABASE AS AN ELEMENT IN INDUSTRY TERMINOLOGY MANAGEMENT**

**Abstract.** In this article, methods applied for creation of terminology management systems (terminological databanks) are reviewed in reference to building the lexico- graphic database of LSP ‘aircraft engine-building’. Special attention is paid to aspects of creation of a bilingual corpus of texts for the subsequent extraction of terms. The speciﬁc issues in corpus work are revealed.

**Keywords.** Lexicographic database, corpus of texts, aviation engine-building, terminology management system, terminological databank.

Владение отраслевой терминологией не только на родном, но и на иностранном (или нескольких иностранных) языках является важной компетенцией специалистов наукоемких отраслей. Отрасль авиаци- онного двигателестроения, в русле которой развивается ряд научных направлений, ведутся опытно-конструкторские работы и осуществляется внешнеэкономическая деятельность, не является исключением. При этом в отрасли существует потребность в поддержке деятельности специали- стов современными автоматизированными средствами выделения, пере- вода, фиксации терминов и их дальнейшего использования для решения разнообразных прикладных задач. Учитывая интеграционные процессы в отрасли, становится актуальной задача создания инструментария, обес- печивающего согласованность употребления терминов в рамках общего процесса управления.

Ключевым элементом в создании инструментария управления отрас-

левой терминологией может стать англо-русская лексикографическая база

© ДЖЕПА Татьяна Леонидовна. — Dzhepa Tatiana. АО «Объединенная двигателестрои- тельная корпорация», Россия. — JSC “United Engine Corporation” (Russia). E-mail: Tatiana. [dzhepa@yandex.ru](mailto:dzhepa@yandex.ru)

данных (ЛБД), основные этапы формирования которой можно автоматизи- ровать при помощи средств компьютерной лексикографии [Беляева, Джепа 2012]. Создание отраслевой двуязычной ЛБД выполняется с применением методов и процедур систем управления операциями с терминологией.

Практика управления операциями с терминологией определяется стандартами Международной организации по сертификации (ISO) и нацелена на создание многоязычных, гармонизированных и стандар- тизованных терминологических ресурсов. В соответствии со стандартом ISO 26162:2012 под *системой управления операциями с терминологией* (terminology management system) понимается программное обеспечение (software tool), специально разработанное для накопления терминологиче- ских данных, обеспечения их обработки и предоставления пользователям доступа к этим данным [ISO 26162:2012]. Методы и принципы создания систем управления операциями с терминологией реализованы в проектах отраслевых, национальных и межнациональных терминологических бан- ков данных (ТБД) [Towards Consolidation… 2006].

Следуя базовым требованиям к системам управления операциями с терминологией, для создания отраслевой ЛБД предметной области ‘авиа- ционное двигателестроение’ используется программное обеспечение ком- пании *TshwaneDJe* — профессиональная система для создания словарей и списков терминов TLex Suite, интегрированная с корпусным менеджером (*corpus query system*) и адаптированная для создания словарей различных типов, включая двуязычные. Работа с системой может осуществляться как одним, так и группой пользователей и позволяет выполнять несколько этапов процедуры создания лексикографического ресурса, используя при- ложения одной программной среды.

Основной источник терминологических данных в современной лек- сикографии и терминографии — специализированный корпус текстов, тематическая организация которого подчиняется выбранной универсаль- ной классификации знаний, с помощью которой определяются подобласти предметной области [Беляева 2014]. В таких международных проектах, как EuroTermBank, IATE, TERMIUM, для этого используется универсальная предметная классификация Леноха, тезаурус Eurovoc, либо специально разработанная в рамках проекта классификационная система, объеди- няющая в подклассы понятия многих предметных областей [Towards Consolidation… 2006; Kudashev 2013].

Базой для извлечения терминологических данных двуязычной ЛБД по авиационному двигателестроению является сопоставимый корпус тек- стов, в который отбираются тексты, охватывающие различные тематики. В качестве источников создания структурной схемы корпуса текстов используются тезаурус NASA, учебные пособия, написанные специали- стами для специалистов, терминологические стандарты, в частности

действующий «ГОСТ 23851-79 Двигатели газотурбинные авиационные. Термины и определения» и др. Тексты учебных пособий используются как основной материал корпуса, так как являются источником кон- венциональных терминов. Формируемый корпус охватывает явления разных уровней коммуникации, исключая явления профессиональной разговорной речи.

В качестве особенностей предметной области, выявленных при формировании корпуса, необходимо отметить ее многоаспектность: теоретическая база отрасли объединяет знания физики (в особенности таких ее разделов, как термодинамика, газодинамика и теплообмен), математика, конструирование, материаловедение, электроника. Другая особенность, тесно связанная с многоаспектностью данной предметной области, — сложность соотнесения исходного англоязычного подкорпуса с русскоязычным, значительную часть которого составляют тексты учеб- ных пособий, характеризующихся различием подходов к описанию пред- метной области. Если для отечественной школы важна структура объекта исследования, то зарубежная школа придерживается функционального подхода, в связи с чем тексты, отобранные в соответствии с общим тема- тическим маркером (например, на тему конструирования того или иного узла газогенератора), отличаются содержанием, объемом, дробностью подачи сопоставляемого материала. Объем параллельных текстов в силу ряда причин в данной предметной области ограничен и, как правило, имеет обратную последовательность (с русского на английский).

В силу того, что формируемый корпус текстов является преимуще-

ственно сопоставимым, извлечение терминологических данных из под- корпусов и терминологическое выравнивание опираются на выявление характерных для обоих подкорпусов однословных терминологических единиц, их сопоставление в качестве кандидатов в переводные экви- валенты, поиск устойчивых словосочетаний с этими однословными терминами в качестве ядер и последующую верификацию у экспертов в предметной области [Беляева 2014]. Автоматизация на данном этапе в процедурах создания ТБД и в лексикографической практике использу- ется при извлечении терминов. Для создания отраслевой ЛБД предметной области ‘авиационное двигателестроение’ в качестве инструмента работы с данными корпуса используется пакет программ WordSmith Tools. При помощи входящих в пакет инструментов Wordlist и Keywords в подкорпу- сах вычисляются ключевые слова. Далее к избранным ключевым словам с помощью инструмента WordSmith Concord формируется конкорданс. Коллокации, повторяющиеся в конкордансе, при условии наличия у них признаков термина, сравниваются с заранее сформированным на основе бумажных словарей списком используемых в данной предметной обла- сти терминов. Не найденные в списке терминов коллокации становятся

кандидатами на включение в ЛБД и подлежат верификации у экспертов в предметной области.

В качестве промежуточного итога работы с корпусом текстов необхо- димо отметить, что использование методов создания системы управления операциями с терминологией при формировании отраслевой ЛБД позво- ляет, в том числе, более полно соотнести логико-понятийные структуры предметной области, выявить их отличительные и общие особенности.

**Литература**

1. *Беляева Л.Н., Джепа Т.Л.* (2012), Автоматизированное рабочее место пере- водчика: лингвистические ресурсы и технологии. Структурная и прикладная лингвистика. Вып. 9: межвуз. сб. / под ред. А.С. Герда. СПб., с. 109–128.
2. *Беляева Л.Н.* (2014), Системы и процедуры выделения терминов из текста. Терминология и знание. Выпуск IV. Материалы международного симпозиума. М., с. 212–223.
3. *ISO 26162:2012* (2012), Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems, available at: [https://www](http://www.iso.org/obp/ui/#iso%3Astd%3Aiso%3A26162%3Aed-1%3Av1%3Aen).iso.or[g/obp/ui/#iso:std:iso:26162:ed-1:v1:en.](http://www.iso.org/obp/ui/#iso%3Astd%3Aiso%3A26162%3Aed-1%3Av1%3Aen) (дата обращения 02.09.2016).
4. *Kudashev I.* (2013), Quality assurance in terminology management: Recom- mendations from the TermFactory project. Helsinki: Unigraﬁa, 2013.
5. *Towards Consolidation of European Terminology Resources* (2006), Experience Recommendations from Euro TermBank Project. Riga: Tilde.

**Literature**

1. *Belyaeva L.N., Dzhepa T.L.* (2012), Avtomatizirovannoe rabocheye mesto perevodchika: lingvisticheskie resursy i tekhnologii. Strukturnaya i prikladnaya ling- vistika. Vyp. 9: mezhvuz. sb. / pod red. A.S. Gerda. SPb., s. 109–128.
2. *Belyaeva L.N*. (2014), Sistemy i protsedury vydeleniya terminov iz teksta. Terminologiya i znanie. Vypusk IV. Materialy mezhdunarodnogo simpoziuma. M., s. 212–223.
3. *ISO 26162:2012* (2012), Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems, available at: [https://www](http://www.iso.org/obp/ui/#iso%3Astd%3Aiso%3A26162%3Aed-1%3Av1%3Aen).iso.or[g/obp/ui/#iso:std:iso:26162:ed-1:v1:en.](http://www.iso.org/obp/ui/#iso%3Astd%3Aiso%3A26162%3Aed-1%3Av1%3Aen) (data obrashcheniya 02.09.2016).
4. *Kudashev I.* (2013), Quality assurance in terminology management: Recom- mendations from the TermFactory project. Helsinki: Unigraﬁa, 2013.
5. *Towards Consolidation of European Terminology Resources* (2006), Experience Recommendations from Euro TermBank Project. Riga: Tilde.

***М.А. Зорина***

**КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ АНГЛОЯЗЫЧНОЙ ТЕРМИНОСИСТЕМЫ КОРПОРАТИВНОГО ПРАВА**

**Аннотация.** Статья посвящена моделированию логико-понятийных схем терминосистемы корпоративного права стран англо-саксонской правовой семьи. В исследовании раскрываются основные проблемы изучаемой предметной обла- сти, такие как многообразие правовых терминосистем, лексико-семантическая историко-территориальная вариантность англоязычных юридических терминов, а также факторы ее возникновения. По результатам анализа автор приходит к выводу, что когнитивное моделирование позволяет комплексно представить сферу корпоративного права, выделить входящие в него разделы, отрасли и под- отрасли, добиться наглядного представления изучаемой терминосистемы, опреде- лить место терминов в ней, а также выявить логико-понятийные связи между понятиями. В свою очередь, все это способствует решению проблем перевода и вариантности терминов.

**Ключевые слова.** Правовая терминосистема, вариантность терминов, логико- понятийная схема, корпоративное право, правовая семья.

***M.A. Zorina***

**COGNITIVE MODEL BUILDING OF ENGLISH TERMINOLOGICAL SYSTEM OF CORPORATE LAW**

**Abstract.** The article is devoted to the model building of a terminological system of the corporate law in common law countries. The research reveals the main problems of the subject area such as a diversity of the terminological systems, lexicological, semantic, historical and territorial variance of English legal terms, as well as the factors for its formation. As a result of the analysis the author comes to the conclusion that the meaning of a term and its translation depends on the concrete cultural and legal terminological system.

**Keywords.** Legal terminological system, variance of terms, logical-and-conceptual scheme, corporate law, legal family.

Степень обращения современных исследователей к методу модели- рования сравнительно возросла. В сфере лингвистики моделирование рассматривается в качестве эффективного метода исследования языковой реальности и состоит в конструировании моделей для целей отражения определенных характеристик моделируемых объектов. В общем смысле моделирование лежит в основе любого метода научного исследования — как теоретического, при котором используются различные знаковые,

© ЗОРИНА Марина Анатольевна. — Zorina Marina. Институт гуманитарного образования и информационных технологий, Россия. — Institute for the Humanities and Information Technologies, Russia. E-mail: [m1810@mail.ru](mailto:m1810@mail.ru)

абстрактные модели, так и экспериментального, использующего пред- метные модели [БЭС 2003: 834].

Моделирование представляет собой «метод исследования объектов действительности на их моделях; построение и изучение моделей реально существующих предметов и явлений и конструируемых объектов для определения, либо улучшения их характеристик, рационализации спо- собов их построения, управления ими и т. п.» [ФЭС: 381]; совокупность способов идеализации и абстрагирования, анализа и синтеза с целью мыслительного и формализованного представления объектов (оригиналов) исследования и изучения на основании этого представления соответствую- щих явлений, признаков, процессов, связей.

Модель обычно рассматривается как аналог (схема, образец, структура, система), воспроизводящий часть действительности в упрощенном виде (материальном или ментальном), сохраняющий наиболее существенные свойства объекта-оригинала, конструируемый для цели его исследования. Это «любая система, которая имеет мысленное представление, матери- ально реализуется и обладает способностью замещать объект исследова- ния в такой мере, чтобы его изучение давало новую информацию об этом объекте» [Алефиренко 2005: 14–15].

Метод моделирования получил широкое распространение в линг- вистике. Языковая модель воспроизводит явление языка в несколько упрощенном, схематизированном виде и представляет собой искус- ственно создаваемое реальное или мысленное устройство. В лингвистике моделирование используется для определения и исследования структур, на которых основываются языковые явления. Этот процесс включает в себя формализацию накопленных эмпирических знаний о языке и его закономерностях и способствует уточнению соответствующих лингвисти- ческих понятий и связей между ними [Кравцова 2014: 181–189].

При построении языковой модели используются средства и методы математической лингвистики. В любой модели фиксируются: объекты, соответствующие данным непосредственного наблюдения (множество звуков, слов, предложений); объекты, конструируемые исследователем для описания (заранее заданные, строго ограниченные наборы категорий, признаков, элементарных смысловых структур и т. п.) [Кравцова 2014: 181–189].

В процессе моделирования принято выделять следующие этапы:

1. постановка проблемы и ее качественный анализ;
2. конструирование модели объекта;
3. исследование и анализ модели;
4. подведение предварительных итогов исследования модели;
5. практическая проверка полученных с помощью модели знаний [Кутугина 2005: 80].

Лингвистическое моделирование имеет свои специфические черты, среди которых А.Ф. Лосев выделяет принцип структуры объекта, саму структуру объекта, структуру, данную в деталях, перенесение структуры на новый объект и соответствующая организация этого объекта [Лосев 1968].

Благодаря моделированию логико-понятийной системы становится возможным лексикографическое описание подъязыков той или иной предметной области, осуществление научного подхода к изучению специ- альных текстов, а также разработку моделей словарей различных типов. Логико-понятийная система является фундаментом любого терминоло- гического словаря, что на данный момент обосновано в ряде диссертаци- онных работ по языкам для специальных целей различных предметных областей, защищенных представителями Ивановской лексикографической школы [Кувшинова 2007].

Существует ряд методов и способов организации понятийных систем специальностей. Терминологические системы могут быть представлены в виде:

* + рубрикатора с последовательным выделением выше- и нижестоящих понятий;
  + тезауруса (составление дискриптивного словаря — тезауруса);
  + «дерева»;
  + «скобочной записи», сгруппированные согласно частным родствен- ным отношениям (видовым, подклассовым);
  + «спутниковой системы» («Satellite system»);
  + «последовательной» терминологической системы («Sequential concept system»), не имеющей определенной графической обработки;
  + фасетов [Кувшинова 2007].

Сфера корпоративного права сложна и многогранна, имеет развет- вленную структуру. В связи с рядом факторов представление данной междисциплинарной области в комплексном и систематизирующем виде достаточно затруднено. Сложность моделирования корпоративного права стран англо-саксонской правовой семьи заключается в отсутствии деления на отрасли, а также иного структурирования отдельных сфер знания, невыраженности разделения права на публичное и частное, казу- альный характер права и его несистематизированность, существенные отличия от романо-германской правовой семьи с ее нормотворческим характером. Внутри самой терминосистемы корпоративного права стран англо-саксонской правовой семьи также присутствуют различия, формирующиеся под влиянием лингвистических и экстралингвисти- ческих факторов. К последним мы относим исторические, правовые, культурные, территориальные и когнитивные факторы. Все это приво- дит к многообразию правовых терминосистем, лексико-семантической

историко-территориальной вариантности англоязычных юридических терминов.

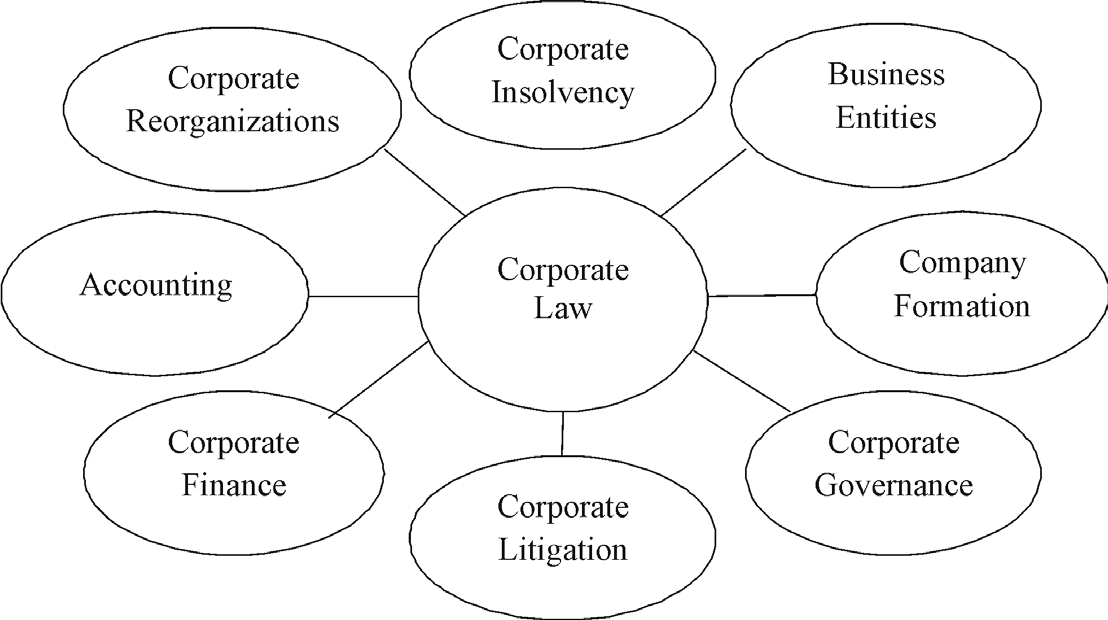
Вариантность терминов объясняется принадлежностью этих единиц разным национально-правовым системам и, следовательно, разным право- вым терминосистемам. Значение термина и его перевод зависит от того, в контексте какой культурно-правовой терминосистемы он употребляется. Сложности, возникающие в процессе перевода, связаны в первую оче- редь с существенными различиями в правовых системах нашей страны и стран, официальным языком которых является английский язык. Ком- плексное моделирование терминосистем позволяет выявить связи между понятиями, их положения относительно других единиц, что позволяет подобрать лучший эквивалентный термин при переводе и способствует решению проблем вариантности терминов.

Накопленный опыт прикладного применения логико-понятийных схем к моделированию структуры, а также солидная теоретическая база, позволяют создать на этих же принципах логико-понятийную систему терминосистемы корпоративного права, которую целесообразно предста- вить в виде схемы, предложенной нами на *рис.* ниже, сконструированной по методу, предложенному учеными западной школы лексикографии, а именно А. Нуоппонен. Графическое изображение схемы носит название

«спутниковой» модели вследствие того, что каждый ее элемент представ- ляет собой своеобразный «спутник», способный как присоединять к себе другие элементы, так и быть присоединенным. Данный метод позволяет иллюстрировать в одной схеме многие графические обозначения взаимо- зависимостей и отношений разных элементов друг к другу [Nuopponen 1997: 364].

Все понятия внутри схемы находятся в отношениях парадигматиче- ского и системного родства. Центральный узел системы — “Corporate Law” — обозначает основное понятие исследования. За каждым вспомо- гательным «спутником» стоит отдельный раздел корпоративного права. Внутренняя логико-понятийная схема корпоративного права состоит из восьми «спутников», каждый из которых соединен лишь с одним, цен- тральным узлом системы [Кувшинова 2007].

Важной задачей моделирования в лингвистике является любое упо- рядочение терминологии, понятийная систематизация в целях осмысления связи теории с ее терминологией [Кантышева 2011]. Логико-понятийная система позволяет выделить основные направления и установившиеся связи, существующие внутри терминологической системы. Моделиро- вание логико-понятийных схем предметной области является одним из наиболее важных этапов в изучении терминологии, а также неотделимой частью работы над терминологическим словарем.



*Рис.* Логико-понятийная система корпоративного права в виде «спутниковой модели»

Проведенное исследование позволяет нам сделать вывод об эффектив- ности когнитивного моделирования терминосистем, которое позволяет представить сферу корпоративного права в комплексном виде, выделить входящие в него разделы, добиться наглядного представления изучаемой терминосистемы, определить место терминов в ней, а также выявить логико-понятийные связи между понятиями.

**Литература**

1. *Алефиренко Н.Ф.* ( 2005), Современные проблемы науки о языке. М.
2. *Большой энциклопедический словарь* [БЭС] (2004) / под ред. А. М. Про- хорова. 2-е изд., перераб., доп. М.-СПб.
3. *Кантышева Н.Г.* (2011), Моделирование терминосистемы «экологический аудит» и ее описание в систематизирующем глоссарии: дис. канд. филол. наук:
   * 1. Тюмень.
4. *Кравцова Ю.В.* (2014), Моделирование в современной лингвистике / Вестник Житомирского государственного университета им. Ивана Франка. № 77.
5. *Кувшинова А.В.* (2007), Моделирование терминосистемы предметной области «текстильное дело»: <http://www.rusnauka.com/8._NPE_2007/Philolo-> gia/20939.doc.htm (дата обращения: 06.05.2016).
6. *Кутугина Е.С.* (2005), Моделирование. Томск : Изд-во ТГУ.
7. *Лосев А.Ф.* (1968), Введение в общую теорию языковых моделей. М.: Наука.
8. Философский энциклопедический словарь [ФЭС] (2003) / Губский Е.Ф., Кораблева Г.В., Лутченко В.А. М.
9. *Nuopponen A.A.* (1997), Model for Systematic Terminological Analysis // Pro- ceedings of 11th European Symposium on LSP. Copenhagen. Vol. 1.

**Literature**

* + - 1. *Aleﬁrenko N.F*. ( 2005), Sovremennye problemy nauki o yazyke. M.
      2. Bol'shoy entsiklopedicheskiy slovar' [BES] (2004) / pod red. A. M. Prokhorova. 2-e izd., pererab., dop. M.-SPb.
      3. *Kantysheva N.G.* (2011), Modelirovanie terminosistemy «ekologicheskiy audit» i eyo opisanie v sistematiziruyushchem glossarii: dis. kand. ﬁlol. nauk: 10.02.21.

Tyumen’.

* + - 1. *Kravtsova Yu.V*. (2014), Modelirovanie v sovremennoy lingvistike / Vestnik Zhitomirskogo gosudarstvennogo universiteta im. Ivana Franka. № 77.
      2. *Kuvshinova A.V*. (2007), Modelirovanie terminosistemy predmetnoy oblasti

«tekstil’noe delo»: <http://www.rusnauka.com/8._NPE_2007/Philologia/20939.doc.htm> (data obrashcheniya: 06.05.2016).

* + - 1. *Kutugina E.S.* (2005), Modelirovanie. Tomsk : Izd-vo TGU.
      2. *Losev A.F*. (1968), Vvedenie v obshchuyu teoriyu yazykovykh modeley. M.: Nauka.
      3. Filosofskiy entsiklopedicheskiy slovar’ [FES] (2003) / Gubskiy E.F., Korableva G.V., Lutchenko V.A. M.
      4. *Nuopponen A.A.* (1997), Model for Systematic Terminological Analysis // Pro- ceedings of 11th European Symposium on LSP. Copenhagen. Vol. 1.

***С.С. Калинин***

**ТЕОРИЯ ПЕРЕВОДА И ЭКОЛИНГВИСТИКА: ПРОЦЕСС ПЕРЕВОДЧЕСКОЙ ДЕЯТЕЛЬНОСТИ КАК ПРОЦЕСС АДАПТАЦИИ СИСТЕМЫ**

**К МЕНЯЮЩИМСЯ ФАКТОРАМ СРЕДЫ**

**Аннотация.** На основе биологической теории познания и эколого- аутопоэзного учения о языке У. Матураны дается анализ процесса переводческой деятельности. Описывается механизм формирования единиц перевода (транс- латем). Показывается, что процесс формирования текста на переводящем языке является по своей сути аутопоэзным. Показывается роль окружающей среды, ее факторов и субстрата в осуществлении живой системой переводческой деятель- ности. Дается общий набросок модели взаимодействия переводящей системы, языковой среды и окружающей среды. Делаются некоторые выводы об особен- ностях функционирования живых и искусственных переводящих систем.

**Ключевые слова.** Биология познания, теория перевода, переводческая дея- тельность, эколингвистика, живая система, языковая среда, окружающая среда, факторы среды, аутопоэзис, наблюдатель.

***S.S. Kalinin***

**TRANSLATION THEORY AND ECOLINGUISTICS:**

**THE TRANSLATION PRACTICE PROCESS AS ADAPTATION OF THE SYSTEM TO THE VARIABLE**

**ENVIRONMENT FACTORS**

**Abstract.** The analysis of the translation practice process is given on the basis of H. Maturana’s biologic theory of cognition and his ecologic-autopoiesis linguistic theory. The ways of the translation units formation are described. It is said that the translating language text formation process is autopoietic. The role of the environment, its factors and the understratum in the translation practice of living system is shown. The draft of the common interaction model of the translating system, verbal environment and environment is given. The main conclusion of this work is about the peculiarities of the function of living and artiﬁcial translating systems.

**Keywords.** Biology of cognition, translation theory, translation practice, ecolinguis- tics, living system, verbal environment, environment, environment factors, autopoiesis, observer.

Биологическая теория когниции У. Матураны дает резко отличное от традиционного понимание языковых процессов, в том числе процесса язы- кового познания. Одним из основных положений данной теории является положение о том, что когда наблюдатель (система, занимающаяся процес-

© КАЛИНИН Степан Сергеевич. — Kalinin Stepan. Кемеровский государственный универ- ситет, Россия. — Kemerovo State University, Russia. E-mail: [rage\_of\_gods@inbox.ru](mailto:rage_of_gods@inbox.ru)

сом языковой когниции мира)1 взаимодействует в своей консенсуальной области с другим наблюдателем или каким-либо объектом в потоке жиз- недеятельности, то, соответственно, наблюдатель описывает эти объекты в языке [Maturana 1995: 155]. Осуществляя координацию координаций взаимодействия себя самого как системы с другой системой, наблюда- тель порождает языковые структуры, которые являются результатами координационной деятельности [Maturana 2008: 19–20]. Все языковые объекты и понятия возникают исключительно как результат координации координаций деятельности системы и существуют только как таковые. Представляется очевидным, что единицы переводческой деятельности (т. н. транслатемы) существуют тоже только как результаты координации координаций взаимодействия системы со внешней средой. В случае про- цесса перевода есть одна особенность: система (можно считать таковой живую систему, например, человека, осуществляющего переводческую деятельность) осуществляет процесс как бы двойной координации, одновременно координируя и со средой исходного языка (далее — ИЯ), и со средой переводящего языка (далее — ПЯ). В этом принципиальное отличие переводческой деятельности системы от ее обычной языковой деятельности (languaging, по У. Матуране). При этом возникают две обла- сти консенсуального поведения системы [Кашкин 2011: 119] — область, относящаяся к ИЯ, и область, относящаяся к ПЯ. Сначала в области ИЯ система, координируя взаимодействие между собой и языковой сре- дой, порождает определенный набор языковых единиц каждого уровня (и соответственно определенных транслатем), затем, переходя в область ПЯ, система координирует форму и содержание этих единиц со средой ПЯ, в результате она воспроизводит новый набор транслатем и соответ- ствующих им языковых единиц. Таким образом система осуществляет переводческую деятельность.

Можно утверждать, что процесс переводческой деятельности, осу-

ществляемый системой, является в некотором отношении аутопоэзным процессом, поскольку среда как ИЯ, так и ПЯ, воспроизводит себя на каждом из языковых уровней, при этом система (в данном случае неважно, живая ли это система или искусственная) является медиатором в этом процессе. Система, осуществляющая переводческую деятель- ность, координирует взаимодействие между собственными структурами и языковой средой, языковым материалом. В том, что языковая среда ИЯ воспроизводит сама себя средствами ПЯ на различных языковых уровнях, проявляется фрактальность языка, его свойство носителя энергейи в гум- больдтианском смысле.

1 В принципе, У. Матурана рассматривает наблюдателя как живую систему, живой орга- низм. Мы же пойдем дальше и будем рассматривать здесь наблюдателя и как искусствен- ную систему (компьютер, ИИ, в перспективе его создания).

В. Н. Комиссаров упоминает о том, что единицей перевода можно считать «текст плюс ситуацию» [Комиссаров 2009: 142], однако он не рас- крывает само понятие «ситуации» в данном контексте и то, что за ним стоит. Нам же важно прояснить само понятие «ситуация» в контексте аутопоэзной и экологической парадигмы исследования языка, предло- женной У. Матураной. Необходимо также прояснить взаимоотношение между системой, средой, в которой находится система, и языковой средой, языковым миром. Здесь и в дальнейшем для простоты изложения будем считать переводящую систему живой (человеком).

С определенной долей уверенности можно сказать о том, что тер- мин «ситуация», который использует В. Н. Комиссаров, подразумевает в эколого-аутопоэзной интерпретации окружающую среду (в которой как раз находится система) и субстрат, лежащий за средой. Окружающая среда состоит, по П. Митчеллу [Митчелл 2001: 99], из четырех компонентов: местообитания, других организмов (в нашей интерпретации, наблюдате- лей), ресурсов и условий. Среда может содержать также и культурный, и религиозный компоненты, если речь идет о среде обитания человека [Митчелл 2001: 100]. В философском смысле среда содержит в себе все вещи в явленной, феноменальной форме. Все вещи в среде даны нам в ощущениях. Субстрат же, в философской интерпретации, непроявлен, его нельзя ощутить непосредственно, но за счет него все вещи как когни- тивные сущности привносятся в феноменальный мир [Имото 2009: 91]. В этом плане интересно высказывание Дж. Спенсера-Брауна о природе вещей и об их происхождении: «Мысль о том, что сотворение должно быть следствием «чего-то», глупа. Ни одна вещь не может иметь каких бы то ни было следствий. Если бы изначально было нечто, оно бы отравило весь процесс творения. Лишь ничто достаточно нестабильно, чтобы породить бесконечные цепочки разных проявлений» [Spencer-Brown 1994: 9]. Инте- ресно сопоставить эту мысль с высказыванием У. Матураны о том, что

«в субстрате есть ничто (не-вещь), потому что вещи принадлежат языку.

В субстрате ничего не существует» [Maturana 1992: 108]. Таким образом, субстрат как бы неявно определяет условия как окружающей среды, так и языковой среды (а в каждой из этих сред находится система), субстрат пронизывает среду, но сам при этом невидим.

Возвращаясь же к определению окружающей среды как структуры, состоящей из четырех компонентов, следует более подробно остановиться на ресурсах среды и условиях среды. Очевидно, что ресурсом для пере- водческой деятельности служит язык и стоящие за ним культурные и кон- цептуальные основы1. Условиями среды же служат различные факторы,

1 Возможно рассматривать концептуальные слои языка как аналог субстрата, поскольку непосредственно концепты недоступны наблюдению, и мы вынуждены исследовать толь- ко их языковые объективации.

повышающие успешность, адекватность переводческой деятельности или же, наоборот, создающие помехи при переводе. Таковыми могут являться, например, помехи при восприятии графических знаков, в случае если осу- ществляется письменный перевод, или помехи при восприятии звуковых знаков, в случае если происходит устная коммуникация. К данным фак- торам можно отнести также языковую и культурную компетенции пере- водчика. Данные факторы детерминируют изменения, адаптации системы к меняющимся условиям среды, что особенно ярко видно на примере помех в каналах чувственного восприятия. Вслед за С. Имото мы будем называть такие факторы триггерами [Имото 2009: 92].

По своему типу большинство факторов и условий окружающей среды относятся к пертурбациям [Имото 2009: 93], поскольку они запускают реакцию системы на изменяющиеся условия окружающей среды, но не приводят к полной ее дезинтеграции (что в случае переводческой деятель- ности вообще невозможно). Эти триггеры определяют то, как ответит система на, образно выражаясь, «очередной вызов» среды. При этом сама система структурно изменяется (например, происходит изменение неко- торых процессов в головном мозге человека), приспосабливается к среде и образует с ней принципиально новое экологическое единство. Конечный результат перевода — текст на ПЯ — также приводит к возникновению нового экологического единства в структуре «переводящая система — окружающая среда — языковая среда». За счет фактора аутопоэзности и принципа накопления и усиления флуктуаций1 [Алифиренко 2009: 236] в тексте ИЯ происходит изменение формы и содержания текста на ПЯ. Так упомянутая выше структура реагирует на изменения условий среды. Таким образом, мы показали с точки зрения биологической теории когниции и аутопоэзно-экологического учения о языке У. Матураны, как организован процесс переводческой деятельности. Мы показали взаи- моотношения переводящей системы с языковой и окружающей средой, показали, что переводящая система образует экологическое единство с данными средами. В дальнейшем можно использовать методы мату- рановской языковой теории для более детального изучения процесса перевода и для изучения переводческой деятельности искусственных

переводящих систем (компьютерные программы, ИИ).

**Литература**

1. *Алефиренко Н.Ф.* (2009), Лингвокогнитивная синергетика: истоки, прин- цип, сущность. Studia linguistica cognitiva. Вып. 2: Наука о языке в изменяющейся парадигме знания. Иркутск: Изд-во БГУЭП, с. 228–256.

1 Он заключается в том, что накопление значительного количества малых изменений в си- стеме приводит, в конечном итоге, к крупным изменениям в самой системе.

1. *Имото С.* (2009), Ничто как наполненность: «ПУТЬ» Лао-Цзы и

«СУБСТРАТ» Матураны / пер. с англ. А. В. Кравченко. Studia linguistica cognitiva. Вып. 2: Наука о языке в изменяющейся парадигме знания. Иркутск: Изд-во БГУЭП, с. 87–100.

1. *Кашкин В.Б.* (2011), Языковые контрасты в истории лингвистики. 2-е изд., перераб. и доп. Воронеж: Наука-Юнипресс. 242 с.
2. *Комиссаров В.Н.* (2009), Лингвистика перевода. 3-е изд. М.: Книжный дом

«ЛИБРОКОМ». 176 с.

1. *Митчелл П.* (2001), 101 ключевая идея: Экология / пер. с англ. О. Перфи- льева. М.: ФАИР-ПРЕСС. 224 с.
2. *Maturana H.R.* (2008), Anticipation and self-consciousness: Are these functions of the brain? Constructivist Foundations, No 4 (1), pp. 18–20.
3. *Maturana H.R.* (1995), Biology of self-consciousness. Consciousness: Distinc- tion and reﬂection. Naples, Italy: Bibliopolis, pp. 145–175.
4. *Maturana H.R.* (1992), The biological foundations of self-consciousness and the physical domain of existence. Beobachter: Konvergenz der Erkenntnistheorien? 2nd ed. Munich, Germany: Wilhelm Fink Verlag, S. 47–117.
5. *Spencer-Brown G.* (1994), Laws of Form. Portland: Cognizer Co., 1994. 141 p.

**Literature**

1. *Alefirenko N.F.* (2009), Lingvokognitivnaya sinergetika: istoki, printsip, sushchnost'. Studia linguistica cognitiva. Vyp. 2: Nauka o yazyke v izmenyayushcheysya paradigme znaniya. Irkutsk: Izd-vo BGUEP, s. 228–256.
2. *Imoto S*. (2009), Nichto kak napolnennost’: «PUT’’» Lao-Tszy i «SUBSTRAT» Maturany / per. s angl. A. V. Kravchenko. Studia linguistica cognitiva. Vyp. 2: Nauka o yazyke v izmenyayushcheysya paradigme znaniya. Irkutsk: Izd-vo BGUEP, s. 87–100.
3. *Kashkin V.B.* (2011), Yazykovye kontrasty v istorii lingvistiki. 2-e izd., pererab. i dop. Voronezh: Nauka-Yunipress. 242 s.
4. *Komissarov V.N.* (2009), Lingvistika perevoda. 3-e izd. M.: Knizhnyi dom

«LIBROKOM». 176 s.

1. *Mitchell P.* (2001), 101 klyuchevaya ideya: Ekologiya / per. s angl. O. Perﬁl’yeva. M.: FAIR-PRESS. 224 s.
2. *Maturana H.R.* (2008), Anticipation and self-consciousness: Are these functions of the brain? Constructivist Foundations, No 4 (1), pp. 18–20.
3. *Maturana H.R.* (1995), Biology of self-consciousness. Consciousness: Distinc- tion and reﬂection. Naples, Italy: Bibliopolis, pp. 145–175.
4. *Maturana H.R.* (1992), The biological foundations of self-consciousness and the physical domain of existence. Beobachter: Konvergenz der Erkenntnistheorien? 2nd ed. Munich, Germany: Wilhelm Fink Verlag, S. 47–117.
5. *Spencer-Brown G.* (1994), Laws of Form. Portland: Cognizer Co., 1994. 141 p.

***М.А. Коряковцев***

**ОСОБЕННОСТИ ПОСТРОЕНИЯ ОНТОЛОГИИ ЭЛЕКТРОННОГО РЕЗЮМЕ ПРИ РАЗРАБОТКЕ КАДРОВОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ**

**Аннотация.** В работе описывается процесс построения и особенности онтологии электронного резюме. Указывается важность построения онтологии в качестве предварительного этапа при разработке кадровой информационной системы с использованием федерализации данных. Важность предварительного построения онтологии заключается в устранении структурных различий, которые присущи электронным резюме из разных источников — сайтов трудоустройства. В качестве особенностей в работе обозначается использование элементов мета- данных процессора федерализации как объектов онтологии: классов, атрибутов и отношений и преобладание относящихся к типу включения отношений между классами онтологии. В заключении, в качестве результата, представляется часть онтологии, построенной на основе бланков электронных резюме сайтов трудо- устройства «HeadHunter», «SuperJob» и «Работа.Ru».

**Ключевые слова.** Кадровая информационная система, онтология, электрон- ные резюме, интеграция данных.

***M.A. Koryakovtsev***

**CHARACTERISTICS OF THE ONTOLOGY**

**OF DIGITAL RESUME FOR THE HUMAN RESOURCE INFORMATION SYSTEM**

**Abstract.** The process of the creation and characteristics of the ontology of digital resume is described in this article. The matter of the ontology creation is pointed out as a preparation step for human resource information system developing. The human resource information system follows the federalization approach for data integration. Disposing of the structural differences of digital resumes from different sources is the main goal at the preparation step. The metadata of federalization processor is used as ontology’s elements. The “part-of” relation between classes is dominant in created the ontology. The part of ontology of digital resume for sites “HeadHunter”, “SuperJob” and “Rabota.Ru” is presented in conclusion.

**Keywords.** Human resource information system, ontology, digital resume, data integration.

Кадровые информационные системы (КИС) относятся к экономиче- ским автоматизированным информационным системам. Автоматизиро- ванная информационная система является человеко-машинной системой, которая обеспечивает сбор и обработку информации, необходимой для управления деятельностью [Горделий 2005: 23]. Основной задачей раз-

© КОРЯКОВЦЕВ Михаил Андреевич. — Koryakovtsev Mikhail. Тюменский государствен- ный университет, Россия. — Tyumen State University. E-mail: [simpfr@mail.ru](mailto:simpfr@mail.ru)

рабатываемой кадровой информационной системы является универсали- зация подбора персонала по электронным резюме, размещаемых на сайтах трудоустройства, таких как «HeadHunter», «SuperJob» и «Работа.Ru» и других. Универсализация заключается в использовании единого поис- кового интерфейса для конечного числа сайтов трудоустройства.

Информационные системы, включая разрабатываемую КИС, состоят из следующих подсистем: информационное обеспечение, техническое обеспечение, программно-математическое обеспечение и организационно- правовое обеспечение.

Остановимся подробнее на подсистеме «информационное обеспече- ние». Информационное обеспечение — «это совокупность баз данных, фай- лов, документов и лингвистических средств, обеспечивающих реализацию информационной составляющей АИС» [Исаев 2013: 35]; «совокупность единой системы информации, унифицированных систем документации и массивов информации, используемых в АИС» [Горделий 2005: 25].

При проектировании описываемой кадровой поисковой информацион- ной системы необходимо в первую очередь решить проблему интеграции данных резюме из различных источников — сайтов трудоустройства.

Таким образом, целью данной статьи является выбор и реализация подхода к интеграции банков резюме трех сайтов трудоустройства «Head- Hunter», «SuperJob» и «Работа.Ru».

Материалом для исследования послужили 14 электронных бланков. Эти электронные бланки используются при создании электронных резюме на перечисленных сайтах трудоустройства и могут быть использованы для восстановления полной структуры электронного резюме конкретного сайта трудоустройства [Коряковцев 2016].

Сайты трудоустройства как источники данных обладают следующими 7 свойствами:

1. отдельная веб-страница сайта трудоустройства с данными резюме является единицей поиска;
2. данные резюме на веб-странице являются частично формализо- ванными;
3. данные резюме не содержат персональной информации (ФИО соис- кателя, контактные данные);
4. адреса веб-страниц электронных резюме можно получить только при использовании предназначенных для человека средств поиска;
5. опубликованные резюме могут быть изменены соискателем;
6. количество резюме на сайтах трудоустройства очень велико (порядка миллионов);
7. количество адресов электронных резюме в поисковой выдаче сред- ствами сайта трудоустройства ничтожно мало по сравнению с общим количеством резюме.

Для интеграции данных в информационных системах обычно при- меняются два фундаментальных подхода — консолидация и федерализа- ция, а также три производных: их комбинация, распространение данных и сервисный подход [Тузовский 2011; Когаловский 2003; Белошицкий 2013]. С учетом приведенных выше характеристик сайтов трудоустройства в качестве источников данных для информационно-поисковой системы была выбрана федерализация.

Федерализация предполагает как трансляцию (преобразование) поискового запроса от пользователя разрабатываемой ИПС к поисковой системе сайта трудоустройства, так и преобразование поисковой выдачи к универсальной, общей структуре. Следовательно, использование феде- ративного подхода к интеграции источников резюме может обеспечить построение некоторой базы знаний. Данная база знаний должна содержать информацию об отображении (“mapping”) структурных элементов резюме сайта трудоустройства на структурные элементы некоторого универсаль- ного резюме, используемого КИС.

В качестве структурных элементов резюме используются рубрики. Под рубриками мы понимаем смысловые элементы электронного резюме, которые пользователь сайта трудоустройства заполняет в спе- циальном электронном бланке при публикации резюме на сайте трудо- устройства. Названия рубрик в используемой структуре данных имеют сложные, составные наименования: «Опыт работы. Место работы. Начало работы. Месяц», «Опыт работы. Место работы. Обязанности, функции, достижения», «Знание языков. Иностранный язык. Уровень владения» и т. д. Это связано с тем, что в электронных резюме рубрики образуют иерархию «рубрика-подрубрика». Однако на разных сайтах трудоустройства используется своя иерархия, отличающаяся как по отношениям между рубриками, так и по наличию отдельных рубрик. Следовательно, использование составного наименования рубрик позволяет упразднить иерархию рубрик, при сохранении уникальности названия рубрики и достаточности названия для описания ее смысло- вого содержания.

Для описания базы знаний для федеративного подхода к интеграции

мы построили специальную онтологию. Основными элементами любой онтологии, ее «строительными блоками», являются *класс, подкласс, экзем- пляр и атрибут класса, отношение* [Верхотурова 2013]. Для описания компонентов разработанной нами онтологии мы использовали только понятия класса, атрибута, отношения и экземпляра. При создании онто- логии нами было выделено 5 различных классов и 7 атрибутов с соответ- ствующими отношениями. Классы, атрибуты и отношения разработанной онтологии представлены в табл. 1.

*Таблица 1.* Элементы онтологии базы знаний для интеграции резюме с различных сайтов трудоустройства

|  |  |  |
| --- | --- | --- |
| **Класс** | **Отношение** | **Атрибут** |
| Сайт трудоустройства | Сайт трудоустройства *имеет* на- звание | Название |
| Электронные резюме сайта трудоустройства *состоят из* рубрик | Список рубрик электронного резюме |
| Универсальное резюме | Универсальное резюме *состоит из* рубрик | Список рубрик универсаль- ного резюме |
| Рубрика | Рубрика *имеет* наименования | Наименование |
| Рубрика *имеет* план содержания | План содержания |
| Отображение | Отображение *преобразует*  рубрику электронного резюме | Рубрика электронного резюме сайта трудоустройства |
| Отображение *заполняет* совокуп- ность рубрик универсального резюме | Рубрики универсального резюме |
| План содержания |  |  |

Полученная в итоге онтология состоит из 3 экземпляров класса «Сайт тру- доустройства», 194 экземпляра класса «Отображение», 344 экземпляра класса

«рубрика». Из 344 экземпляров класса «Рубрика» 116 относятся к рубрикам универсального резюме, 84 — к рубрикам резюме сайта трудоустройства

«HeadHunter», 66 — к рубрикам «SuperJob» и 78 — к рубрикам «Работа.Ru». Класс «План содержания» представлен 4 экземплярами, которые обозначают способ представления рубрики: «Текст», «Слово», «Выбор» и «Группа».

Полученная онтология описывает базу знаний для интеграции дан- ных с помощью федеративного подхода. Следовательно, основными элементами такой онтологии являются экземпляры класса «Отображе- ния». Таким образом, база знаний представляет собой набор экземпляров класса «Отображения», состоящих из рубрики электронного резюме сайта трудоустройства и набора рубрик универсального резюме. В табл. 2 пред- ставлены примеры экземпляров класса «Отображения»:

*Таблица 2*. Примеры экземпляров класса «Отображения» онтологии базы знаний федеративного подхода

|  |  |  |
| --- | --- | --- |
| **Сайт трудо-**  **устройства** | **Рубрика** | **Рубрики универсального резюме** |
| HeadHunter | Образование. Учебное заве- дение. Учебное заведение | Образование. Место обучения. Название учебного заведения |
| SuperJob | Личная информация. Теле- фон. Номер телефона | Персональная информация. Телефон. Код страны |
| Персональная информация. Телефон. Идентификационный код |
| Персональная информация. Телефон. Номер абонента |
| Работа.Ru | Опыт, навыки. Опыт работы. Последнее место работы.  Обязанности, достижения | Опыт работы. Место работы. Обязанности, функции, достижения |

**Выводы**. При разработке кадровой информационной системы основ- ное внимание отводится проектированию подсистемы информационного обеспечения. Проектирование информационного обеспечения кадровой информационной системы включает в себя проектирование универ- сального структурного представления резюме и механизмов интеграции с сайтами трудоустройства, которые являются основным источником электронных резюме. Тем не менее интеграция с сайтами трудоустрой- ства является нетривиальной задачей, которая требует выбора и реализа- ции конкретного подхода интеграции. В силу такой особенности сайтов трудоустройства в качестве источника данных, как огромное количество и различная структура электронных резюме, был выбран федеративный подход к интеграции. Федеративный подход основан на отображении структурных элементов электронного резюме на структурные элементы универсального резюме в момент выполнения поискового запроса КИС. Для описания отображений была построена специальная онтология. Раз- работанная онтология описывает 194 отображения между 228 рубриками электронных резюме и 116 рубриками универсального резюме.

**Литература**

1. *Белошицкий Д.А.* (2013) , Интеграция данных в информационных системах: [http://sntbul.bmstu.ru/ﬁle/out/602642](http://sntbul.bmstu.ru/%EF%AC%81le/out/602642) (12.06.2016).
2. *Верхотурова Ю. С.* (2013), Модель предметной области на языке описания онтологий / Вестник Бурятского государственного университета, 9/2013, с. 63–68.
3. *Горделий А.В.* (2005), Кадровые информационные системы. М.
4. *Исаев Г.Н.* (6-е изд) (2013), Информационные системы в экономике: учеб- ник для студентов вузов, обучающихся по специальностям «Финансы и кредит»,

«Бухгалт. учет, анализ и аудит». М.

1. *Когаловский М.Р*. (2010), Методы интеграции данных в информационных системах: <http://www.ipr-ras.ru//articles/kogalov10-05.pdf>(18.06.2016).
2. *Коряковцев М.А.* (2016), Электронный бланк резюме как прототип фрей- мовой модели «Электронное резюме». Наука и образование в социокультурном пространстве современного общества. Смоленск, с. 130–134.
3. *Тузовский А.Ф.* (2011), Интеграция информации с использованием техноло- гий Semantic Web. Проблемы информатики. Новосибирск, с. 51–58.

**Literature**

1. *Beloshitskiy D.A*. (2013) , Integratsiya dannykh v informatsionnykh sistemakh: [http://sntbul.bmstu.ru/ﬁle/out/602642](http://sntbul.bmstu.ru/%EF%AC%81le/out/602642) (12.06.2016).
2. *Verkhoturova Yu. S*. (2013), Model' predmetnoy oblasti na yazyke opisaniya ontologiy / Vestnik Buryatskogo gosudarstvennogo universiteta, 9/2013, s. 63–68.
3. *Gordeliy A.V*. (2005), Kadrovye informatsionnye sistemy. M.
4. *Isaev G.N.* (6-e izd) (2013), Informatsionnye sistemy v ekonomike: uchebnik dlya studentov vuzov, obuchayushchikhsya po spetsial'nostyam «Finansy i kredit»,

«Bukhgalt. uchet, analiz i audit». M.

1. *Kogalovskiy M.R.* (2010), Metody integratsii dannykh v informatsionnykh siste- makh:<http://www.ipr-ras.ru//articles/kogalov10-05.pdf> (18.06.2016).
2. *Koryakovtsev M.A*. (2016), Elektronnyi blank rezyume kak prototip freymovoy modeli «Elektronnoe rezyume». Nauka i obrazovanie v sotsiokul’turnom prostranstve sovremennogo obshchestva. Smolensk, s. 130–134.
3. *Tuzovskiy A.F*. (2011), Integratsiya informatsii s ispol’zovaniem tekhnologiy semantic web. Problemy informatiki. Novosibirsk, s. 51–58.

***Петр Михаловский***

**ТРЕХЪЯЗЫЧНЫЙ ТЕРМИНОЛОГИЧЕСКИЙ СЛОВАРЬ ДЛЯ ПЕРЕВОДЧИКОВ — ПОТРЕБНОСТИ И ПЕРСПЕКТИВЫ**

**Аннотация**. В настоящей статье представляется концепция создания русско- польско-английского словаря экономической терминологии. Адресатами словаря являются прежде всего (будущие) переводчики специальных текстов. Рассматри- ваются также потребности по отношению к данному словарю с точки зрения как профессиональной работы переводчика, так и учебного процесса. В статье обсуж- даются возможности создания терминографической серии на базе описываемого лексикографического продукта.

**Ключевые слова**. Русско-польско-английский словарь, отраслевой терми- нологический словарь, экономическая терминология, переводчик, специальный язык.

***Piotr Michałowski***

**TRILINGUAL TERMINOLOGICAL DICTIONARY FOR TRANSLATORS — NEEDS**

**AND PERSPECTIVES**

**Abstract**. The article presents an idea of creating a Russian-Polish-English eco- nomic dictionary. The dictionary is aimed at assisting primarily (future) translators of texts for special purposes. The article also attempts to describe the needs of potential users of the dictionary, from the point of view of both professional work of a translator and the educational process. Furthermore, it discusses the possibilities of creating a terminographical series based on the described lexicographical product.

**Key words**. Russian-Polish-English dictionary, branch terminological dictionary, terminology of economics, translator, LSP (language for special purposes).

В наше время в связи с постоянным развитием международного торгового и экономического сотрудничества, в частности между Польшей и ее восточными соседями, т. е. Россией, Белоруссией, Украиной и др., наблюдается острая нехватка качественных дву- и многоязычных терминологических словарей, которые смогли бы обеспечить современную лексикографическую базу переводчикам специальных текстов в области экономики в Польше и других странах. В настоящей статье представляется концепция русско-польско-английского словаря экономической

© МИХАЛОВСКИЙ Петр. — Michałowski Piotr. Варшавский университет, Польша. — University of Warsaw, Poland. Е-mail: [p.j.michalowski@uw.edu.pl](mailto:p.j.michalowski@uw.edu.pl)

терминологии, адресованного в первую очередь переводчикам1. Согласно классификациям терминологогических словарей В. М. Лейчика [Лейчик 1989] и В. Змарзер [Zmarzer 1991], данный лексикографический продукт представляет собой отраслевой терминологический словарь (ОТС). В словаре этого типа предметом описания является терминология одной области специальных знаний (в данном случае — экономики). В настоящее время работу над созданием такого словаря проводит авторский коллектив под руководством проф. В.В. Дубичинского, сформировавшийся в Лаборатории лексикографических исследований Института русистики Варшавского университета (в состав которого входят А. Глоговская, П. Михаловский, С. Хватов).

В словник словаря входит около 7000 терминов и терминологических словосочетаний, в том числе 2500 заглавных единиц, отобранных по принципам:

* частотности употребления в экономических текстах;
* учебно-методической целесообразности.

Словник словаря представляет собой трехъязычный терминологиче- ский минимум, обучение которому необходимо для овладения экономи- ческими специальностями в высшем учебном заведении.

Макроструктура проектируемого словаря состоит из следующих элементов:

1. словарная часть;
2. алфавитные индексы терминологических единиц с номерами на польском, английском и русском языках;
3. список наименований денежных единиц различных государств;
4. список наиболее употребительных аббревиатур.

Микроструктура словаря состоит из трех языковых частей (русской, польской и английской). В связи с учебным характером словаря, прин- цип лексикографической симметрии реализуется полностью на всех уровнях его микроструктуры [Michałowski 2014: 86]. Все языковые части построены из одинаковых элементов, которые, в свою очередь, имеют одинаковую структуру во всех словарных статьях.

Микроструктура словаря включает в себя следующие компоненты:

1. заглавный термин, выделенный жирным шрифтом;
2. краткая грамматическая характеристика заглавного термина;
3. предикативная дефиниция заглавного термина;

1 Словарь будет также полезен бизнесменам, менеджерам и экономистам, а также сту- дентам и аспирантам переводческих, экономических и филологических специально- стей.

1. терминологическая сочетаемость заглавного термина;
2. словообразовательное гнездо заглавного термина.

Заглавные термины представляют собой однокомпонентные терми- нологические единицы — имена существительные, например: *акция*, *налог*, *цена*. С одной стороны, они являются компонентами (чаще всего центрами) терминополей, с другой стороны, также опорными термино- элементами терминологических гнезд, которые представляются в дальней- шей части словарной статьи. Наличие краткой предикативной дефиниции на трех языках подчеркивает учебный характер словаря и будет полезным для студентов и начинающих переводчиков.

Терминологическая сочетаемость заглавных единиц обозначается знаком **#** и описывается следующим образом:

В позиции 1 приводятся терминологические словосочетания, в состав которых входит заглавный термин с прилагательным или причастием в функции согласованного определения, например: *биржевая сделка*1, *валютная сделка*, *лицензионная сделка*.

В позиции 2 приводятся терминологические словосочетания, в кото- рых заглавная единица выступает с зависимыми существительными в косвенных падежах (с предлогами и без предлогов), например: *договор купли-продажи*, *договор на поставку*, *договор о найме*.

В позиции 3 приводятся терминологические словосочетания, в которых заглавный термин является зависимым членом и высту- пает в косвенных падежах (с предлогами и без предлогов), например: *концепция маркетинга*, *управление маркетингом*, *управляющий по маркетингу*.

В позиции 4 приводятся терминологические словосочетания заглав- ного термина с управляющими им глаголами (с предлогами и без пред- логов), например: *взыскивать убытки*, *возмещать убытки*, *терпеть убытки*.

Следующим элементом микроструктуры словаря является слово- образовательное гнездо, которое обозначается знаком **\*** и включает в себя однокоренные терминологические единицы, производные от заглавного термина. В словаре приводятся только те компоненты словообразовательного гнезда, которые активно упротребляются в специальных текстах или необходимы с методической точки зрения. Например: *цена* — *ценник*, *оценивать*, *ценные бумаги*, *ценовая поли- тика*, *ценностные ориентиры*.

1 Во всех примерах заглавные термины подчеркиваются.

*Таблица*. Пример словарной статьи предлагаемого словаря

|  |  |  |
| --- | --- | --- |
| **НАЛОГ** *м*  *обязательный, индиви- дуально безвозмездный платеж, принудитель- но взимаемый органами государственной власти различных уровней с орга- низаций и физических лиц*  #  акцизный налог земельный налог косвенный налог подоходный налог прогрессивный налог прямой налог | **PODATEK** *m*  *obowiązkowe świadczenie pie- niężne pobierane przez zwią- zek publicznoprawny (pań- stwo, jednostka samorządu te- rytorialnego) bez konkretnego, bezpośredniego świadczenia wzajemnego*  #  podatek akcyzowy / akcyza podatek gruntowy  podatek pośredni podatek dochodowy podatek progresywny podatek bezpośredni | **TAX / DUTY**  *a financial charge or other levy imposed upon a taxpayer (an individual or legal entity) by a state or the functional equivalent of a state*  #  excise tax (duty) / excise land tax  indirect tax income tax progressive tax direct tax |
| **#** | # | **#**  coupon tax  value added tax (VAT) personal income tax (PIT) corporate income tax (CIT) inheritance and gift tax  sales tax  **#**  before tax after tax taxable after tax tax-free  **#**  to raise (to collect) tax to deduct tax  to tax  to withhold tax  to evade (to dodge) a tax to pay tax  \*  **tax / ﬁscal**  tax base  tax declaration (form / return / sheet)  tax relief tax burden  tax haven (exile) |
| налог на дивиденды | podatek od dywidendy |
| налог на добавленную | podatek od wartości dodanej |
| стоимость (НДС) | (VAT) |
| налог на доходы физиче- | podatek od dochodów osób |
| ских лиц | ﬁzycznych |
| налог на прибыль орга- | podatek od dochodów osób |
| низаций | prawnych |
| налог с наследства и да- | podatek od darowizn i spad- |
| рений | ków |
| налог с продаж | podatek od sprzedaży |
| **#** | **#** |
| до удержания налога | przed potrąceniem podatku |
| за вычетом налога | po potrąceniu podatku |
| облагаемый налогом | podlegający opodatkowaniu |
| после удержания налога | po potrąceniu podatku |
| свободный от налога | wolny od podatku |
| **#** | **#** |
| взимать налог | ściągać (pobierać) podatek |
| вычитать налог | odliczać (potrącać) podatek |
| облагать налогом | opodatkowywać / nakładać |
| удерживать налог | podatek |
| уклоняться от оплаты | potrącać podatek |
| налога | uchylać się od płacenia podatku |
| уплачивать налог | uiszczać (płacić) podatek |
| \* | \* |
| **налоговый** | **podatkowy** |
| налоговая база | podstawa opodatkowania |
| налоговая декларация | zeznanie podatkowe |
| налоговая льгота | ulga podatkowa |
| налоговое бремя | ciężar podatkowy |
| налоговый рай | oaza podatkowa / raj podat-  kowy |

|  |  |  |
| --- | --- | --- |
| **налогооблагаемый**  налогооблагаемый доход **налогообложение** льготное налогообложе- ние  прямое налогообложение налогообложение капи- тала **налогоплательщик** обложенный налогом | **podlegający opodatkowaniu**  dochód podlegający opodat- kowaniu  **opodatkowanie**  opodatkowanie ulgowe  opodatkowanie bezpośrednie opodatkowanie kapitału  **podatnik opodatkowany** | **taxable / chargeable / assessable**  taxable income  **taxation / cessing**  soft taxation  direct taxation capital taxation  **taxpayer**  taxed |

Вышеописанный русско-польско-английский словарь экономической терминологии представляет интерес прежде всего для переводчиков, потому что он содержит базовую экономическую терминологию с экви- валентами на двух языках. Словарь имеет ярко выраженый учебный характер, о чем свидетельствует наличие дефиниций заглавных терминов на трех языках и полное соблюдение принципа лексикографической сим- метрии. Такие компоненты микроструктуры словаря, как терминологи- ческая сочетаемость и словобразовательное гнездо заглавного термина, несомненно, будут полезными и для переводчиков, и для изучающих русский, польский или английский язык экономики.

Создание вышепредставленного словаря в традиционной форме является первым этапом работы авторского коллектива. Следующий этап — это конструирование электронной версии словаря, которая будет представлять собой терминологическую базу данных с поисковой систе- мой. Электронная версия предоставит пользователям возможность легко менять исходный язык словаря и в зависимости от потребности пользо- ваться польско-английско-русским, англо-русско-польским или русско- польско-английским словарями. Она сделает также возможным находить необходимый термин в любой части терминологического минимума при помощи системы гиперссылок. Планируются также:

* повышение степени точности словаря путем добавления интенсио- нальных дефиниций;
* расширение словника за счет новых терминов и тех единиц, частот-

ность употребления которых значительно повысится;

* добавление к словарю новых языковых модулей;
* создание электронного трехъязычного тезауруса экономической терминологии;
* создание УМК по языку экономики, содержащего мультимедийные

материалы, доступные в сети интернет.

**Литература**

* 1. *Лейчик В.М.* (1989), Опыт построения классификаци терминологических словарей // Теория и практика научно-технической лексикографии. Москва.
  2. *Michałowski P.* (2014), Принципы конструирования двуязычного пред- метного словаря — торгово-финансовая терминология. KOMUNIKACJA SPECJALISTYCZNA, vol. 7, pp. 84–95.
  3. *Zmarzer W.* (1991), Leksykograﬁa terminologiczna // Teoretyczne podstawy terminologii. Wrocław.

**Literature**

1. *Leychik V.M.* (1989), Opyt postroeniya klassifikatsi terminologicheskikh slovarey // Teoriya i praktika nauchno-tekhnicheskoy leksikograﬁi. Moskva.
2. *Michałowski P*. (2014), Printsipy konstruirovaniya dvuyazychnogo predmetnogo slovarya — torgovo-finansovaya terminologiya. KOMUNIKACJA SPECJALISTYCZNA, vol. 7, pp. 84–95.
3. *Zmarzer W.* (1991), Leksykograﬁa terminologiczna // Teoretyczne podstawy terminologii. Wrocław.

***A.В. Мочалова, В.А. Мочалов***

**ЯЗЫК ОПИСАНИЯ ОНТОЛОГИЧЕСКИ-СЕМАНТИЧЕСКИХ ПРАВИЛ ДЛЯ ВОПРОСНО-ОТВЕТНОЙ СИСТЕМЫ**1

**Аннотация.** В работе предлагается новый язык описания онтологически- семантических правил, которые используются в семантическом анализаторе, на котором базируется разрабатываемая авторами ВОС. Демонстрируются при- меры языковых конструкций, позволяющих извлекать семантические отношения из анализируемого текста. Приводится пример генерируемого кода онтологически- семантического правила на языке Drools.

**Ключевые слова.** Вопросно-ответная система, онто-семантический анали- затор, онто-семантические шаблоны.

***A.V. Mochalova, V.A. Mochalov***

**A LANGUAGE TO DESCRIBE ONTOLOGICAL-SEMANTIC RULES FOR QUESTION-ANSWER SYSTEM**

**Abstract.** In the paper we propose a new language to describe ontological-semantic rules used in a semantic analyzer laying in the base of the question answering system being developed by us. We demonstrate examples of language constructs that allow to eliminate semantic relations from the analyzed text. We also provide an example of an ontological-semantic rule to generate the code, written in Drools language.

**Keywords.** question answering system, semantic analyzer, ontological-semantic rules.

**Введение**

В современном мире с увеличением количества текстов, представлен- ных в электронном виде (web-странцы, оцифрованные печатные издания, электронные библиотеки и т. п.), все более актуальным становится раз- работка систем автоматической обработки текста. К системам такого рода относятся и вопросно-ответные системы (ВОС). Архитектура каждой ВОС, как правило, имеет свои уникальные особенности. Особенности эти могут зависеть от разных факторов, например, от предметной области ВОС, от программных продуктов, от имеющихся у разработчиков ВОС программных продуктов и баз данных на начало программной реализации

1 Работа выполнена при финансовой поддержке РГНФ в рамках научного проекта

№15-04-12029 «Программная разработка электронного ресурса с онлайн-версией русско- язычной вопросно-ответной системы».

© МОЧАЛОВА Анастасия Викторовна. — Mochalova Anastasia. Институт космофизических исследований и распространения радиоволн ДВО РАН, Россия. — Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS, Russia. E-mail: [a.mochalova@ikir.ru](mailto:a.mochalova@ikir.ru)

© МОЧАЛОВ Владимир Анатольевич. — Mochalov Vladimir. Институт космофизических исследований и распространения радиоволн ДВО РАН, Россия. — Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS, Russia. E-mail: [vmochalov@ikir.ru](mailto:vmochalov@ikir.ru)

системы, от требуемого качества ответа и др. Различные архитектуры ВОС могут базироваться на использовании различных методов. Напри- мер, система TextAnalyst [TextAnalyst 2016] формирует ответы на вопросы пользователя с помощью статистических методов, в основе работы ВОС, описанной в работе [Мозговой 2006], лежит использование семантиче- ского словаря Тузова В.А. [Тузов 2004], работа ВОС, представленной в статье [Богуславский и др. 2015], также базируется на использовании результатов семантического анализатора. Некоторые из ВОС интегриро- ваны с различными базами данных и онтологий. Например, ВОС QASYA [Moussa et al. 2011] использует YAGO Ontology [YAGO 2016]. Система, описанная в работе [Athira et al. 2013], при ответе на вопрос также исполь- зует онтологическую информацию. Отличительная особенность ВОС WolframAlpha [WolframAlpha 2016] заключается в том, что ответы она не ищет, а вычисляет посредством обработки данных из уникальной БД. В работе предлагается новый язык описания онтологически- семантических правил, которые используются в семантическом ана- лизаторе, алгоритм работы которого описан в работе [Мочалова 2014]. На результатах работы этого анализатора базируется ВОС [Mochalova

2015] (находится в стадии разработки).

### Онтологически-семантические правила

Онтологически-семантические правила (ОСП) работают на тексте, в котором выделены синтаксемы — минимальные, далее неделимые семантико-синтаксические единицы русского языка, выступающие одно- временно как носитель элементарного (категориально-семантического) смысла и как конструктивный элемент более сложных синтаксических построений [Золотова 1988]).

Онтологически-семантические правила могут использовать инфор- мацию из онтологии. В этой работе будет рассматриваться объектно- ориентированная онтология, для которой определены функции ее изме- нения. Более подробно структура такой онтологии и функции для работы с ней описаны в статье [Kuznetsov et al. 2015].

Онтологически-семантическим правилом назовем правило, по кото- рому экспертная система (ЭС) в анализируемом тексте (где выделены синтаксемы и каждая из них отнесена к определенному классу или объекту онтологии) находит семантические зависимости между классами и объ- ектами, из которых он состоит. ОСП, являющееся правилом ЭС, состоит из левой и правой части. Левая часть описывает условия, при выполнении которых выполняются действия, описанные в правой части правила. Так, например, в левой части правила всегда описывается двусвязный список фактов, а также могут описываться булевы функции, аргументами которых являются факты из этого двусвязного списка [Мочалова и др. 2016].

Правая часть правила содержит список действий, каждое из которых может: модифицировать любой факт ЭС (путем изменения отношения в соответствующем ему Классе или Объекте из общей онтологии); доба- вить в очередь на удаление факт ЭС, имеющий определенный приоритет удаления; другие действия.

### Грамматика языка описания онто-семантических правил

Ниже приведен программный Java-код парсера на основе приведенной грамматики получен с помощью набора средств для написания парсеров ANTLR [ANTLR 2016] .

*grammar OntoDrools;*

*prog: property\* template+ ; property: ID '=' type ';' ;*

*template: ((initTemplateCommand ';')+)? when (left ';' )+ then (right ';' )+ ; initTemplateCommand: ID '(' params? ')' ;*

*left: (type '=')? (ID '(' params? ')' '&'?)+ ; right: ID '(' params? ')' ;*

*when: 'when' ;*

*then: '=>' ;*

*type: STRING | ID | DOUBLE | INT ;*

*typeList: type (',' type)\* ; arrayList: '[' typeList ']';*

*params: type # typeParam*

*| arrayList # arrayListParam*

*| params (op=('&'|'|') params)+ # logicParams*

*| params (',' params)+ # listParams*

*| '!'? '(' params ')' # parens*

*;*

*STRING: '"' (ESC|.)\*? '"' ;*

*fragment*

*ESC : '\\"' | '\\\\' ; // 2-char sequences \" and \\*

*ID : ([0-9]\*[a-zA-ZабвгдеёжзийклмнопрстуфхцчшщъыьэюяАБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ\_$]*

*+[0-9]\*)+ ; // match identiﬁers INT : [0-9]+ ; // match integers*

*DOUBLE : [0-9]+'.'[0-9]+ ; // match double*

*NEWLINE:'\r'? '\n' -> skip; // return newlines to parser (is end-statement signal) WS : [ \t]+ -> skip;*

*COMMENT : '/\*' .\*? '\*/' -> skip; LINE\_COMMENT : '//' ~[\r\n]\* '\r'? ‘\n’ -> skip;*

На текущий момент ОСП преобразуются с помощью Java-программы в скрипт экспертной системы Drools. Онтологически-семантический анализатор [Мочалова 2015] выполняет преобразование исходного текста в набор фактов, помещает их в рабочую память экспертной системы и запускает блок логического вывода на выполнение [Kuznetcov et al. 2015].

Все ОСП в рамках одного проекта логически разбиваются на файлы. Пример файла с описанием ОСП для определения семантического отно- шения «ПРИЗНАК»:

*ПРИОРИТЕТ\_ПРАВИЛА\_ПО\_УМОЛЧАНИЮ = 100;*

*ПРИОРИТЕТ\_ДВУ\_ПО\_УМОЛЧАНИЮ = 6;*

*when 0=МОРФ(П);*

*1=МОРФ(С); СОГЛ([0,1],[МОРФ\_ПАДЕЖ,МОРФ\_ЧИСЛО]);*

*=> СЕМ\_ДВУ(ПРИЗНАК,1,0,DEFAULT);*

*when 0=МОРФ(С);*

*1=МОРФ(П); СОГЛ([0,1],[МОРФ\_ПАДЕЖ,МОРФ\_ЧИСЛО]);*

*=> СЕМ\_ДВУ(ПРИЗНАК,0,1,DEFAULT);*

В приведенном выше файле атрибут *«ПРИОРИТЕТ\_ПРАВИЛА\_ПО\_ УМОЛЧАНИЮ»* определяет приоритет по умолчанию срабатывания правила экспертной системы (см. ниже: salience в правиле Drools), атрибут *«ПРИОРИТЕТ\_ДВУ\_ПО\_УМОЛЧАНИЮ»* определяет приоритет по умолчанию добавления факта в очередь на удаление, условие when определяет начало левой части ОСП, *МОРФ* определяет морфологические характеристики факта (могут записываться логическим выражением), логическая функция *СОГЛ* определяет согласованность списка фактов по списку атрибутов (в приведенном примере согласованными должны быть два подряд идущих факта *[0,1]* по морфологическим характеристикам *[МОРФ\_ПАДЕЖ,МОРФ\_ЧИСЛО])*, оператор “=>” определяет начало правой части ОСП, действие *СЕМ\_ДВУ* осуществляет изменение онтологии путем добавления семантического отношения

*«ПРИЗНАК»* между *1* и *0* фактом и добавляет в очередь на удаление факт 0

с приоритетом по умолчанию.

Ниже представлено сгенерированное с помощью Java-программы для экспертной системы Drools первое правило:

*rule "ПРИЗНАК1"*

*salience 100 when*

*$w0 : Word( partOfSpeech == "П" )*

*$w1 : Word( prev == $w0, partOfSpeech == "С" )*

*eval ( Unit.hasAnyIntersectionsInSelectedOntoTypes(new Unit[]{$w0, $w1}, new String[]{"МОРФ\_ ПАДЕЖ", "МОРФ\_ЧИСЛО"}) )*

*then*

*SemanticRelation sem0 = new SemanticRelation("ПРИЗНАК");sem0.setLeftAutoPosInText($w1); sem0.setRightAutoPosInText($w0);String indexSem0 = sem0.getIndexString(); if(hsAllIndexedSemanticRelations.contains(indexSem0) == false){*

*hsAllIndexedSemanticRelations.add(indexSem0);insert(sem0);} boolean changed0 = myQueue.addOrUpdateCheckToDelete($w0, 6); if(changed0)*

*update(myQueue);*

*end*

Пример ОСП для определения семантического отношения «ОДНО- РОДНЫЕ»:

*when 0=МОРФ(П|С);*

*1=ЗНАЧ\_Н(","|"или"|"и"|"да");*

*2=МОРФ(П|С);*

*СОГЛ([0,2],[МОРФ\_ПАДЕЖ,МОРФ\_ЧИСЛО,МОРФ\_РОД, МОРФ\_ЧАСТЬ РЕЧИ]);*

*=>*

*СЕМ\_ДВУ(ОДНОРОДНЫЕ,0,[2, 1],DEFAULT);*

Пример ОСП для модификации онтологии по найденному семанти- ческому отношению:

*when*

*СЕМ\_ОТН(ДействиеКтоЧтоКому(ПОДАРИЛ, [$X,$Y,$Z]));*

*=>*

*СЕМ\_УДАЛИТЬ($X, ИМЕЕТ, $Y); СЕМ\_ДОБАВИТЬ($Z, ИМЕЕТ, $Y);*

**Литература**

1. *Богуславский И.М., Диконов В.Г., Иомдин Л.Л., Лазурский А.В., Сизов В.Г., Тимошенко С.П.* (2015), Семантический анализ и ответы на вопросы: система в стадии разработки // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог». Вып. 14 (21): в 2 т. Т. 1: Основная программа конференции. М.: Изд-во РГГУ, с. 62–79.
2. *Золотова Г.А.* (1988), Синтаксический словарь: Репертуар элементарных единиц русского синтаксиса. М.: Наука.
3. *Мозговой М.В.* (2006), Простая вопросно-ответная система на основе семантического анализатора русского языка // Вестник СПб университета. Сер. 10. Вып. 1, с. 116–122.
4. *Мочалова А.В.* (2014), Алгоритм семантического анализа текста, основан- ный на базовых семантических шаблонах с удалением // Научно-технический вестник информационных технологий, механики и оптики. № 5, с. 126–132.
5. *Мочалова А.В.* (2016), Модификация онтологии с помощью базовых онто- семантических правил, содержащих глаголы // Теория и практика современных гуманитарных и естественных наук. Вып. 6. Ч. II: сборник научных статей еже- годной межрегиональной научно-практической конференции, Петропавловск- Камчатский, 08–12 февраля 2016 г. Петропавловск-Камчатский: КамГУ им. Витуса Беринга, с. 192–197, 205 с.
6. *Мочалова А.В.* (2015), Свидетельство о государственной регистрации про- граммы для ЭВМ «Программа семантического анализа текста, основанная на базо- вых семантических шаблонах с удалением», № 2015613430, Россия. 28.01.2015.
7. *Тузов В.А.* (2004), Компьютерная семантика русского языка. СПб.: Изд-во С.-Петерб. ун-та.
8. *ANTLR.* (2016): [www.antlr.org](http://www.antlr.org/) (дата обращения 17.08.2016)
9. *Athira P.M., Sreeja M., Reghuraj P.C.* (2013), Architecture of an Ontology-Based Domain-Speciﬁc Natural Language Question Answering System // International Jour- nal of Web & Semantic Technology (IJWesT) Vol. 4, No. 4, October 2013, pp. 31–39.
10. *Kuznetsov V.A., Mochalov V.A., Mochalova A.V.* (2015), Ontological-semantic text analysis and the question answering system using data from ontology // ICACT Transactions on Advanced Communications Technology (TACT) Vol. 4, Issue 4, July 2015, pp. 651–658.
11. *Mochalova A.* Search for answers in ontological-semantic graph. Proceedings of the AINL-ISMW FRUCT, Saint-Petersburg, Russia, 9–14 November 2015, ITMO University, FRUCT Oy, Finland, рр. 174–180.
12. *Moussa, Abdullah M. & Rehab , Abdel-Kader* (2011), QASYO: A Question Answering System for YAGO Ontology . International Journal of Database Theory and Application. Vol. 4, No 2, June, 2011, р. 99.
13. *TextAnalyst* [(2016), http://www.analyst.ru](http://www.analyst.ru/) (дата обращения 10.08.2016)
14. *WolframAlpha* (2016): <http://www.wolframalpha.com/>(дата обращения 15.08.2016).
15. *YAGO*. (2016), [www.yago-knowledge.org/](http://www.yago-knowledge.org/) (дата обращения 15.08.2016).

**Literature**

1. *Boguslavskiy I.M., Dikonov V.G., Iomdin L.L., Lazurskiy A.V., Sizov V.G., Timoshenko S.P.* (2015), Semanticheskiy analiz i otvety na voprosy: sistema v stadii razrabotki // Komp’yuternaya lingvistika i intellektual’nye tekhnologii: Po materi- alam yezhegodnoy Mezhdunarodnoy konferentsii «Dialog». Vyp. 14 (21): v 2 t. T. 1: Osnovnaya programma konferentsii. M.: Izd-vo RGGU, s. 62–79.
2. *Zolotova G.A.* (1988), Sintaksicheskiy slovar’: Repertuar elementarnykh yedinits russkogo sintaksisa. M.: Nauka.
3. *Mozgovoy M.V*. (2006), Prostaya voprosno-otvetnaya sistema na osnove seman- ticheskogo analizatora russkogo yazyka // Vestnik SPb universiteta. Ser. 10. Vyp. 1, s. 116–122.
4. *Mochalova A.V*. (2014), Algoritm semanticheskogo analiza teksta, osnovannyi na bazovykh semanticheskikh shablonakh s udaleniem // Nauchno-tekhnicheskiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki. № 5, s. 126–132.
5. *Mochalova A.V*. (2016), Modiﬁkatsiya ontologii s pomoshch’yu bazovykh onto-semanticheskikh pravil, soderzhashchikh glagoly // Teoriya i praktika sovremen- nykh gumanitarnykh i yestestvennykh nauk. Vyp. 6. Ch. II: sbornik nauchnykh statey yezhegodnoy mezhregional’noy nauchno-prakticheskoy konferentsii, Petropavlovsk- Kamchatskiy, 08–12 fevralya 2016 g. Petropavlovsk-Kamchatskiy: KamGU im. Vitusa Beringa, s. 192–197, 205 s.
6. *Mochalova A.V*. (2015), Svidetel’stvo o gosudarstvennoy registratsii programmy dlya EVM «Programma semanticheskogo analiza teksta, osnovannaya na bazovykh semanticheskikh shablonakh s udaleniem», № 2015613430, Rossiya. 28.01.2015.
7. *Tuzov V.A.* (2004), Komp’yuternaya semantika russkogo yazyka. SPb.: Izd-vo S.-Peterb. un-ta.
8. ANTLR. (2016): [www.antlr.org](http://www.antlr.org/) (data obrashcheniya 17.08.2016)
9. *Athira P.M., Sreeja M., Reghuraj P.C.* (2013), Architecture of an Ontology-Based Domain-Speciﬁc Natural Language Question Answering System // International Jour- nal of Web & Semantic Technology (IJWesT) Vol. 4, No. 4, October 2013, pp. 31–39.
10. *Kuznetsov V.A., Mochalov V.A., Mochalova A.V*. (2015), Ontological-semantic text analysis and the question answering system using data from ontology // ICACT

Transactions on Advanced Communications Technology (TACT) Vol. 4, Issue 4, July 2015, pp. 651–658.

1. *Mochalova A.* Search for answers in ontological-semantic graph. Proceedings of the AINL-ISMW FRUCT, Saint-Petersburg, Russia, 9–14 November 2015, ITMO University, FRUCT Oy, Finland, рр. 174–180.
2. *Moussa, Abdullah M. & Rehab , Abdel-Kader* (2011), QASYO: A Question Answering System for YAGO Ontology . International Journal of Database Theory and Application. Vol. 4, No 2, June, 2011, р. 99.
3. Te[xtAnalyst (2016), http://www.analyst.ru](http://www.analyst.ru/) (dataobrashcheniya 10.08.2016)
4. WolframAlpha (2016): <http://www.wolframalpha.com/>(dataobrashcheniya 15.08.2016).
5. YAGO. (2016), [www.yago-knowledge.org/](http://www.yago-knowledge.org/) (dataobrashcheniya 15.08.2016).

***О.А. Невзорова, А.М. Галиева, В.Н. Невзоров***

**К РАЗРАБОТКЕ АВТОМАТИЧЕСКИХ МЕТОДОВ ВЫЯВЛЕНИЯ СООТВЕТСТВИЯ ГРАММАТИЧЕСКИХ МОДЕЛЕЙ ТЕРМИНОВ В РАЗНОСТРУКТУРНЫХ ЯЗЫКАХ**1

**Аннотация.** В статье обсуждаются первые результаты, полученные при раз- работке русско-татарского общественно-политического тезауруса. При переводе понятий тезауруса актуальной является задача выявления переводных моделей на уровне грамматики и разработка автоматических методов выявления соответ- ствий переводных моделей. Построены частотные классы переводных моделей, исследованы их системные признаки.

**Ключевые слова.** Словосочетание, перевод, грамматические модели, общественно-политическая лексика.

***O.A. Nevzorova, A.M. Galieva, V.N. Nevzorov***

**ON DEVELOPING AUTOMATIC METHODS OF DETECTING CORRESPONDENCE BETWEEN GRAMMATICAL MODELS IN STRUCTURALLY DIFFERENT LANGUAGES**

**Abstract.** This paper presents preliminary results of developing the Russian-Tatar thesaurus of social and political vocabulary. In the course of compiling and translating the concepts of the thesaurus the task of deﬁning grammatical translation models and developing automatic methods of detecting correspondence between these models is a topical matter. We built frequency classes of translation models and studied their characteristics on derivational and syntactic levels.

**Key words.** Сollocations, translation, grammar patterns, social and political vocabulary.

Системное исследование лексической системы языка является важной задачей лингвистики. Одним из методов изучения системных отношений между лексическими единицами является разработка тезауруса. В статье изложены первые результаты, полученные в ходе разработки русско- татарского общественно-политического тезауруса. Данный тезаурус

1 Исследование выполнено при финансовой поддержке Российского научного фонда (про- ект № 16-18-02074 «Разработка моделей связывания терминологии в разных языках (на материале русского и татарского языков)»).

© НЕВЗОРОВА Ольга Авенировна. — Nevzorova Olga. Академия наук Республики Татарстан, Казанский федеральный университет, Россия. — Tatarstan Academy of Sciences, Kazan Federal University, Russia. E-mail: [onevzoro@gmail.com](mailto:onevzoro@gmail.com)

© ГАЛИЕВА Альфия Макаримовна. — Galieva Alphia. Академия наук Республики Татар- стан, Россия. — Tatarstan Academy of Sciences, Russia. E-mail: [amgalieva@gmail.com](mailto:amgalieva@gmail.com)

© НЕВЗОРОВ Владимир Николаевич. — Nevzorov Vladimir. Казанский национальный исследовательский технический университет им. А. Н. Туполева, Россия. Kazan National Research Technical University named after A.N. Tupolev, Russia. — E-mail: [nevzorovvn@gmail.com](mailto:nevzorovvn@gmail.com)

разрабатывается на основе тезауруса РуТез [Лукашевич 2011], поэтому татарский тезаурус использует концептуальную структуру этого тезауруса. Главной задачей на данном этапе является перевод понятийных единиц тезауруса на татарский язык, уточнение набора понятий и их содержания. При переводе понятий тезауруса актуальной является задача выявления пере- водных моделей на уровне грамматики и разработка автоматических методов выявления соответствий переводных моделей. При разработке переводных моделей учитывался многолетний опыт исследований по машинному пере- воду и лексикографированию специальной лексики [Беляева 2010; Беляева 2011; Рычкова 2016; Neelameghan et al. 2007; Déjean et al. 2002].

В статье анализируются способы перевода наиболее частотных грам- матических моделей словосочетаний на русском языке, представленных в тезаурусе РуТез. Перевод на татарский язык выполняется на основе раз- личных источников — двуязычных русско-татарских словарей, текстов офи- циальных документов, размещенных на портале «Официальный Татарстан» (http://tatarstan.ru), текстов СМИ, а также данных Татарского национального корпуса «Туган тел» (http://corpus.antat.ru).

Рассмотрим наиболее частотные грамматические модели в тезаурусе РуТез и способы их перевода на татарский язык. Частотные грамматические модели выявлены на основе словообразовательного и частичного синтак- сического анализа, выполненного на основе лингвистических технологий системы “OntoIntegrator” [Невзорова и др. 2012]. Для построения списка грамматических моделей был проанализирован представительный список понятий тезауруса (cвыше 3000 понятий) и выявлены классы грамматических моделей. Некоторые частотные классы моделей обсуждаются ниже в статье.

1. Перевод словосочетаний, образованных по модели *«сущ. + сущ в Р.п.»*

Основная масса терминологических сочетаний русского языка в выборке представляет следующие частные падежные значения — родительный определительный и родительный объекта. При их переводе на татарский язык наблюдается четкая корреляция: в абсолютном боль- шинстве случаев словосочетания, имеющие в своем составе родительный определительный, переводятся изафетом второго или третьего типа, родительный объекта — татарским словосочетанием, содержащим имя действия *+* существительное в аккузативе.

Табл. 1 дает общее представление о распределении грамматических моделей переводных соответствий русским словосочетаниям.

Русские словосочетания, при переводе которых используется форма аккузатива, характеризуются следующими особенностями: первое (опор- ное) существительное обозначает отвлеченное действие (как правило, раз- ные типы отглагольных существительных), зависимое существительное в генитиве — объект этого действия (*схрон оружия, убийство человека, эксплуатация труда, аренда помещений).*

*Таблица 1.* Распределение способов перевода русских словосочетаний, образованных по модели «сущ. + сущ в Р.п.»

|  |  |
| --- | --- |
| **Способы перевода словосочетаний** | **Общее количество в выборке** |
| Изафет 2 | 310 |
| Изафет 3 | 57 |
| Имя действия *+* сущ. в аккузативе | 84 |
| Другое | 36 |
| Всего в выборке | 486 |

Системных различий при переводе родительного субъекта и родитель- ного определительного нами не выявлено — в обоих случаях при переводе используется изафет, например:

*труд осужденных* (родительный субъекта) — *хөкем ителгəннəр хезмəте* (изафет 2);

*театр оперы* (родительный определительный) — *опера театры*

(изафет 2).

В ряде случаев (11 примеров) при переводе на татарский язык зави- симое существительное имеет аффикс локатива:

*участники акции* — *акциядə* (LOC) *катнашучылар.*

Во всех подобных случаях в татарском языке опорным словом явля- ется существительное, образованное от имени действия при помощи аффикса -*чы/-че* и обозначающее лицо по действию.

Представляет интерес выяснение того, какими факторами объясняется выбор при переводе изафета второго или третьего типа.

Изафет 2 — словосочетание типа:

*шəһəр башлыгы* город глава,-POSS\_3 ‘градоначальник’.

Изафет 3 — словосочетание типа:

*шəһəрнең башлыгы*

город-GEN глава-POSS\_3 ‘глава [данного] города’.

Выбор изафета 3 в большинстве случаев обусловлен структурными факторами. Как правило, изафет 3 используется в случаях, когда в сло- восочетании непосредственно перед опорным словом (крайним правым элементом словосочетания) стоит определение, т. е. основные компоненты словосочетания (существительные или имена действия) разделены опреде- лением. В этом случае использование аффикса генитива позволяет четко фиксировать грамматические и смысловые отношения между словами:

*высший орган государственной власти* — *дəүлəт хакимиятенең*

(GEN) ***югары*** *органы.*

В ряде случаев предпочтение изафета 3 обусловлено семантическими факторами — опорным словом является существительное, обозначающее человека: *дневник ученика* — *укучының* (GEN) *көндəлеге.*

Во многих случаях выбор изафета 2 или изафета 3 во многом может быть обусловлен контекстом [Галиева 2014].

1. Перевод словосочетаний, образованных по модели «*прил. + сущ.»* В исходной выборке основная масса русских терминологических соче- таний имеет в своем составе относительное прилагательное. Поскольку в тюркских языках относительные прилагательные образуют сравнительно небольшой класс слов, заимствованных из других языков (преимуще- ственно арабского, персидского и западноевропейских), в большинстве случаев при переводе используются изафетные конструкции, образуемые существительными. Изафетные конструкции служат средством выраже- ния относительного признака, отвлекаемого от значения компонента- определения; наличие такого языкового инструмента значительно снижает потребность тюркских языков в относительных прилагательных [Гузев

и др. 2007].

Как показывает материал, в большинстве случаев русские произ- водные относительные прилагательные переводятся мотивирующими существительными или их татарскими эквивалентами/соответствиями:

*аттестационный документ* — *аттестация документы; больничная палата* — *хастаханə палатасы.*

Табл. 2 дает общую картину распределения основных способов пере- вода русских терминологических сочетаний, образованных по модели

«прил. *+* сущ.»

*Таблица 2.* Распределение типовых моделей перевода русских атрибутивных словосочетаний

|  |  |
| --- | --- |
| **Способы перевода словосочетаний** | **Количество в выборке** |
| Прил. *+* сущ. | 131 |
| Сущ. *+* сущ. в разных формах Cущ. (PL) *+* сущ. | 362  29 |
| Причастие *+* сущ. | 7 |
| Всего в выборке | 500 |

Сохранение при переводе прилагательного происходит в следующих основных случаях:

1. замена относительного прилагательного на качественное:

*верховный совет* — *югары* (буквально: 'высокий') *совет*;

1. сложно восстановить исходное существительное из относительного прилагательного:

*гуманитарное образование* — *гуманитар белем бирү*;

1. в татарском языке имеется соответствующее относительное при- лагательное: *военное сотрудничество* — *хəрби хезмəттəшлек.*

Татарское существительное, функционирующее в качестве определе- ния, может стоять в форме множественного числа. Это происходит в двух основных случаях:

* 1. если определяемое существительное имеет собирательное значение и называет совокупность чего-либо (используются лексемы, обозначаю- щие такие понятия, как партия, объединение, система и т. п.): *коммуни- стическая партия* — *коммунистлар* (PL) *партиясе*;

*аптечная система* — *даруханəлəр* (PL) *системасы.*

* 1. если объект предполагает использование значительным числом субъектов: *детский сад* — *балалар* (PL) *бакчасы*;

Интересно, что в ряде случаев одно и то же русское прилагательное может по-разному переводиться на татарский язык:

*авторский коллектив* — *авторлар* (PL) *коллективы* (буквально «кол- лектив авторов»);

*авторское право* — *авторлык* (NMLZ) *хокукы* (буквально «право авторства»).

Таким образом, при переводе на татарский язык акцентируется то, что слово *авторский* (‘относящийся к автору’) имеет разную референцию, в одном случае обозначается группа самих авторов, во втором — явления, так или иначе соотносимые с авторством.

Проведенный анализ позволяет выявлять типовые соответствия грамматических моделей терминов в русском и татарском языках. Спо- собы перевода во многих случаях позволяют дифференцировать частные значения русских падежей и семантические классы лексики; полученные модели могут быть использованы для машинного перевода.

**Литература**

1. *Беляева Л.Н.* (2011), Корпусная лингвистика и перевод: потенциал и огра- ничения. Труды международной конференции «Корпусная лингвистика — 2011». СПб.: Изд-во филологического ф-та СПбГУ, с. 87–91.
2. *Беляева Л.Н.* (2010), Потенциал автоматизированной лексикографии и при- кладная лингвистика. Известия РГПУ им. А.И. Герцена, № 134, с. 70–79.
3. *Галиева А.М.* (2014), Индивидуализирующая семантика в системе татарских падежей. Типология морфосинтаксических параметров: материалы международ- ной конференции «Типология морфосинтаксических параметров 2014». Вып. 1, с. 12–22.
4. *Гузев В.Г., Бурыкин А.А.* (2007), Общие строевые особенности агглю- тинативных языков. Acta linguistica Petropolitana; Труды ИЛИ РАН; Т. 3; Ч. 1, с. 109–117.
5. *Лукашевич Н.В.* (2011), Тезаурусы в задачах информационного поиска. М.: Изд-во МГУ, 512 с.
6. *Невзорова О.А., Невзоров В.Н.* (2012), Интеллектуальная инструментальная система «OntoIntegrator» для задач автоматической обработки текстов. Труды конференции КИИ-2012. Т. 4. Белгород: Изд-во БГТУ, с. 92–99.
7. *Рычкова Л.* (2016), Социопрагматический аспект лексикографирования специальной лексики. Proceedings of the XVII EURALEX International Congress:

Lexicography and Linguistic Diversity. Tbilisi, Ivane Javakhishvili Tbilisi State University, pp. 541–552.

1. *Neelameghan, A., Raghavan, K. S.* (2007), Online bilingual thesaurus for subjects in the humanities: a case study, ICSD, pp. 489–505.
2. *Déjean, H., Gaussier, E., Sadat, F.* (2002, August). An approach based on multilingual thesauri and model combination for bilingual lexicon extraction, Proceedings of the 19th international conference on Computational linguistics, Vol. 1, pp. 1–7.

**Literature**

1. *Belyaeva L.N*. (2011), Korpusnaya lingvistika i perevod: potentsial i ograni- cheniya. Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2011». SPb.: Izd-vo ﬁlologicheskogo f-ta SPbGU, s. 87–91.
2. *Belyaeva L.N*. (2010), Potentsial avtomatizirovannoy leksikograﬁi i prikladnaya lingvistika. Izvestiya RGPU im. A.I. Gertsena, № 134, s. 70–79.
3. *Galieva A.M*. (2014), Individualiziruyushchaya semantika v sisteme tatarskikh padezhey. Tipologiya morfosintaksicheskikh parametrov: materialy mezhdunarodnoy konferentsii «Tipologiya morfosintaksicheskikh parametrov 2014». Vyp. 1, s. 12–22.
4. *Guzev V.G., Burykin A.A.* (2007), Obshchie stroevye osobennosti agglyutinativ- nykh yazykov. Acta linguistica Petropolitana; Trudy ILI RAN; T. 3; Ch. 1, s. 109–117.
5. *Lukashevich N.V*. (2011), Tezaurusy v zadachakh informatsionnogo poiska. M.: Izd-vo MGU, 512 s.
6. *Nevzorova O.A., Nevzorov V.N*. (2012), Intellektual'naya instrumental'naya sistema «OntoIntegrator» dlya zadach avtomaticheskoy obrabotki tekstov. Trudy konferentsii KII-2012. T. 4. Belgorod: Izd-vo BGTU, s. 92–99.
7. *Rychkova L.* (2016), Sotsiopragmaticheskiy aspekt leksikografirovaniya spetsial'noy leksiki. Proceedings of the XVII EURALEX International Congress: Lexi- cography and Linguistic Diversity. Tbilisi, Ivane Javakhishvili Tbilisi State University, pp. 541–552.
8. *Neelameghan, A., Raghavan, K. S.* (2007), Online bilingual thesaurus for subjects in the humanities: a case study, ICSD, pp. 489–505.
9. *Déjean, H., Gaussier, E., Sadat, F.* (2002, August). An approach based on mul- tilingual thesauri and model combination for bilingual lexicon extraction, Proceedings of the 19th international conference on Computational linguistics, Vol. 1, pp. 1–7.

***З.Г. Прошина***

**НЕОБХОДИМОСТЬ ИЗУЧЕНИЯ ОПОСРЕДОВАННОГО ПЕРЕВОДА СПЕЦИАЛИСТАМИ ЗАПАДНОЕВРОПЕЙСКИХ ЯЗЫКОВ**

**Аннотация.** В докладе обосновано изучение перевода восточных реалий с западноевропейских языков, выступающих в качестве языков-посредников, пре- жде всего с английского как глобального языка общения. Использование латиницы для записи, например китайских слов, нередко вызывает перенос законов прямого англо-русского перевода при транслитерации как приеме перевода, что искажает звуковой облик слова и делает невозможной его переводную обратимость. Ана- логичные проблемы возникают и при переводе соответствующих слов с русского языка на английский и другие европейские языки. С расширением культурно- языковых контактов между Россией и странами Востока проблема опосредован- ного перевода приобретает все большую актуальность. В докладе автор делится опытом внедрения данной проблематики в курс теории и практики перевода.

**Ключевые слова.** Перевод, варианты английского языка, латинизация, транс- литерация, транскрипция.

***Z.G. Proshina***

**NECESSITY TO STUDY INTERMEDIARY TRANSLATION BY STUDENTS OF WEST-EUROPEAN LANGUAGES**

**Abstract.** The presentation discusses the need to teach translators from European languages to deal with Asian culture-loaded words written in English as a Lingua Franca. The use of Romanized writing, especially for Chinese words, often makes translators transfer regularities of direct English-Russian translation to transliterating Asian words. This also takes place in translating from other European languages, as well as from Russian into English. With the increase in culture and language contacts between Russian and the East, the challenges of intermediary translation have become more demanding. In the presentation, the author will share her experience in introducing intermediary translation into translation programs.

**Keywords.** Translation, world Englishes, Romanization, transliteration, transcription.

Плюрицентричность современного английского языка проявляется в возникновении множества его вариантов и их постепенной стандарти- зации, по-разному проявляющейся на разных уровнях языка. Менее всего признана стандартизация вариантов, которые относят к так называемому Расширяющемуся кругу, по классификации Б. Качру [Kachru 1985], функ- ционирующих главным образом как лингва франка в ситуациях межкуль- турной коммуникации.

© ПРОШИНА Зоя Григорьевна. — Zoya Proshina. Московский государственный универ- ситет им. М.В. Ломоносова, Россия. Lomonosov Moscow State University, Russia. E-mail: [proshinazoya@yandex.ru](mailto:proshinazoya@yandex.ru)

Вместе с тем и в этих вариантах английского языка стандартизация имеет место не только в виде экзонорм, т. е. внешних норм, но и эндонорм, внутренних стандартов, что наблюдается прежде всего на фонографи- ческом уровне. Так, например, китайский вариант английского языка использует стандартизированную систему латинизации, известную как китайский фонетический алфавит, или пиньинь, разработанную в 1958 г. и утвержденную в качестве официальной системы латинизации в КНР в 1979 г. В Республике Корея новый стандарт латиницы был утвержден в 2000 г. и сегодня используется во всем мире. В создании этих систем передачи слов восточного происхождения буквами латинского алфавита принимали ученые именно этих стран, так что по праву эти фонографи- ческие системы могут считаться эндонормами, созданными не по образцу других стран, а внутри своей собственной культуры и далее распростра- нившейся в мире.

Сегодня латинизированное письмо нередко ассоциируется с англий-

ским языком как средством глобального общения, и многие реалии и соб- ственные имена, записанные на китайской или корейской латинице, оши- бочно воспринимаются так, словно они исконно возникли как английские, т. е. по «чужому» стандарту, или экзонорме. Такое восприятие, например китайской латиницы, сопряжено с целым рядом переводческих ошибок, возникающих в результате прямого переноса англо-русских корреляций на перевод китайских слов в английском тексте. Например, на сайтах неко- торых вузов находим явную транскрипцию с английского наименования китайской гимнастики *tai ji (quan)*: МГПУ — *По базовому образованию Стюарт Хеллер математик, специалист по системным операциям. В начале 70-х годов он увлекся восточными единоборствами и медицин- скими практиками. В этой сфере он достиг высоких результатов, став обладателем черных поясов 6-го и 7-го уровня (дана) по трем различным видам боевых искусств и профессором* ***Тай-джи***1. Институт Конфуция, ЕГЛУ им. В. Брюсова — ***Тай джи*** *ушу считается одним из видов китай- ского боевого искусства и пользуется мировой славой*2.

Однако в данном случае мы имеем дело не с прямым переводом, а с

опосредованным, имеющим свои собственные закономерности. Опо- средованный перевод — это перевод с языка-посредника, или лингва франка, обслуживающего две исконно «чужие» для него культуры, как, например, русскую и китайскую. В сущности, это перевод с одного из вариантов языка, отличающегося от прототипического британского. Одна из главных особенностей опосредованного перевода, который обычно имеет письменную форму, заключается в том, что он опирается на тради-

1 URL: <http://мпгу.рф/anonsyi/psihologii-prakticheskih-obrazovaniya/> (дата обращения 18.07.2016)

2 URL: <http://confucius.brusov.am/learning/tai_ji_wushu.php?l=R>(дата обращения 18.07.2016)

ции устного прямого перевода, например, с китайского языка на русский. Эти традиции нередко не совпадают с традициями англо-русской транс- литерации и транскрипции, что как раз и вызывает серьезные проблемы у переводчиков с английского, не знающих китайского языка. Подобного рода проблемы возникают не только в отношении перевода с английского, но и других европейских языков. Например, имя французского писателя китайского происхождения *Dai Sijie*, автора романа “Balzac et la petite tailleuse chinoise”, переведенного на русский язык, представляют в России не как *Дай Сыцзе*, а как *Дэ Сижи.* Такой «офранцуженный» (как и англи- зированный) перевод приводит к проблеме ксенонимической обратимости [Кабакчи 1998] или обратимости перевода [Прошина 2014], т. е. неузна- ваемости имени при обратном переводе с русского языка — *Дэ Сижи* может приобрести форму *De Sizhi*, или у человека, знакомого с китайской латиницей, — *De Xiri.*

Переводческие проблемы возникают не только в связи с нетрадици-

онностью корреляций опосредованного перевода в сравнении с прямым переводом с западноевропейских языков, но также и по другим причинам, в частности, потому что современному стандарту латинизации восточных слов предшествовали другие системы латинизации — например, китай- ские слова долгое время в английских текстах записывались в системе, предложенной британскими синологами Т. Уэйдом и Г. Джайлсом, в результате чего в русском языке при прямом переводе с английского *tai chi chuan*, соответствующего в пиньине форме *tai ji quan,* появляется *тай чи чуань*: *Во всем мире* ***тай-чи*** *(таково более распространенное название) считается именно оздоровительной гимнастикой1.*

Эти факторы свидетельствуют о настоятельной необходимости модер- низировать программы по переводу на факультетах западноевропейских языков отечественных вузов — особенно сегодня, когда политико- экономическая ситуация, обслуживаемая нашими переводчиками, явно смещается своим фокусом на Восток. Наши студенты, будущие перевод- чики (как, впрочем, и будущие педагоги, которым, возможно, предстоит работать со студентами из азиатских стран) должны получить навыки работы с текстами, включающими азиатские имена и реалии, т. е. навыки опосредованного перевода.

В курс теории и практики перевода необходимо ввести закономерно- сти транслитерации, по крайней мере китайских и японских имен. Послед- ние имеют гораздо меньше нетрадиционных корреляций, по сравнению с китайскими именами; однако очень многие японские слова заимствуются русским языком в англизированном виде: *суши* вместо *суси*, *сашими* вме-

1 2011 Будь здорова, URL: [http://medportal.ru/budzdorova/ﬁtness/354/](http://medportal.ru/budzdorova/%EF%AC%81tness/354/) (дата обращения 21.09.2014)

сто *сасими* «блюдо из сырой рыбы, нарезанной ломтиками»; *тамагочи* вместо *тамаготи* «детская игрушка в виде яйца, внутри которого живет существо, которое нужно поить, кормить и воспитывать с помощью кнопок»*, Хитачи* вместо *Хитати* «название компании и ее продукции»*, ваза* вместо *вадза* «техника борьбы в японских боевых искусствах» и т. п. Проблемными буквами в японских словах являются следующие: *сh = ть, sh = сь, j = дзь, z = дз*. Китайские имена в пиньине имеют достаточно много специфичных буквенных соответствий: *с = ц (ci — цы), j* = *цзь (ji — цзи), -n = -нь (yin — инь), ng = -н (yang — ян), q = ць (qi — ци), r = ж (ren — жэнь), x = сь (xi — си), z = цз (Ze — Цзэ),* не считая буквосочетаний (например, *Guo = Го; wu = у, -ong = -ун* и др.) (Прошина 2005). Следует обратить внимание, что в востоковедческой практике принято считать единицей перевода слог — именно поэтому таблицы транскрипций пиньинь и кириллицы Палладия обычно предлагают соответствия слогов. Вместе с тем западноевропейские переводчики, пользующиеся приемом транскрипции/транслитерации, обычно ориентируются на более мелкие единицы, звуки/буквы. Наша практика работы с переводчиками с англий- ского языка показывает, что переход на побуквенную транслитерацию вполне возможен и воспринимается студентами как логичная реализация привычного им принципа побуквенного перевода, хоть и дающего нетра- диционные корреляции.

На объяснение и закрепление особенностей опосредованного пере-

вода с каждого из азиатских вариантов английского языка уходит одно занятие; правда, для полного закрепления перевода китайских реалий требуется как минимум еще пара занятий. При этом хороший эффект дает сначала интенсивная тренировка англо-русского перевода одних слов (как нарицательных, так и собственных — антропонимов и топонимов), затем слов — в тексте, а потом осуществляется аналогичный перевод слов и текстов с русского на английский язык. Необходимость представления студентам англоязычных текстов с восточными реалиями необходимо для того, чтобы у них не создавалось впечатление, что их заставляют учить китайский язык и китайские слова — они переводят английский текст на китайскую тематику, который естественным образом включает китай- ские наименования. Как правило, при таком подходе студенты ощущают соприкосновение с экзотической для них восточной культурой, и это вызывает особенный интерес.

При обсуждении текстов с восточными наименованиями следует также

обращать внимание на то, как читаются эти наименования в англоязыч- ном тексте, поскольку в этом случае обычно работают законы аналогии и слово восточного происхождения адаптируется к системе принявшего его языка. Например, буква *q* в китайских заимствованиях в английском языке читается не как [k], а подобно первой букве слова *ji*; *x* произносится

как в слове *she,* а буква *z* читается как и в других английских словах, т. е. как [z], но в русском языке она соответствует *дз* в случае японского про- исхождения слова или *цз* — в случает китайского слова и т. д.

Подобного рода материал можно также включать в курс практического английского языка, профессионально ориентированного языка, лингво- культурологии, вариантологии английского языка и т. п.

Подведем итог сказанному. Внедрение в программы подготовки спе- циалистов с западноевропейскими языками материала, связанного с пере- дачей восточноазиатских реалий, обусловлено как лингвистическими, так и экстралингвистическими причинами: а именно изменением политико- экономической ситуации и необходимостью усиления контактов со стра- нами Востока, а также глобальной распространенностью английского языка, служащего языком-посредником в контактах России и азиатских стран. Подготовка к таким контактам должна вестись целенаправленно, а значит, требует педагогических усилий и пересмотра соответствующих программ подготовки переводчиков, педагогов, специалистов межкуль- турной коммуникации.

**Литература**

* 1. *Кабакчи В.В.* (1998), Основы англоязычной межкультурной коммуникации. СПб.: Изд-во РГПУ им. А.И. Герцена.
  2. *Прошина З.Г.* (2014), Опосредованный перевод и принцип обратимости пере- вода. Социальные и гуманитарные науки на Дальнем Востоке, № 1 (41), с. 179–184.
  3. *Прошина З.Г.* (2005), Передача китайских, корейских и японских слов при переводе с английского языка на русский и с русского языка на английский: Теория и практика опосредованного перевода. М.: Восток-Запад.
  4. *Kachru B.B.* (1985**)** Standards, codiﬁcation and sociolinguistic realism. The English language in the outer circle. In: Randolph Quirk and H.G. Widdowson (eds.). English in the World: Teaching and Learning the Language and Literatures. Cambridge: Cambridge University Press, pp.11–30.

**Literature**

1. *Kabakchi V.V.* (1998), Osnovy angloyazychnoy mezhkul'turnoy kommunikatsii. SPb.: Izd-vo RGPU im. A.I. Gertsena.
2. *Proshina Z.G.* (2014), Oposredovannyi perevod i printsip obratimosti perevoda. Sotsial'nye i gumanitarnye nauki na Dal'nem Vostoke, № 1 (41), s. 179–184.
3. *Proshina Z.G.* (2005), Peredacha kitayskikh, koreyskikh i yaponskikh slov pri perevode s angliyskogo yazyka na russkiy i s russkogo yazyka na angliyskiy: Teoriya i praktika oposredovannogo perevoda. M.: Vostok-Zapad.
4. *Kachru B.B.* (1985**)** Standards, codiﬁcation and sociolinguistic realism. The English language in the outer circle. In: Randolph Quirk and H.G. Widdowson (eds.). English in the World: Teaching and Learning the Language and Literatures. Cambridge: Cambridge University Press, pp.11–30.

***В.Э. Рогачева***

**МЕТОДЫ ИЗВЛЕЧЕНИЯ ТЕРМИНОЛОГИЧЕСКИХ ЭКВИВАЛЕНТОВ ИЗ ДВУЯЗЫЧНОГО КОРПУСА ТЕКСТОВ**

**Аннотация.** В статье рассматриваются проблемы извлечения терминов — универбов и терминологических словосочетаний из сопоставимого корпуса предметно-ориентированных текстов на русском и английском языках. Термины, извлеченные из подкорпусов на каждом языке, составляют терминосистемы, кото- рые могут быть сопоставлены. Сопоставление терминосистем дает возможности стандартизировать передачу терминов с одного языка на другой.

**Ключевые слова.** Извлечение терминологии, сопоставимый корпус текстов, критерии терминологичности, лингвостатистический анализ.

***V.E. Rogacheva***

## EXTRACTION OF EQUIVALENT TERMS FROM A BILINGUAL TEXT CORPUS

**Abstract.** The paper concerns the problems of term extraction from bilingual comparable text corpora, sampled of specialized Russian and English texts. The terms, extracted from Russian or English subcorpora, make parts of term systems that may be compared. By comparison of bilingual term systems terminology translation may be standardized.

**Keywords.** Terminology extraction, comparable text corpus, termhood criteria, statistic-linguistic analysis.

Межъязыковая коммуникация в сфере науки и техники в условиях глобализации — одно из важнейших направлений обмена информацией, которое актуализирует создание лексикографических источников, соот- ветствующих современному уровню развития науки и техники и предна- значенных для поддержки переводческого процесса.

В таких взаимосвязанных задачах, как создание терминологических словарей и корректный перевод терминологической информации научно- технического текста, важной становится адекватная передача терминоси- стемы одного языка на другой.

Потенциал терминологических ресурсов, создаваемых в научно- технической лексикографии и используемых в работе специалистов в области анализа текстов на естественном языке (*language workers*), делает актуальной разработку методов автоматического извлечения терминологических единиц из источников научно-технической инфор- мации — текстов.

© РОГАЧЕВА Валерия Эдуардовна. — Rogacheva Valeriya. Российский государствен- ный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University, Russia. E-mail: Valery. xxx. [zzz@gmail.com](mailto:zzz@gmail.com)

В основе подобных процедур извлечения терминов лежит идея сопо- ставимости контекстов слов с одинаковым значением в текстах на разных языках, представленных и выровненных в виде параллельных корпусов. Параллельные корпуса включают в себя тексты, которые являются пере- водами друг друга. Но у параллельных корпусов есть ряд недостатков, затрудняющих адекватную передачу терминосистемы с одного языка на другой. К ним относятся недостаточное качество переводов, в которых отражается влияние исходных текстов, невозможность дать переводам экс- пертную оценку, наблюдается недостаток доступных для использования параллельных текстов в различных предметных областях [Беляева 2014; Delpech, Daille 2010].

Сопоставимые корпуса текстов дают возможность наблюдать реальное использование терминов в тексте. Материалы такого корпуса отражают единую предметную область, но не являются переводами друг друга.

Исследование предполагает формирование корпуса сопоставимых текстов научных конференций «Терминология и знание» за 2008–2012 гг. и «Terminology and Knowledge engineering» за 2010–2014 гг. на английском и русском языках.

Материалы конференций — важный и доступный материал для состав- ления сопоставимых корпусов, так как они ориентированы на единую предметную область, отражают современные тенденции лингвистической науки и включают в себя новые термины [Беляева 2014, с. 41].

Извлечение терминов проводится при помощи сетевой программы AntConc — конкордансера, предназначенного для лингвостатистического анализа текста [Беляева, Пиотровская 2014, с. 27]. Инструментарий про- граммы позволяет выделить из корпуса наиболее частотные лексические единицы (ЛЕ) и исследовать их контексты в словосочетаниях. Семантика частотных терминов уточняется именно в терминологических словосо- четаниях, что позволяет представлять терминополе исследуемой области знаний в виде концептуальных структур, которые могут быть использо- ваны при проектировании онтологий и тезаурусов.

Алгоритм извлечения концептуальной иерархии терминов из корпуса текстов включает в себя следущие шаги:

* составление частотного словаря лексических единиц, зафиксирован- ных в русском и английском подкорпусах;
* выделение из частотного словаря терминов-универбов, учитывая существующие подходы к разграничению терминов и общеупотреби- тельных слов;
* разработка процедуры исследования контекстов на основе исполь- зования конкордансов и кластеров, в которых встречаются наиболее частотные термины-универбы.

Частотный словарь по каждому подкорпусу текстов строится на основе сервиса World list программы AntConc. В получившемся русском словаре 25 795 токенов — лексических единиц, в английском — 10 339.

Критерии, по которым ЛЕ из частотного словаря отбираются для дальнейшего исследования, включают принадлежность к словам откры- тых классов (существительные, прилагательные, полнозначные глаголы и наречия), чей состав постоянно изменяется [Gillam et al., 2005], а также ранг в частотном словаре.

Слова, имеющие высокую частоту в корпусе специальных текстов, могут появляться с низкой частотой в корпусе текстов общего языка, либо не появляться в нем вообще [Gillam et al., 2005]. Следовательно, частота в случае специализированного корпуса является показателем важности лексической единицы для специальной коммуникации [Кудашев, 2007].

Отраженное в толковании лексическое значение выбранных ЛЕ должно быть постоянным, не пересекаться семантически с лексическими значениями других единиц в терминосистеме, а отграничиваться от них [Шелов, 1998]. ЛЕ могут приобретать терминологичность, будучи вклю- ченными в терминополе специального текста [Табанакова, 2014]. Их значение раскрывается в терминологических словосочетаниях.

При помощи сетевой программы лингвостатистического анализа тек- ста AntConc можно получить данные об окружении лексических единиц, используя сервисные инструменты: кластеры и конкорданс.

Рассмотрим процедуру выделения словосочетаний с опорной ЛЕ *тер- минология (terminology)*. В соответствии с частеречной характеристикой этих ЛЕ предполагается, что программно выделяемый кластер — это совокупность именных групп (ИГ).

Инструмент Clusters позволяет выделять словосочетания, в которых ЛЕ является ядром, а поясняющие ее лексические компоненты находятся или в препозиции, или в постопозиции от ядра. Частота появления каждого слова в каждой позиции вокруг ядра регистрируется, и самые частотные словосочетания, компоненты которых связаны между собой синтаксически и семантически, могут быть признаны значимой для лек- сикографа моделью сочетаний слов. Инструмент Concordance позволяет рассматривать окружение ЛЕ с обеих сторон. Это позволяет увидеть появление ЛЕ в двух значимых моделях сочетаний слов одновременно, что невозможно заметить в кластере. Поэтому результаты, выдаваемые инструментами Clusters и Concordance, должны быть сопоставлены и объединены.

Для получения кластеров был задан минимум (1) и максимум (2) слов, окружающих ядерное слово справа и слева. Особенности морфологии русского и английского языков определяют процедуру анализа кластеров.

Развитая система падежных окончаний в русском языке делает необ- ходимым лемматизацию ЛЕ для рассмотрения ее появления в кластере в разных падежных формах. В результате получаем лемму терминолог\* (символом «\*» обозначаются падежные окончания).

Следующим шагом является составление списка всех ЛЕ, входящих в именные группы (ИГ) и включающих в себя искомую лемму. Затем следует компонентный анализ выделенных ИГ. ЛЕ *терминология* вклю- чает в себя семантический компонент *область знаний.* Соответственно ИГ, в которых этот компонент отсутствует, дальнейшему анализу не под- лежат.

Из множества результатов фиксации ЛЕ в конкордансе также выби- раются те, в которых присутствует сема *область знаний,* с последующей лемматизацией. Для слова *терминология* получен конкорданс, состоящий из 501 употребления слова в контексте. В каждой строке конкорданса выделены ИГ, в которые входит слово *терминология*, притом учитыва- ются однородные члены словосочетания. Эта мера позволяет рассмотреть исследуемый термин в более широком контексте, так в кластере упуска- ются однородные определения к термину, управляющие и управляемые им существительные, которые также могут иметь определения. С помощью инструментов поиска в документе выделены и подсчитаны наиболее часто встречающиеся слова, вступающие в синтаксические отношения с термином. При этом фактор, стоят ли эти слова в непосредственной близости к термину и имеют ли определения и дополнения, игнорируется. В списках, полученных с помощью конкорданса и кластеров, ищутся соответствия.

Для англоязычного подкорпуса была проделана аналогичная работа: в конкордансе выделены коллокации с ЛЕ *terminology*. На основе сравне- ния конкорданса с выделенными кластерами определена частота колло- каций.

ЛЕ *terminology* оказалась зависимым элементом словосочетания в большинстве выделенных коллокаций. В переводных эквивалентах именных групп ЛЕ выражается либо соответствующим прилагательным (*terminology resources — терминологические ресурсы)*, либо существитель- ным, выступающим в роли дополнения к ядерному элементу словосоче- тания (*terminology teaching — обучение терминологии*). Словосочетания, в которых ЛЕ является ядром, отбираются и сравниваются со списком словосочетаний на русском языке, в которых ядро — русскоязычный эквивалент ЛЕ.

В табл. дано сопоставление списков словосочетаний, в которых пред- ставлена пара переводных эквивалентов *терминология — terminology.*

*Таблица.* Списки сопоставляемых словосочетаний

|  |  |
| --- | --- |
| Научная терминология | Frame-based terminology |
| Народная терминология | Concept-oriented terminology |
| Лингвистическая терминология | Ontology-based terminology |
| Профессиональная терминология | Discourse-purposed terminology |
| Терминология православия | Bibliographic terminology |
| Археологическая терминология | Application terminology |
| Отраслевая терминология | Speciﬁc terminology |
| Православная терминология | Industrial practice terminology |
| Медицинская терминология | EU terminology |
| Предметная терминология | Theoretical terminology |
| Терминология прикладной науки | Applied terminology |
| Обрядовая терминология | Phraseological terminology |
| Научнонормативная терминология | Legal terminology |
| Терминология языка для специальных целей | Acquired terminology |
| Понятийно четко очерченная терми- нология | Well-developed terminology |
|  | Mature terminology |
|  | Corresponding terminology |

Из списков, представленных в табл., можно выделить соответствия:

1. Терминология языка для специальных целей — speciﬁc terminology
2. Терминология прикладной науки — applied terminology
3. Отраслевая терминология — industrial practice terminology
4. Понятийно четко очерченная терминология — well-developed ter- minology

В сопоставимом корпусе текстов соответствия между всеми термино- логическими словосочетаниями не обнаруживаются, однако в них всех прослеживается общая синтаксическая модель. Терминологические словосочетания, к которым эквиваленты не найдены, необходимо пере- вести, чтобы можно было из всех словосочетаний из сопоставимого списка составить единую семантическую структуру. Принимая, что ЛЕ *терминология (terminology) —* это гипероним и вершина в синтаксиче- ской иерархии, можно считать, что выделенные с помощью инструмен- тария AntConc словосочетания являются потенциальными гипонимами в семантической структуре, которые могут классифицироваться по различным признакам и объединяться между собой как гиперонимы. Иерархическая структура, в которую входят терминологические слово- сочетания, может быть использована для построения онтологий и сло- варных статей.

**Литература**

* 1. *Беляева Л.Н., Данилова О.А., Джепа Т.Л., Камшилова О.Н., Карнуп Е.В., Нымм В.Р., Чумилкин С.В* (2014), Лексикографический потенциал современных лингвистических технологий. СПб.: Книжный дом, 168 с.
  2. *Беляева Л.Н., Пиотровская К.Р.* (2014), Сетевой инструментарий филолога: учебное пособие. СПб.: Книжный дом, 48 с.
  3. *Беляева Л.Н.* (2014), Системы и процедуры выделения терминов из текстов. Терминология и знание: материалы IV Международного симпозиума. М.: Вест- Консалтинг, с. 212–223.

1. *Кудашев И.С.* (2007), Проектирование переводческих словарей специальной лексики. Helsinki.: Helsinki University Print, 443 с.
2. *Табанакова В.Д.* (2014), Переводчик-лингвист, переводчик-терминолог, переводчик-специалист: стратегия и тактика перевода термина. Вестник Тюмен- ского государственного университета (1), с. 72–81.
3. *Шеллов С.Д.* (1998), Определение терминов и понятийная структура тер- минологии. СПб.: Изд-во С.-Петерб. ун-та, 233 с.
4. *Delpech E., Daille B.* (2010), Dealing with lexicon acquired from comparable corpora: validation and exchange. Proceedings of 9th Conference on Terminology and Knowledge Engineering (TKE), pp. 229–223.
5. *Gillam L., Tariq M., Ahmad K.* (2005), Terminology and the construction of ontology, available at: [http://clara.b.uib.no/ﬁles/2011/06/Gillam-Tariq-and-Ahmad.pdf](http://clara.b.uib.no/%EF%AC%81les/2011/06/Gillam-Tariq-and-Ahmad.pdf) (дата обращения 28. 03. 2016).

**Literature**

1. *Belyaeva L.N., Danilova O.A., Dzhepa T.L., Kamshilova O.N., Karnup E.V., Nymm V.R., Chumilkin S.V* (2014), Leksikograﬁcheskiy potentsial sovremennykh lingvisticheskikh tekhnologiy. SPb.: Knizhnyi dom, 168 s.
2. *Belyaeva L.N., Piotrovskaya K.R*. (2014), Setevoy instrumentariy ﬁlologa: uchebnoe posobie. SPb.: Knizhnyi dom, 48 s.
3. *Belyaeva L.N.* (2014), Sistemy i protsedury vydeleniya terminov iz tekstov. Ter- minologiya i znanie: materialy IV Mezhdunarodnogo simpoziuma. M.: Vest-Konsalting, s. 212–223.
4. *Kudashev I.S*. (2007), Proektirovanie perevodcheskikh slovarey spetsial’noy leksiki. Helsinki.: Helsinki University Print, 443 s.
5. *Tabanakova V.D.* (2014), Perevodchik-lingvist, perevodchik-terminolog, perevodchik-spetsialist: strategiya i taktika perevoda termina. Vestnik Tyumenskogo gosudarstvennogo universiteta (1), s. 72–81.
6. *Shellov S.D*. (1998), Opredelenie terminov i ponyatiynaya struktura terminologii. SPb.: Izd-vo S.-Peterb. un-ta, 233 s.
7. *Delpech E., Daille B.* (2010), Dealing with lexicon acquired from comparable corpora: validation and exchange. Proceedings of 9th Conference on Terminology and Knowledge Engineering (TKE), pp. 229–223.
8. *Gillam L., Tariq M., Ahmad K.* (2005), Terminology and the construction of ontology, available at: [http://clara.b.uib.no/ﬁles/2011/06/Gillam-Tariq-and-Ahmad.pdf](http://clara.b.uib.no/%EF%AC%81les/2011/06/Gillam-Tariq-and-Ahmad.pdf) (data obrashcheniya 28. 03. 2016).

***В.Д. Табанакова, А.Ю. Левенкова***

**СЕМАНТИЧЕСКИЕ КРИТЕРИИ ОТБОРА ТЕРМИНОЛОГИЧЕСКИХ КОЛЛОКАЦИЙ В СЛОВАРЬ КОМБИНАТОРНОГО ТИПА**

**Аннотация**. В данной статье рассматриваются семантические критерии отбора терминологических коллокаций, выделенных из специальных текстов в области солнечной энергетики корпусным методом. В результате ранжирования терминологических коллокаций вводятся и обосновываются критерии термино- логичности данных единиц: критерий семантической отнесенности к сфере сол- нечной энергетики; критерий семантической целостности; логико-интуитивный критерий. Данные критерии признаются авторами решающими при формирова- нии итогового списка терминологических коллокаций с целью их семантизации в терминологическом словаре комбинаторного типа.

**Ключевые слова.** Терминологическое словосочетание, терминологическая коллокация, семантические критерии терминологичности, терминологический словарь комбинаторного типа.

***V.D. Tabanakova, A.Y. Levenkova***

**SEMANTIC CRITERIA OF SELECTING TERMINOLOGICAL COLLOCATIONS FOR A SPECIALIZED DICTIONARY**

**OF COLLOCATIONS**

**Abstract**. This article focuses on semantic criteria for selecting terminological col- locations extracted from specialized texts in the ﬁeld of solar energy using corpus-based methods. Grading of the terminological collocations allowed to introduce and justify the criteria of their termhood: semantic reference to the ﬁeld of solar energy; semantic integrity, and logical-and-intuitive criterion. The authors consider these criteria to be crucial for compiling a ﬁnal list of the terminological collocations with a view to their specifying in a specialized dictionary of collocations.

**Key words.** Multiword term**,** terminological collocation, semantic termhood cri- teria, specialized dictionary of collocations.

Терминологическое словосочетание как особая знаковая форма для обозначения специального понятия является одновременно единицей языка и единицей речи. При этом «сфера функционирования термина первична, сфера фиксации — вторична» [Лейчик 1994: 152]. В рамках проблемы отбора терминов в словарь функциональные параметры термина для нас имеют первостепенное значение, и предметом нашего исследова- ния становится терминологическое словосочетание в сфере речи.

© ТАБАНАКОВА Вера Дмитриевна. — Тabanakova Vera. Тюменский государственный университет, Россия. — Tyumen State University, Russia. E-mail: [veratab@inbox.ru](mailto:veratab@inbox.ru)

© ЛЕВЕНКОВА Анна Юрьевна. — Levenkova Anna. Тюменский государственный универ- ситет, Россия. — Tyumen State University, Russia. E-mail: [annapopova31@rambler.ru](mailto:annapopova31@rambler.ru)

В нашем исследовании терминологическое словосочетание в тексте предлагается называть «терминологической коллокацией». Под термино- логической коллокацией мы понимаем устойчивое, часто встречающееся сочетание слов в речи, формирование которого обусловлено понятийной системой области знания, в которой оно используется [Левенкова 2015: 1392]. В рамках статистического подхода коллокации — это характерные, часто встречающиеся сочетания слов, появление которых рядом друг с другом основывается на регулярном характере взаимного ожидания и задается не грамматическими, а чисто семантическими факторами [Хохлова 2008: 343].

Разница между терминологическим словосочетанием и терминологи- ческой коллокацией носит не семантический, а чисто функциональный характер. Терминологическая коллокация — это то же терминологическое словосочетание в сфере речи, характеризующееся высокой частотой встре- чаемости и не имеющее статуса термина в силу отсутствия дефиниции.

Сочетаемостные свойства лексических единиц традиционно рас- сматриваются в словарях комбинаторного типа. Опираясь на работы М.В. Влавацкой [Влавацкая 2009, 2012], мы определяем комбинаторный словарь как особый тип словаря, цель которого — представить лексиче- скую, семантическую и синтаксическую сочетаемость. Терминологиче- ская лексикография, в отличие от общей, не разрабатывает комбинаторные словари как самостоятельные виды словарей, поэтому терминологические словосочетания семантизируются в общих и специальных словарях в соот- ветствии с принципами описания общелитературной лексики. Разработка принципов описания терминологических словосочетаний в специальном комбинаторном словаре представляется отдельной задачей теоретической терминографии.

В данном исследовании мы останавливаемся на задаче отбора терми- нологических словосочетаний в специальный словарь комбинаторного типа и предлагаем процедуру установления семантических критериев отбора терминологических коллокаций в тексте.

Материалом исследования была выбрана стремительно развивающаяся терминология солнечной энергетики (СЭ), представленная главным образом многословными терминами как в сфере фиксации (стандарты и специальные словари), так и в сфере функционирования (научные статьи, тексты докладов, учебные пособия и т. д.). Анализ единиц в сфере функционирования проводился корпусным методом. Общее количество словоформ и словоупотреблений в корпусе составило 67 572 и 1 605 404 соответственно. В результате корпусного исследования, проведенного с помощью корпусного менеджера AntConc с использованием функций Wordlist и Clusters, было выделено 2473 коллокации. Коллокации отбирались по частотному принципу.

Отсутствие дефиниции у термина речи диктует определенные ограничения при описании его семантики. В связи с этим из-за невозможности проследить понятийную устойчивость терминологических коллокаций встает вопрос определения степени терминологичности номинаций такого рода. Учитывая данные ограничения, мы приняли решение распределить полученные коллокации на три группы. В первую группу вошли коллокации с высокой степенью терминологичности: словосочетания, зафиксированные в государственных стандартах и терминологических словарях по солнечной энергетике (*пассивное использование, перезаряжаемая энергетика, лучепоглощающая поверхность, производство биогаза, фотоэлектрическая установка и др.)*. Во вторую группу мы отнесли коллокации, которые потенциально могут быть терминами: словосочетания, в которых один из компонентов является узкоспециальным термином или в которых два и более компонентов являются межсистемными заимствованиями (*традиционная энергетика, работа электростанции, поверхность коллектора, ветер переменной скорости, производство водорода и др.)*. В третью группу вошли коллокации, которые сложно назвать терминами и к которым необходимы применение универсальных критериев терминологичности и даже оценка эксперта: словосочетания, полностью состоящие из слов общеупотребительного языка; семантически незаконченные сочетания, требующие смысловой развертки; многокомпонентные словосочетания с несколькими семантическими ядрами, сочетания с сочинительной связью между компонентами; словосочетания, не имеющие никакой ассоциативной связи с областью солнечной энергетики (*эффективность работы, поверхность воды, зеркальная поверхность, развитие региона, базовые материалы, проблемы технологии микроэлектроники и особо чистых материалов, жаркое солнце, дни без солнца*, *строение растворов электролитов литий-ионных источников тока и т. д).*

В результате распределения мы получили следующие данные. В пер-

вую группу вошли 417 коллокаций (**17%**), во вторую группу — 1866 еди- ниц (**75%**), и третья группа представлена словосочетаниями в количестве 190 (**8%**).

К третьей группе, которая оказалась самой малочисленной, мы решили применить критерии терминологичности с целью включения ее единиц в итоговый список для лексикографического описания.

Критерии терминологичности, выделяемые лингвистами, в полной мере можно использовать только для терминов сферы фиксации. Напри- мер, такие критерии, как дефиниционный критерий, критерий концепту- альной целостности, информационный критерий и критерий логических теорем, требуют обращения к дефиниции термина, чтобы соотнести знак со специальным понятием. Для описания терминов речи мы можем

использовать статистический и семантический критерий. Но в рамках нашего исследования данные критерии не в состоянии определить степень терминологичности коллокаций, так как изначально единицы отбирались по статистическим показателям из научно-технических текстов, посвящен- ных направлению солнечной энергетики. В связи с этим целесообразно выработать критерии, которые способны определить степень терминоло- гичности конкретных единиц речи.

Внимательно изучив характер коллокаций, вошедших в третью группу, мы посчитали нужным применение следующих критериев:

1. критерий семантической отнесенности к сфере СЭ (если коллока- ция не имеет никакой ассоциативной связи с областью СЭ и может быть использована в других терминосистемах, то данная коллокация не явля- ется терминологической коллокацией СЭ);
2. критерий семантической целостности (коллокация должна обла- дать смысловой законченностью и не должна требовать дальнейшей развертки);
3. логико-интуитивный критерий (соотнесенность коллокации с еди- ным, хотя и расчлененным понятием).

Данные критерии были применены к 190 коллокациям.

Применив вышеописанные критерии терминологичности к данному списку коллокаций, мы пришли к некоторым заключениям.

Абсолютное большинство коллокаций (123 единицы) не имеют тер- минологического характера, так как не отражают семантическую связь с областью солнечной энергетики и не обладают смысловой целостностью. Рассмотрим несколько примеров из данной группы коллокаций. Колло- кации «*эффективное использование», «особенности использования»,*

*«низкая эффективность»* не имеют ассоциативной связи с исследуемой предметной областью в силу отсутствия слов, которые непосредственно соотносят номинацию со специальным понятием. Коллокации подобного рода не обладают смысловой законченностью, так как требуют смысловой завершенности в виде объекта.

Ряд коллокаций (*районирование территории, холодное время, зави- симость от времени года),* хотя и обладают смысловой завершенностью, но с трудом могут быть отнесены к сфере солнечной энергетики.

В исследуемой группе коллокаций высокой степенью терминоло- гичности обладают многокомпонентные словосочетания с несколькими семантическими ядрами. Рассмотрим, например, коллокацию «*строение растворов электролитов литий-ионных источников тока».* В многоком- понентных терминологических коллокациях определить семантическое ядро довольно сложно, поскольку только дефиниция может ограничить объем понятия. Поэтому на уровне терминологической коллокации есть смысл говорить о нескольких семантических ядрах, которые являются

компонентами составного понятия, а формально являются главными сло- вами в словосочетаниях. Рассматриваемое нами словосочетание можно разбить на два смысловых отрезка: «*строение растворов электролитов*» и «*литий-ионные источники тока*», в которых выделяются два семантиче- ских ядра: «*строение*» и *«источники*». Несмотря на наличие нескольких семантических узлов, данное словосочетание обладает единой понятийной структурой.

О целостности понятийной структуры свидетельствует наличие под- чинительной связи между компонентами словосочетания, а также система зафиксированных в нем понятий. В политехническом словаре находим определение термина «*электролит*» — вещество, которое проводит электрический ток вследствие диссоциации на ионы, что происходит в растворах и расплавах, или движения ионов в кристаллических решет- ках твердых электролитов [Фединский 2007: 902]. Опираясь на данное определение, можно заключить, что понятие «*литий-ионные источники тока»* является видовым признаком родового понятия *«вещество*», что говорит о понятийной целостности всего словосочетания. Коллокации такого рода должны быть включены в словарь, так как они выражают профессиональное понятие, а сфера их функционирования ограничена областью солнечной энергетики.

Еще одна подгруппа терминов, требующая пояснения, представляет

собой словосочетания с сочинительной связью между компонентами. Приведем несколько примеров: *геотермальные и приливные станции, полезная нагрузка и энергия рабочего потока, энергосбережение и повы- шение энергетической эффективности.* Рассматривая словосочетания такого рода, возникает вопрос: выражают ли данные словосочетания одно понятие или несколько понятий? Ответ на данный вопрос можно найти, обратившись к самой природе коллокаций. Поскольку колло- кация — это речевая единица, не имеющая дефиниции, мы не можем рассматривать словосочетания с сочинительной связью как единые понятия. В связи с этим для включения данных единиц в словарь счи- таем целесообразным их разбиение на отдельные составляющие, выражающие единое понятие: *геотермальные станции, приливные станции, полезная нагрузка, энергия рабочего потока, повышение энергетической эффективности.*

Таким образом, на семантическом уровне терминологическая коллока-

ция может быть рассмотрена относительно степени терминологичности. Функциональные особенности коллокаций диктуют выделение особых критериев, которые смогли бы послужить дополнительным средством семантизации коллокаций в словаре. Предложенные нами критерии осно- ваны на универсальных критериях с проекцией на единицы речи в сфере солнечной энергетики.

**Литература**

1. *Влавацкая М.В*. (2009), Словари комбинаторного типа в современных типологиях и классификациях и критерии их выделения: Филологические науки // Вопросы теории и практики. Тамбов: Грамота, с. 80–87.
2. *Влавацкая М.В.* (2012), Лексикографическое описание сочетаемости слов. Модель построения англо-русского учебного комбинаторного словаря: LAPLAM- BERTAcademicPublishingGmbHandCo.KG, 189 c.
3. *Лейчик В.М*. (1994), Обоснование структуры термина как языкового знака понятия // Терминоведение. Вып. 2, с. 152.
4. *Левенкова А.Ю.* (2015)*,* Статус терминологической коллокации в теории сочетаемости: Научный альманах. Филологические науки, № 7(9), с. 1386–1393.
5. *Фединский Ю.И.* (2007), Большой нормативно-технический словарь. М.: Астрель, с. 902.
6. *Хохлова М.В*. (2008), Экспериментальная проверка методов выделения коллокаций. SlavicaHelsengiensia 34. Инструментарий русистики: Корпусные подходы. Хельсинки, с. 343.

**Literature**

1. *Vlavatskaya M.V.* (2009), Slovari kombinatornogo tipa v sovremennykh tipologi- yakh i klassiﬁkatsiyakh i kriterii ikh vydeleniya: Filologicheskie nauki // Voprosy teorii i praktiki. Tambov: Gramota, s. 80–87.
2. *Vlavatskaya M.V.* (2012), Leksikograﬁcheskoe opisanie sochetaemosti slov. Model' postroeniya anglo-russkogo uchebnogo kombinatornogo slovarya: LAPLAM- BERTAcademicPublishingGmbHandCo.KG, 189 s.
3. *Leychik V.M.* (1994), Obosnovanie struktury termina kak yazykovogo znaka ponyatiya // Terminovedenie. Vyp. 2, s. 152.
4. *Levenkova A.Yu*. (2015), Status terminologicheskoy kollokatsii v teorii socheta- emosti: Nauchnyi al’manakh. Filologicheskie nauki, № 7(9), s. 1386–1393.
5. *Fedinskiy Yu.I*. (2007), Bol’shoy normativno-tekhnicheskiy slovar’. M.: Astrel’, s. 902.
6. *Khokhlova M.V*. (2008), Eksperimental’naya proverka metodov vydeleniya kollokatsiy. SlavicaHelsengiensia 34. Instrumentariy rusistiki: Korpusnye podkhody. Khel’sinki, s. 343.

***А.Н. Трофимов***

**ГРАММАТИЧЕСКАЯ И КАТЕГОРИАЛЬНАЯ ПРЕДИКАТИВНОСТЬ ГЛАГОЛА**

**В ТЕКСТЕ РУССКОЙ НАУЧНОЙ АННОТАЦИИ**

**Аннотация.** В статье рассматриваются грамматические категории русского глагола — вид и возвратность с точки зрения их грамматического и предикатив- ного значений. В ходе категориального анализа автор приходит к выводу о том, что, помимо собственно грамматических значений, данные категории несут модальную и предикативную нагрузку высказывания. Были выделены такие семантические оппозиции, как уверенность/неуверенность и согласие/несогласие. Это означает, что предикация в тексте научной аннотации выражается не только лексическим и грамматическим значением глагола, но также и его категориальным значением. Категориальное значение русского глагола ложится в основу предика- тивной модели перевода научной аннотации с русского на английский.

**Ключевые слова.** Предикация, глагол, грамматическая категория, модаль- ность, предикативная модель перевода.

***A.N. Troﬁmov***

**GRAMMATICAL AND CATEGORICAL PREDICATION OF A VERB IN THE TEXT OF A RUSSIAN**

**SCIENTIFIC ABSTRACT**

**Abstract.** This article deals with some grammatical categories of a Russian verb such as aspect and reﬂexivity according to their grammatical and predicative mean- ing. After categorical analysis the author comes to the conclusion that along with their grammatical meaning these categories convey modal and predicative meaning. For example, such semantic oppositions have been observed: certainty\uncertainty, agree- ment\disagreement. This means that scientiﬁc predication can be expressed with the help of categorical meaning of a verb, which can be used in the predicative model for translation of scientiﬁc abstracts from Russian into English.

**Key words**. Predication, verb, grammatical category, modality, predicative transla- tion model.

Классическое представление о предикации сводится к точке зрения, что предикат — это глагол в чистом виде. И вообще предикат для мно- гих — синоним слова «сказуемое», и мы обычно ассоциируем это слово именно с глаголом, не принимая во внимание незначительные отклоне- ния от нормы. Безусловно, это представление оказывается ошибочным, когда мы вспоминаем, что сказуемое может выражаться различными частями речи: от существительного до числительного и прилагательного

© ТРОФИМОВ Алексей Николаевич. — Troﬁmov Aleksei. Тюменский государственный университет, Россия. — Tyumen State University, Russia. E-mail: [lextrif@mail.ru](mailto:lextrif@mail.ru)

(составное именное сказуемое). Если же взглянуть на простое глагольное сказуемое с точки зрения семантики, то глагол в нем, в принципе, ничем не отличается от имени существительного или другой номинативной единицы, только он называет действие, а не предмет. Эти примитивные взгляды на предикацию не имеют никакого значения в современном мире, они устарели и не могут быть объективными. Их опровергал еще Ю.С. Степанов в своих трудах 1980-х гг. В качестве аргументации свей точки зрения он приводит слова Гегеля из «Науки логики», согласно которым номинация противопоставляется предикации как «единичное» к «всеобщему» [Степанов 1980: 311]. Автор доказывает эту точку зре- ния на примере двух простых предложений: Учителя много работают и Мои родители *—* учителя. В первом случае учителя является субъектом и выступает в роли номинативной единицы, во втором — предикатом, но семантическая сущность данной единицы в обоих примерах является абсолютно разной. Главным отличием можно считать то, что в первом при- мере учителя — это денотат, обозначающий целый класс людей, объеди- ненных родом деятельности, а во втором — сигнификат, обозначающий свойства референта-денотата. Таким образом, предикаты стоит четко противопоставлять субъектам, оставляя неоспоримым мнение о равной значимости этих компонентов высказывания.

Центральными единицами нашего исследования являются именно

предикаты, которые представляют собой сложную и порой непредсказуе- мую систему значений, играющую колоссальную роль в формировании высказывания и реализации его коммуникативной функции. Предика- ция в научном дискурсе, а именно в научных аннотациях, — еще более противоречивый вопрос, так как именно текст аннотации характеризуется высокой степенью компрессии и имплицитности излагаемой информации. Глагол совмещает в себе не только предикативное значение, но и ряд других сем, связанных с субъектом и объектом высказывания, а также временно-пространственными аспектами ситуации. Ю.С. Степанов под- тверждает эту точку зрения мнением, что глагол только частично является предикатом, обладая при этом иными оттенками значения и имея связь с другими компонентами высказывания [Степанов 1980: 312].

Мы все же не будем опровергать точку зрения, что зачастую преди- кат — это глагол. Наоборот, мы считаем, что глагол вносит наибольший вклад в предикативное значение высказывания и имеет больше средств выражения предикативного и модального значений. В данной статье под глаголом мы понимаем любую личную форму глагола, использованную в тексте научной аннотации.

Глагол — явление многогранное и объединяет в себе значение преди- ката и несколько других значений: категориальное, модальное, временное и др. Ю.С. Степанов утверждал, что глагол — поверхностная форма пре-

диката и «чего-то еще» [Степанов 1980: 236]. Загадочная формулировка

«чего-то еще» заставляет нас задуматься, что же еще может выражать глагол. Возможно, его внутренняя грамматическая структура обладает чем-то большим, чем категориальным, временным или аспектным значе- ниями? Д.В. Савова также отмечала, что глагол-предикат стоит во главе предикативной структуры высказывания, обладая широчайшим спектром предикативных значений, которые выражаются его семантикой, прагма- тикой и грамматикой [Савова 2012: 70].

Сфокусируемся на двух типично-глагольных категориях: вид и воз- вратность. На примере оппозиций соверешенность/несовершенность, возвратность/невозвратность мы попытаемся выделить дополнительные предикативные значения, которые эти категории могут выражать. Обратим внимание на то, что мы рассматриваем глагол в тексте научной аннотации, который, как отмечалось выше, является текстом особого типа, характери- зующимся высокой степенью сжатости и предикативности. Наше мнение перекликается с выводами Ю.Д. Апресяна, который в своих трудах «Лек- сическая семантика» писал, что научный текст (в нашем случае — текст научной аннотации) нуждается в предикации больше других типов текста и выражает ее всеми доступными языковыми средствами [Апресян 1995: 134].

Мы проанализировали более 50 русских научных аннотаций, опи- сывающих исследования в различных сферах науки: лингвистика, журналистика, математика, экономика, физика, история и философия науки, машиностроение, информационные технологии, химия, биология, медицина. Статьи и их аннотации, использованные в исследовании, были опубликованы в ведущих научных журналах России. В результате анализа синтаксической и предикативной структуры предложений в аннотациях мы пришли к выводу, что предикативность и модальность в таких текстах почти не выражается посредством вводных и модальных слов, она заклю- чена в самой предикативной структуре, которая, в свою очередь, зачастую представляет собой простую ключевую структуру «субъект — предикат — объект», предикат в которой в большинстве случаев (что так же очень примечательно) является глаголом. Наша задача — пристальнее изучить все глаголы-предикаты выбранных аннотаций и попытаться помимо чисто грамматических и категориальных значений выделить то самое «что-то еще», отмеченное Ю.С. Степановым и многими другими лингвистами.

Рассмотрим два предложения: *1) Мы делаем вывод об отсутствии*

*системных связей <…> при таком методе обработки данных. 2)Мы пришли к выводу, что термин определено и зачастую одновременно входит в несколько терминосистем <…>.* С первого взгляда может показаться, что оба этих предложения будут иметь одну и ту же прагма- тическую задачу: сообщить читателю о выводе, сделанном в результате

научного исследования. Различаются только использованные сказуемые, а вернее, их грамматическая форма: несовершенный вид/совершенный вид. Именно это отличие отражает не случайность выбора автором катего- рии. Обратимся ко всему предложению и контексту. В первом случае автор сначала приводит один вывод, а далее по тексту противопоставляет его с другим более рациональным и правильным выводом, тем самым указы- вая на ложность первого. Второе предложение согласно аннотации и самой статье описывает изначально верно сформулированный вывод. Все эти наблюдения сводятся к тому, что несовершенный вид глагола в научной аннотации выражает неуверенность автора или ложность высказывания, когда совершенный вид глагола как бы констатирует правдивость той или иной мысли. В данном случае мы имеем дело с первой семантической оппозицией «уверенность/неуверенность». Примечательно, что глаголы несовершенного вида превалируют среди предикатов в рассмотренных нами аннотациях, несмотря на то что такие тексты направлены преиму- щественно на описание результатов. Но, ознакомившись со статьями, мы пришли к выводу, что зачастую статьи и их аннотации посвящены обзорным, экспериментальным исследованиям, где результат работы либо только предполагается, либо его вовсе нет, или же в работах описывается только процесс исследования.

Первый вывод касательно категорий глагола как носителей модально-

предикативного значения заключается в том, что категория вида глагола необходима для передачи семантических значений процессуальности/ сверешенности и модально-предикативных сем уверенности/неуверен- ности.

Категория возвратности, как и категория вида, вызывает наш интерес с точки зрения наличия в ней предикативного значения. Мы склонны считать, что использование возвратных и невозвратных глаголов может быть обусловлено задачами автора, посылом и основной идеей, которую он хочет донести до референта.

Проследим, влияет ли выбор формы глагола относительно категории возвратности на общий смысл высказываний, а также на их предикативное наполнение. В качестве примеров возьмем одно предложение, в котором употреблен возвратный глагол, и второе, в котором использован невоз- вратный глагол: *1) Данное предположение подтверждается статисти- ческими данными; 2) Также автор подтверждает выводы М.И. Романова относительно наличия в таких органах стволовых клеток.* Оба выска- зывания имеют приблизительно одинаковую коммуникативную задачу: доказать правдивость каких-либо данных, результатов, выводов. Нас же интересует вопрос выбора формы глагола: чем руководствуется автор, когда осуществляет выбор между возвратной и невозвратной формой глагола. Рассуждая поверхностно, можно предположить, что категория

возвратности зависит от контекста и может быть использована во избе- жание повторения структуры предложения для создания синонимической пропозиции без нежелательного совпадения структуры высказывания. Д.В. Савова, размышляя о синтаксических синонимах, пишет, что дей- ствие, которые участвует в неком референтом событии, сохраняет свою валентность при создании синонимической пропозиции. Автор приводит пример, в котором предикативное значение не меняется при изменении структуры предложения и категориальной формы глагола: *Пища гото- вится здесь — Здесь готовится пища* (2). Это замечание бесспорно объективно. Однако данные примеры являются синтаксическими сино- нимами исходного предложения, приведенного автором: *Здесь повариха готовит пищу* (1) [Савова 2012: 71]. Если сравнивать это высказывание с его синонимами, становится очевидна разница между ними: в первом случае субъект выражен семантически окрашенным словом «повариха», которое, безусловно, обладает конотативным значением, позволяющим автору сконцентрировать внимание реципиента на этом слове; во втором случае предложение также имеет субъект, но уже неодушевленный, вслед- ствие чего высказывание приобретает безличность и логический акцент переходит на само действие. Значит ли это, что категория переходности отвечает только за тема-рематическое членение высказывания?

Вернемся к нашим примерам: *1) Данное предположение подтверж-*

*дается статистическими данными; 2) Также автор подтверждает выводы М.И. Романова относительно наличия в таких органах стволовых клеток.* Выбор возвратной формы в первом предложении также обуслов- лен необходимостью задать рему высказывания, но стоит заметить, что во втором предложении тот же глагол в невозвратной форме обладает боль- шей категоричностью и некой «одобрительностью» со стороны автора. Иными словами, безличная возвратная форма глагола отстраняет автора от предикации и ликвидирует любую модальность, когда слово «автор» в качестве субъекта в сочетании с невозвратной формой глагола выражают категоричность и добавляют сему «согласие» к предикативному значению пропозиции. Следовательно, категория возвратности представляет собой предикативную оппозицию согласия/несогласия.

Изучив категории вида и возвратности глагола с точки зрения преди- кативности их значений, мы можем сделать вывод, что они обладают экс- тралингвистическими модально-предикативными компонентами значения, представленными нами в виде оппозиций уверенность/неуверенность, согласие/несогласие. Данные результаты указывают на то, что предикатив- ная нагрузка высказывания строится не только на основе синтаксической структуры предложения и лексических средств языка, но и может быть выражена грамматическими формами глаголов-предикатов, а возможно, и других членов предложения. Во-первых, такое заключение дает почву

для дальнейших исследований в этом направлении, ведь категориальная предикативность, как и научная предикативность в целом, — сферы малоизученные и неразработанные. Во-вторых, предикация, выраженная категориальным значением, может послужить одним из компонентов пре- дикативной модели перевода русской научной аннотации на английский язык, которая нацелена на наиболее точную передачу предикативного значения высказывания при переводе.

**Список литературы**

1. *Апресян Ю.Д.* (1995), Лексическая семантика. Синонимические средства языка. Язык. Семиотика. Культура, 364 с.
2. *Савова Д.В.* (2012), Предикативная структура высказывания в болгарском языке. Вопросы языкознания, № 3, c. 69–90.
3. *Степанов В.С.* (1980), К универсальной классификации предикатов. Изве- стия АН СССР. Серия литературы языка, 190, Т. 39, № 4, c. 311–323.

**Literature**

1. *Apresyan Yu.D*. (1995), Leksicheskaya semantika. Sinonimicheskie sredstva yazyka. Yazyk. Semiotika. Kul’tura, 364 s.
2. *Savova D.V*. (2012), Predikativnaya struktura vyskazyvaniya v bolgarskom yazyke. Voprosy yazykoznaniya, № 3, s. 69–90.
3. *Stepanov V.S*. (1980), K universal’noy klassifikatsii predikatov. Izvestiya AN SSSR. Seriya literatury yazyka, 190, T. 39, № 4, s. 311–323.

***К.Р. Турганова***

**ВЫБОР ВАРИАНТНОГО ПЕРЕВОДНОГО СООТВЕТСВИЯ С ПОМОЩЬЮ КОРПУСНОГО МЕНЕДЖЕРА BNCWEB**

**Аннотация.** В статье описана роль корпусных менеджеров на этапе вычлене- ния термина в контексте и подбора к нему переводного соответствия. При помощи функции «сравнения» корпусного менеджера Британского национального корпуса устанавливаются принципы отбора терминологических словосочетаний в техни- ческом тексте по тематике «капитальный ремонт скважин». Логико-понятийный анализ терминологических словосочетаний позволяет обосновать выбор перевод- ного соответствия.

**Ключевые слова.** Переводное соответствие, вариант перевода, корпус тек- стов, двухкомпонентные терминологические словосочетания.

***K.R. Turganova***

**VARIANT TRANSLATION EQUIVALENTS CHOICE USING BNCWEB CORPUS MANAGER**

**Abstract.** The article describes corpus manager’s role at the stage of term extraction in the context and translation equivalents matching. Terminological word combinations choice principles in technical texts on the topic “wells workover” are deﬁned using the function “comparison” British National Corpus Manager. Logical-conceptual analyze of term word combinations determines the choice of translation equivalent.

**Keywords.** Translation equivalents, translation variant, corpus, bicomponent ter- minological word combinations.

Нефтегазовая отрасль, достаточно обширная, включающая в себя такие крупные направления, как разведка, добыча и транспорт углеводо- родного сырья, делится, в свою очередь, на различные подразделы: гео- физические методы исследования месторождений, капитальный ремонт скважин, сбор и подготовка нефти и газа, переработка углеводородов. Особенности данной отрасли не могли не отразиться на структуре ее терминосистемы. Термины, используемые в нефтегазовой отрасли, отли- чаются неоднородностью, самобытностью, так как нефтегазовая сфера формировалась достаточно хаотично, временами форсировано, находясь в прямой зависимости от погодных условий, природных ресурсов. Осо- бого внимания требует перевод технической документации, инструкций, руководств, стандартов, так как неточно подобранная единица текста перевода, эквивалентная единице текста оригинала, может привести к воз- можным экономическим убыткам, аварийному состоянию объекта и, как следствие, к катастрофе.

© ТУРГАНОВА Камила Рашидовна. — Turganova Kamila. Тюменский индустриальный уни- верситет, Россия. — Industrial University of Tyumen, Russia. E-mail: [ktur](mailto:kturganova@gmail.com)[ganova@gmail.com](mailto:ganova@gmail.com)

В данной статье рассматриваются возможности корпусного менеджера XAIRA Британского национального корпуса в процессе выбора пере- водного соответствия терминологического словосочетания подотрасли нефтяного дела «Капитальный ремонт скважин».

Материалом исследования послужили двухкомпонентные термино- логические словосочетания, отобранные методом сплошной выборки из текстов инструкций, договоров, международных стандартов, по теме

«Капитальный ремонт скважин», предложенные Международным центром повышения квалификации бурильщиков для перевода в период с октября 2015 г. по март 2016 г.

В процессе предпереводческого анализа было выделено 326 терми- нологических единиц, из которых количество однокомпонентных тер- минов составило 62 терминологические единицы (19% от общего числа терминов), трехкомпонентных — 19 терминологических единиц (6%) и подавляющее большинство составили двухкомпонентные термины — 203 терминологических единиц (62%).

Количественная пропорция терминов не противоречит результатам исследования, полученным другими лингвистами в ходе анализа раз- личных терминосистем [Гринев 1993: 142–150; Додонова 2000: 47–48;

Беляева 2007: 89–90].

Исследование проводилось в несколько этапов. Поиск переводных соответствий осуществлялся следующим образом: на первом этапе были использованы электронные тематические русско-английский и англо- русский словари нефтегазовых терминов, находящиеся в свободном доступе в [интернете:http://ruseng.com/anglo\_russkiy\_slovar\_neftegazovoy\_](http://ruseng.com/anglo_russkiy_slovar_neftegazovoy_) promyishlennosti/, [http://slovar-vocab.com/english-russian/big-oil-gas-vocab.](http://slovar-vocab.com/english-russian/big-oil-gas-vocab) html, также был использован словарь общей лексики, наиболее объемный электронный словарь — Мультитран [www.multitran.ru.](http://www.multitran.ru/) На этом этапе нашей задачей являлся подбор эквивалентных соответствий для выбран- ных 203 терминологических словосочетаний при помощи данных сло- варей. Я.И. Рецкер в учебном пособии по переводу с английского языка на русский трактует эквиваленты как соответствия между словами двух языков, которые являются постоянными, равнозначащими. В процессе работы были подобраны эквивалентные соответствия для более чем поло- вины двухкомпонентных терминов, однако для 47 терминологических еди- ниц (23%) эквиваленты в электронных словарях не были выявлены. Этот список составили следующие терминологические единицы: *ball diverter, ball dropper, bed wrap, belt effect, bleedoff line, chemical diversion, chrome tubing, counterbalance winch, electrical coupon, equalizing loop, ﬁshing diagram, foam diversion, frac balls, frac crew, frac gel, gripper blocks, heavy pipe, hydraulic bypass, hydraulic centralizer, kill pump, pumping schedule, residual bend, rod elevator, rod unit, sand lifting, shear stock, slugging pill,*

*snubbing basket, snubbing force, snubbing jack, squeeze manifold, squeeze tool, stationary slips, stationary snubbers, stimulation byproduct, liner punch, live cement, mast unit, mechanical diversion, neutralizing solution, nominal ﬁlter, ovality limit, workover string, travelling slips, travelling snubbers, tubular jar, pipe heavy.*

Электронные словари предоставляют различные возможности, позво- ляющие облегчить работу с терминами, например: функции запоминания и гипертекст, различные виды поиска, а также грамматическая характе- ристика слов. Тем не менее в ряде случаев было невозможно подобрать эквивалент двухкомпонентного термина, особенно в тех случаях, когда ядро термина являлось словом общего языка, например *ball diverter* или *bed wrap*. Именно потребность в дополнительной информации обусло- вила использование корпусного менеджера Британского национального корпуса.

На втором этапе исследования были использованы функции данного корпуса текстов, в попытке найти вариантные соответствия для пере- численных выше терминологических единиц. Согласно характеристике структуры двухкомпонентных терминов, данной Р. Стреловым, в его работе Стандартизация технической терминологии: принципы и при- менение [Стрелов 1983: 26–34]: «Степень автономности понятий, репре- зентируемых каждым из компонентов двухкомпонентной модели, и место репрезентируемых понятий в общей онтологии понятий конкретной терминосистемы является приоритетной. Именно автономность понятий, просматриваемая в каждом из компонентов, представляет наибольшую трудность при выборе переводного соответствия». Если эквиваленты являются постоянными, равнозначащими и, как правило, не завися- щими от контекста, поскольку эквивалентное соответствие всегда одно [Рецкер 1974: 10–11; Швейцер 1988: 76], то для поиска вариантного соответствия важную роль играет контекст. Информация о контексте содержится в корпусах текстов. Рассмотрим отличительные характе- ристики Британского национального корпуса. При помощи корпусных менеджеров можно получить данные о частотности и коллокатах интере- сующего переводчика термина или терминологического словосочетания и при выборе переводного соответствия ориентироваться на данные показатели (частотность употребления в жанрах и поджанрах, контекст). Британский национальный корпус — один из первых национальных корпусов, по образцу которого создавались многие современные корпуса различных языков, поставляемый с программным обеспечением Xaira. Британский национальный корпус (BNC от *англ*. British National Corpus)

— это корпус текстов из 100 миллионов слов, содержащий образцы

письменного и разговорного британского английского языка из широкого круга источников. Корпус охватывает британский английский конца XX в.,

представленный широким разнообразием жанров, и задуман как образец типичного разговорного и письменного британского английского языка того времени.

На третьем этапе исследования мы приступили к переводу терминоло- гических единиц с русского языка на английский, также используя функ- ции Британского национального корпуса. На примере терминологического словосочетания “*oil drilling*” мы рассмотрим основные функции BNC. Корпус имеет разметку в соответствии с рекомендациями консорциума Text Encoding Initiative и включает полную лингвистическую аннотацию и контекстную информацию. Те же результаты можно получить в виде таблицы, выбрав соответствующую функцию: «отображение в виде таблицы». В результате на экране появится сравнительная таблица, показывающая частотность употребления данного словосочетания в различных секциях корпуса (в различных жанрах). Также возможно рассмотреть частотность употребления слов или словосочетаний по поджанрам. При необходимости можно также рассмотреть контекст, кликнув на соответствующую кнопку. В данной статье приведен лишь фрагмент данной таблицы.

Следующая функция корпусного менеджера — это возможность рас- смотрения списка коллокатов — слов, которые сочетаются с искомым словосочетанием, данная функция позволяет легко разобраться в значении исходного слова (или словосочетания), а также в его использовании.

Рассмотрим еще одну функцию Британского национального кор- пуса — возможность анализа списка конкордансов. Выбрав данную функцию, можно посмотреть список всех примеров, с которыми данное словосочетание употребляется.

Функция «сравнение слов/словосочетаний» Британского нацио- нального корпуса помогла нам при выполнении перевода с русского языка на английский при выборе варианта перевода термина. Остано- вимся на семантической структуре терминологических словосочетаний. В двухкомпонентных и многокомпонентных терминах обычно выделяют семантическое ядро, как правило, субстантивное, это можно видеть на примерах терминологических словосочетаний, полученных методом сплошной выборки из текстов, которые предстояло перевести: *ACID (acid inhibitor, acid tank, acetic acid, acid job, acid stimulation, acid job, hydroﬂuoric hydrochloric acid), PRESSURE (maximum treating pressure, maximum allowable pressure, pressure sender).*

Благодаря сравнению коллокатов двух слов или словосочетаний, можно разглядеть их тонкие смысловые различия и случаи употребления. Для сравнения мы взяли словосочетание “*oil drilling*” (WORD 1) и “*oil development*” (WORD 2).

*Таблица***.** Частота словоупотреблений терминологических словосочетаний

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | WORD | W1 | W2 | W1/W2 | SCORE |
| 1 | OIL | 27 | 7 | 3.9 | 1.0 |
| 11 | ARCTIC | 3 | 0 | 6.0 | 1.6 |
| 12 | MINING | 3 | 1 | 3.0 | 0.8 |
| 13 | IS | 3 | 1 | 3.0 | 0.8 |
| 14 | ON | 3 | 2 | 1.5 | 0.4 |

Как видно из табл., словосочетание “*oil drilling*” гораздо более употре- бительное (W1-27 словоупотреблений) и соответственно в определенном ряде случаев именно этот вариант подходит в качестве перевода словосо- четания «*бурение нефти*» на английский язык. По мнению Г.В. Чернова, существуют следующие причины вариативности: 1) многозначность лек- сических единиц и грамматических форм; 2) многозначность сочетаний слов во внетекстовом высказывании (если только речь не идет о посло- вицах и афоризмах; 3) отсутствие строгих критериев минимального (необходимого и достаточного) контекста, который необходим для снятия многозначности; 4) неоднозначность понимания связана также с неодно- значностью когнитивно-тезаурусного вывода (разными получателями текста, следовательно, разными переводчиками делаются разные выводы, они зависят от тезауруса знаний переводчика); 5) неоднозначность понимания, связанного с ситуативной неопределенностью (отсутствие наглядно-чувственного представления о ситуации, которая описывается в данном высказывании). Вывод об описываемой предметной ситуации является результатом взаимодействия тезауруса знаний о мире (следова- тельно, о типичных ситуациях) и предметно чувственного представления об описываемой ситуации. [Чернов 1988].

Перевод словосочетания «*бурение нефти*» при помощи корпус-

менеджера показал, что при использовании подобных программ причины вариативности могут быть сняты: наличие информации о контексте, наглядные примеры коллокатов позволяют осуществлять логико- понятийный анализ в каждом отдельно взятом случае и таким образом повысить точность перевода терминов и терминологических словосоче- таний.

В ходе исследования выяснилось, что корпус текстов, в отличие от электронных словарей, обладает большим количеством функций, что способствует повышению качества перевода, проверив наличие в корпусе 47 терминологических единиц, в корпусе текстов мы обнаружили 40. Таким образом, 7 терминологических единиц — это менее 5% (а именно 3,4%) от общего числа терминов (203), что в 4 раза меньше того, что мы не смогли обнаружить в электронных словарях — у 47 двухкомпонентных терминов (23%) не было найдено эквивалента на русском языке. Данные

исследования подтвердили эффективность Британского корпусного менед- жера по сравнению электронными словарями, находящимися в свободном доступе в интернете.

В заключение хочется отметить, что изучение корпусов позволяет получать точные данные о лексическом составе языков, об относительных частотах употребления тех или иных слов.

**Литература**

1. *Беляева Л.Н.* (2007), Специальные тексты в аспекте машинного и ручного перевода: сб. докл. междунар. науч-практ. конф. М.
2. *Додонова Н.Э.* (2000), Англо-русские соответствия в акцентно-ритмической реализации многокомпонентных терминов: дис. … канд. филол. наук. Пятигорск.
3. *Гринёв С.В.* (1993), Введение в терминоведение. М. 309 с.
4. *Pецкер Я.И.* (1974), Теория перевода и переводческая практика. М. 216 с.
5. *Чернов Г.В.* (1988), Фильтры на входе и выходе и множественность вари- антов перевода. М.
6. *Швейцер А.Д*. (1988), Теория перевода: статус, проблемы, аспекты. М. 215 с.

**Literature**

1. *Belyaeva L.N.* (2007), Spetsial’nye teksty v aspekte mashinnogo i ruchnogo perevoda: sb. dokl. mezhdunar. nauch-prakt. konf. M.
2. *Dodonova N.E*. (2000), Anglo-russkie sootvetstviya v aktsentno-ritmicheskoy realizatsii mnogokomponentnykh terminov: dis. … kand. ﬁlol. nauk. Pyatigorsk.
3. *Grinyov S.V*. (1993), Vvedenie v terminovedenie. M. 309 s.
4. *Petsker Ya.I*. (1974), Teoriya perevoda i perevodcheskaya praktika. M. 216 s.
5. *Chernov G.V.* (1988), Fil'try na vkhode i vykhode i mnozhestvennost' variantov perevoda. M.
6. *Shveytser A.D.* (1988), Teoriya perevoda: status, problemy, aspekty. M. 215 s.

***С.В. Чебанов***

**ЛИБЕРАЛИЗМ ТЕРМИНОЛОГИИ И СТАНДАРТИЗАЦИЯ НОМЕНКЛАТУРЫ КАК СРЕДСТВА ОБЕСПЕЧЕНИЯ РАЗВИТИЯ НАУЧНОГО ТВОРЧЕСТВА**

**Аннотация.** Обсуждая вопрос о стандартизации специальной лексики, необ- ходимо принимать во внимание непредсказуемость развития любой предметной области и различать фундаментальные изыскания и прикладные разработки; научные школы и научные партии; термины, номены и онимы. Стандартизовать номены можно с помощью универсальных номенклатурных типов, а иллюстри- ровать смысл терминов — с помощью эталонов, выделяемых конкретными исследователями. Оперирование с типами и эталонами обеспечивает свободу научного творчества при сохранение референциального значения используемой специальной лексики.

**Ключевые слова.** Фундаментальные изыскания, прикладные разработки, научные школы, научные партии, термин, номен, оним, номенклатурный тип, эталон, референциальное значение.

***S.V. Chebanov***

**LIBERALISM OF TERMINOLOGY AND STANDARDIZATION OF THE NOMENCLATURE AS MEANS OF ENSURING**

**OF DEVELOPMENT OF SCIENTIFIC CREATIVITY**

**Abstract.** Discussing a question of standardization of special lexicon it is necessary to take into account the unpredictability of development of any data subject area and to distinguish fundamental researches and application-oriented development applied research; scientiﬁc schools and scientiﬁc parties; terms, nomena and onima. It is possible to standardize nomena by means of the universal nomenclature types, and to illustrate a sense of terms — by means of standards selected by speciﬁc researchers. Operation with types and standards provides freedom of scientiﬁc creativity when saving referential value of the used special lexicon.

**Keywords.** Fundamental research, applied research, scientiﬁc schools, scientiﬁc party, term, nomen, onim, nomenclature type, standard, referential value.

Проблема стандартизации специальной лексики давно и интенсивно обсуждается терминологами и специалистами по смежным областям. Тем не менее, как представляется, в этих дискуссиях не всегда четко различены особенности персональных точек зрения специалистов и специфика тех предметных областей, специальная лексика которых является предметом обсуждения. Осуществление такого различения и является предметом настоящей работы.

© ЧЕБАНОВ Сергей Викторович. — Chebanov Sergey. Санкт-Петербургский государственный университет, Россия. Saint-Petersburg State University (Russia). E-mail: [s.chebanov@gmail.com](mailto:s.chebanov@gmail.com)

Во-первых, нужно принять во внимание, что развитие любой пред- метной области есть процесс непредсказуемый и малоуправляемый. Поэтому говорить о каком-либо нормировании лексики в области фунда- ментальных наук, основополагающих сочинениях и устных дискуссиях по проблемам техники, работах по искусствоведению или политике просто бессмысленно (если только не иметь в виду сознательное сдерживание их развития). При этом будут существовать различия в лексике специалистов разных стран, школ (ср. Московскую и Петербургскую фононологиче- ские школы, московскую и петербургскую систематику бескилевых птиц и т. д.), наконец, индивидуальных идиолектов (последние наиболее ярко выражены у философов). Любые нормативные ограничения (в том числе связанные с терминологией) в этой области неприемлемы и приводят к остановке развития той или иной сферы деятельности (преждевремен- ные рефлексия и нормирование просто преступны с точки зрения эври- стики). Хотя различение нормативной и ненормативной лексикографии, в том числе терминологической и номенклатурной, хорошо известно, антагонизм нормативности и эвристичности не всегда получает заслужи- вающее того внимание.

Во-вторых, существует значительное различие областей фундамен-

тальных исследований и разработок (включая изобретательскую деятель- ность) и прикладных разработок с точки зрения допустимости нормирова- ния, которое в большей мере нужно для прикладных сфер и должно быть минимизировано для фундаментальных изысканий.

В-третьих, необходимо четко различать научные (технические, научно- технические, художественные, философские) школы и научные (и прочие) партии. Отличительной чертой школы является постоянное развитие идей, причем такое развитие, которое может включать в себя и опровержение каких-то представлений предшествующего этапа разработок. Нередко лидером таких разработок является сам глава школы, некоторые ученики которого могут рвать со своим учителем из-за того, что они оказываются более консервативными, чем их наставник. С точки зрения представлений Т. Куна о парадигме, научная школа для того, чтобы оставаться школой, все время должна разрушать ту парадигму, которую создает. Задачей же партии является отстаивание своих идей, причем порою любыми сред- ствами и любой ценой. Вместе с тем нередко получается, что научная школа в целом или какие-то ее представители начинают действовать, следуя логике партии, причем такую подмену не замечают как изнутри сами участники событий, так и внешние наблюдатели.

Учитывая сказанное, можно утверждать, что отечественная школа тер-

миноведения, идущая от Д.С. Лотте, ориентирована прежде всего на нор- мативную [Гринёв-Гриневич 2008], прикладную (учебную, справочную) и фактически партийную лексикографию (если иметь в виду апелляцию

к нормативным документам, стандартам и т. д., которые носят отраслевой или общегосударственный характер, но создаются ограниченной группой конкретных людей, борющихся за утверждение и распространение своих мнений).

В интенсивно же развивающихся новых предметных областях, тем более фундаментальных, а не прикладных, плодотворны именно научные школы, теряющие плодотворность с превращением в научные партии.

Описательное естествознание (биолого-палеонтологическая система- тика, стратиграфия) показывает, что терминоведение неполно рефлекти- рует свое приложение, порождая дискуссии вместо требуемого размеже- вания. Начать последнее целесообразно с различения терминов, номенов и онимов (Чебанов, Мартыненко 1999).

*Таблица.* Сопоставление трех классов специальной лексики

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Аспект различия** | | **Термин** | **Номен** | **Оним** |
| Тип языка1 | | описательный | назывательный | – |
| Роль мотивировки | | введение  в терминосистему | мнемоническое удобство | – |
| Условие успешности использования | | адекватность внутренней формы смыслу | орфографическая стандартность, благозвучие | |
| Систем- ность | синтаг- матики | есть | нет | нет |
| парадиг- матики | нет | есть | нет |
| Отнесенность  к уровню реаль- ности | | теоретический | эмпирический | эмпирический когнитивной реальности |
| С чем соотносится | | объектом | вещью | сущностью |
| Способ демонстра- ции семантики | | эталон | тип2 | представление |
| Место хранения объекта | | сад | музей | сознание |
| Носителем чего является объект | | признаков | имени3 | – |
| Что демонстрирует | | признаки референта | отмеченный референт4 | само себя |
| Оппозиция в парадигме | | сигнификатов | референтов | имен |
| Ориентация на | | интенсионал | экстенсионал | интенсионал = экстенсионал |
| Характер употребления | | индивидуальный | конвенциональ- ный | конвенциональ- ный |

|  |  |  |  |
| --- | --- | --- | --- |
| **Аспект различия** | **Термин** | **Номен** | **Оним** |
| Принцип отбора | адекватность | приоритет5 (не выше семейства) | приоритет |
| Требование к значению | развитие | стабильность | – |
| Судьба при переводе | передача внутрен- ней формы | транслитерация | транслитерация, заимствование |
| Доминирующий тип словообразования6 | аналитический | синтетический | |

1 По Уэвеллу [Шелов, 1985].

2 В анатомии (как и в стратиграфии) типами являются фрагменты одного референта, что порождает сложности категоризации.

3 У Уэвелла — типичных свойств [Шелов, 1985].

4 [Реформатский, 1959].

5 Не выше семейства. Это аргумент в пользу принятия трактовки Н3 [Шелов, 1985].

6 Не принимая во внимание частот их омонимии в биологии.

Опираясь на осуществленное различение, можно утверждать, что в прикладной систематизации допустима стандартизация всех слоев лексики, а для научной необходима стандартизация номенклатуры (регу- лируется кодексами номенклатуры — напр., [Международный... 2005, Павлинов 2015**]**) и свобода терминологии. Специальные эмпирические референты (конкретные экземпляры) — номенклатурные типы, хранимые в гербариях, зоологических музеях и коллекциях культур микроорганиз- мов — являются носителями номенов, признаваемых всеми исследовате- лями, а другие референты — эталоны (стратоэталоны — [Стратиграфи- ческий... 2006]), признаваемые отдельными исследователями (хранимые в их личных коллекциях, которыми могут быть и ботанические и зоологи- ческие сады с живыми организмами), иллюстрируют использование тер- минов (бывающие омонимичными номенам). Экспертное отождествление типов и эталонов делает номены синонимами авторского термина, которые обозначают архетипы таксонов, поименованных соответствующими номе- нами. При этом номенклатурные комитеты как научные партии отстаивают завоеванные позиции, а систематизация развивается в идеале научными школами, постоянно подвергающими сомнению свои основания.

Пользуясь описанным аппаратом, можно делать сложные ревизии

таксонов, позволяющие сохранять понимание того, с каким эмпирическим материалом соотносятся вновь вводимые номены и термины. Так, напри- мер, если взять папоротник Голокучник обыкновенный, то окажется, что он имеет множество синонимов: Голокучник Линнея, Голокучник дриоп- терисовый, Голокучник трёхпёрый, Голокучник трёхраздельный, Голо- кучник щитовниковый, Щитовник Линнея. Латинская синонимика этого вида *Gymnocarpium dryopteris* (L.) Newman еще более обширна: *Dryopteris*

*dryopteris* (L.) Britton, *Dryopteris disjuncta* (Rupr.) C.V.Morton, *Dryopteris linnaeana* C. Chr., *Lastrea dryopteris* (L.) Bory, *Phegopteris dryopteris* (L.) Fée, *Thelypteris dryopteris* (L.) Sloss., *Aspidium dryopteris* Baumg., *Carpo- gymnia dryopteris* (L.) Á. Löve & D. Löve, *Currania dryopteris* (L.) Wherry, *Dryopteris pulchella* (Salisb.) Hayek, *Dryopteris pumila* V.I. Krecz., *Filix pumila* Gilib., *Nephrodium dryopteris* (L.) Michx., *Polypodium dryopteris* (L.), *Polypodium pulchellum* Salisb. Причина этого, в частности, в том, что по вполне весомым морфологическим основаниям (строения и ветвле- ния корневища, отсутствия индузия — покрывальца, прикрывающего сорусы — кучки спорангиев) в 1960-е гг. этот вид был вынесен из рода Щитовник (*Dryopteris*) во вновь выделенный род *Carpogymnia*, который, как оказалось через некоторое время, омонимичен зоологическому роду. Поэтому он был переименован в *Gymnocarpium*, что отчасти и определило дальнейшую номенклатурную нестабильность вида.

Характеризуя такую практику работы, Ж. Уолд в 1952 г. писал

* трудностях, с которыми ему пришлось встретиться при выяснении того, к каким животным относятся названия Cynocephalus mormon и Cynocephalus sphinx: «До того времени я считал, что одно из них — ман- дрил, другое — бабуин. После того как Нэтол в 1904 г. опубликовал свою работу, эти названия претерпели следующие удивительные превращения: Cynocephalus mormon стал Papia mormon или Papio maimon, который превратился в Papia sphinx. Это название легко можно было бы спутать с Cynocephalus, который стал теперь Papio sphinx, если бы последний не превратился тем временем в Papio papio. Едва устранив эту опасность, Papio sphinx переименовали в Mandrillus sphinx, a Papio papio — в Papio comatus. Все что я могу сказать по этому поводу, так это — слава богу, что одно животное называется мандрилом, а другое гвинейским бабуином» [Вилли 1964:94].

Однако, не смотря на такие сложности, предлагаемые альтернативы

[Павлинов 2015] оказываются неэффективными при оперировании с мил- лионами таксонов, и описанный способ работы не только сохраняется, но и распространяется на новые предметные области (фитосоциологию, стратиграфию, палеонтологию и т. д.).

**Литература**

1. *Вилли К*. (1964), Биология. М.
2. *Гринев-Гриневич С. В.* (2008), Терминоведение. М.
3. Международный кодекс фитосоциологической номенклатуры (2005). Рас- тительность России. СПб., № 7.
4. *Павлинов И.Я.* (2015)**,** Таксономическая номенклатура. Кн. 3. Современные кодексы. М.
5. *Реформатский А.А.* (1959), Что такое термин и терминология? М.
6. Стратиграфический кодекс России (2006), СПб.
7. *Чебанов С.В., Мартыненко Г.Я.* (1999), Семиотика описательных текстов (типологический аспект). СПб.
8. *Шелов С.Д.* (1985), Об одном классе научно-технической лексики (три подхода к выделению номенклатурных наименований). Научно-техническая информация. Сер. 1, № 3, с. 1–7.

**Literature**

1. *Villi K*. (1964), Biologiya. M.
2. *Grinev-Grinevich S.V*. (2008), Terminovedenie. M.
3. Mezhdunarodnyi kodeks ﬁtosotsiologicheskoy nomenklatury (2005). Rasti- tel’nost’ Rossii. SPb., № 7.
4. *Pavlinov I.Ya*. (2015), Taksonomicheskaya nomenklatura. Kn. 3. Sovremennye kodeksy. M.
5. *Reformatskiy A.A*. (1959), Chto takoe termin i terminologiya? M.
6. Stratigraﬁcheskiy kodeks Rossii (2006), SPb.
7. *Chebanov S.V., Martynenko G.Ya*. (1999), Semiotika opisatel’nykh tekstov (tipologicheskiy aspekt). SPb.
8. *Shelov S.D.* (1985), Ob odnom klasse nauchno-tekhnicheskoy leksiki (tri podkhoda k vydeleniyu nomenklaturnykh naimenovaniy). Nauchno-tekhnicheskaya informatsiya. Ser. 1, № 3, s. 1–7.

# КОРПУСНАЯ ЛИНГВИСТИКА

***Л.Д. Бакурадзе М.М. Беридзе***

**МОДЕЛИРОВАНИЕ ЛИНГВОКУЛЬТУРНОГО ПРОСТРАНСТВА ДИАЛЕКТНОГО ОСТРОВА В ГРУЗИНСКОМ ДИАЛЕКТНОМ КОРПУСЕ**

**(ферейданский диалект грузинского языка)**

**Аннотация.** Широкомасштабное корпусное документирование как лингви- стической, так и лингвокультурной модели идиоэтнического компонента языка является важнейшей задачей Грузинского диалектного корпуса (ГДК). В статье представлена концепция репрезентативности ГДК, которая ориентирована не только на фиксирование и изучение системных лингвистических особенностей, но также и на максимальное документирование и представление лингвокультур- ной модели диалекта. Помимо этого, данная концепция отражает специфические языковые процессы, протекающие внутри «островных» коллективов.

**Ключевые слова**. Грузинский диалектный корпус, языковой остров, лингво- культурное пространство, ферейданский диалект.

* 1. ***Bakuradze***
  2. ***Beridze***

**THE LINGUO-CULTURAL MODEL OF A “DIALECT ISLAND” IN GDC**

**(Fereydanian dialect)**

**Abstract.** Large-scale corpus documentation of both the linguistic and linguo- cultural models of an idio-ethnic component has been the primary task of the Georgian Dialect Corpus. The paper presents the concept of representativeness of GDC oriented both towards detection and study of systemic linguistic peculiarities and ultimate docu- mentation and representation of a linguo-cultural model of a dialect. The concept is also aimed at reﬂecting of speciﬁc linguistic phenomena within “island” communities. **Keywords.** Georgian Dialect Corpus, Language Island, Linguo-cultural Area,

Fereydanian Dialect.

Грузинский языковой остров сформировался в Иране в виде ферейдан- ского диалекта в результате нескольких вынужденных масштабных мигра- ционных потоков в XVII в. Он развивался в условиях особой изоляции от

© БАКУРАДЗЕ Лиа Давидовна. — Bakuradze Lia. — ТГУ Институт языкознания им. Арн. Чико- бава, Грузия. — TSU Arn.Chikobava Institute of Linguistics, Georgia. E-mail: [l.bakuradze@gmail.com](mailto:l.bakuradze@gmail.com)

© БЕРИДЗЕ Марина Михайловна. — Beridze Marina. ТГУ Институт языкознания им. Арн. Чикобава, Грузия. — TSU Arn.Chikobava Institute of Linguistics, Georgia. E-mail: marine- [beridze@yahoo.com](mailto:beridze@yahoo.com)

ядерного культурного и языкового ареала. На протяжении четырех веков диаспоре удалось почти полностью сохранить «культурную изоляцию». Разрыв этой изоляции, которой сопутствовали ускоренные темпы проте- кающих ассимиляционных процессов, начался только в 60-е гг. прошлого столетия [Беридзе и др. 2015]. Активизация ассимиляционных процессов совпала с периодом восстановления контактов с исторической родиной (70-е гг.), после чего начался интенсивный этап культурной и языковой реинтеграции, что, естественно, сказалось и на языковых процессах. В ферейданских текстах, представленных в Грузинском диалектном кор- пусе, отображаются эти два противоположных явления.

Для репрезентативного моделирования языковой картины главным фактором является полнота текстового и лексикографического материала. В ГДК интегрированы все тексты, опубликованные на сегодняшний день, а также материал, записанный нами во время экспедиций (2008–2014 гг.). Помимо этого, в общем словнике корпуса полно представлен и словарный материал (как изданный словарный материал, так и лексикографическая информация, полученная нами во время экспедиций, — около 6000 сло- варных единиц); в лексикографическом материале рядом со словарными формами приводятся также фонетические, словообразовательные и фор- мообразовательные варианты.

Задача описания ферейданского лингвокультурного пространства учитывалась в самом процессе подготовки текстового и лексикографиче- ского материала. Использовались традиционные методы фиксирования диалектного материала, такие как мотивация тематического повествова- ния, инициирование спонтанной речевой коммуникации, использование составленных специально для этих целей отраслевых и идиографических вопросников и т. д. Для повышения качества репрезентативности материал записывался с учетом возрастного и гендерного баланса, тематического разнообразия, региональной вариативности и др.

Нами была поставлена цель представить языковую картину на всех уровнях номинативными и функциональными возможностями языка, выявлением лакунарности, метафоризмом [Попова и др. 2007: 45].

Эти компоненты языковой картины активизируются в ГДК по-разному: номинативный компонент представлен в текстовых и лексикографических базах, функциональные возможности отражены в системе лингвисти- ческого аннотирования, маркирование тропных и фоносемантических средств возможно в информационном блоке словаря и различных струк- турных полях.

Когда мы говорим о диалектной языковой картине, естественно, имеем в виду языковую личность, являющуюся центральной фигурой этой картины.

**Языковая личность представляет один из основных феноменов лингвокультурного исследования. Она является носителем нацио- нальной языковой картины мира**.

«Диалектная языковая личность» также находится в центре созданной ею идиоэтнической картины мира.

**Диалектная языковая личность** становится объектом анализа согласно трем основным составляющим:

* + словарю диалектной языковой личности;
  + тексту, репродуцированному диалектной языковой личностью;
  + метаязыковому сознанию диалектной языковой личности [Иванцова 2002].

Исследование феномена диалектной языковой личности иногда длится в течение ряда лет и завершается изучением полного словарного запаса этой личности, языкового своеобразия и характерными особенностями способов концептуализации мира.

Как осуществляется в корпусе моделирование и репрезентация диа- лектной языковой личности?

Нами были записаны наиболее «яркие» носители ферейданского языкового кода. Среди них были и такие, речь которых уже ранее записы- валась исследователями старшего поколения.

Мы несколько лет подряд целенаправленно собирали информацию от одного и того же информатора, и он сам был активно заинтересован репродукцией своего «языкового знания»: он постоянно записывал всплывавшие в его памяти различные события, отдельные слова или фразеологизмы, пытался систематизировать, классифицировать мате- риал, «готовился» к тому, чтобы фиксировать этот материал вместе с научной группой. В результате такой работы у нас имеется довольно интересный текстовой и лексикографический материал для субкорпуса

«одного информатора». Несмотря на это, данного материала все-таки недостаточно для наглядного представления диалектной языковой личности.

Во время экспедиций мы пытались выявить феномен **коллективной языковой личности**, так как в островном языковом коллективе реально не существует одной личности, объемлющей всю языковую картину мира. Социум сам лично определяет/воспринимает фрагментарность общей языковой картины мира, и с целью преодоления этой фрагментарности он сам выделяет личности, «ответственные за определенные» участки. В таком коллективе свободно называют несколько человек, которые лучше других могут вести беседу на ту или иную тему. Здесь свои «хозяева» имеются у топонимов, антропонимов, шуточных историй, хозяйственной лексики, тематических историй из прошлой жизни... Эти люди не просто нарраторы, а «исполнители» определенного репертуара. Именно с уче-

том такой реальности мы собирали материал в Ферейдане и считаем, что в результате этого в корпусе представлена приблизительная модель язы- ковой картины мира, созданной ферейданской коллективной диалектной языковой личностью.

Основные компоненты языковой личности — организованный ею текст и словарный запас, как уже отмечалось выше, представлены в кор- пусе в виде текстовой базы и словаря ферейданского диалекта.

Метааннотирование корпуса и система лингвистического аннотиро- вания, а также информационная структура словаря организованы таким образом, чтобы с максимальной точностью представить метаязыковое сознание диалектной языковой личности.

Под метаязыковым сознанием личности подразумевается осмысление и оценка языковой личностью своей собственной речи и речи других.

Метаязыковое сознание играет особую роль в диалектном острове. Именно это сознание регулирует языковые процессы, возникающие при контакте с иноязычной средой и ядерным культурным ареалом. В ГДК по возможности наиболее полно сохранены «отпечатки» метаязыкового сознания диалектной языковой личности — текстовые базы корпуса и платформы словаря оформлены таким образом, что позволяют:

* + - выявить в словарях языковой материал, общий с ядерным культур-

ным ареалом;

* + - связать друг с другом посредством определенных ссылок существую- щие в диалекте параллельные формы;
    - выявить внутридиалектные различия и ареал распространения таких форм;
    - зафиксировать инновации и псевдодиалектные формы, возникающие в диалекте под влиянием литературного языка.

Специальный редактор словаря при помощи ссылок («смотри», «то же, что») связывает друг с другом параллельные формы, будь то персидско- грузинские параллелизмы, внутридиалектные различия или же параллель- ные формы диалекта и грузинского литературного языка.

Появление параллелизмов рядом с литературными формами представ- ляет один из признаков сближения с ядерным языковым ареалом. Такие параллелизмы легко, почти сразу же укореняются. Поляризация «мы — вы» часто представлена как результат языкового «наблюдения», как про- явление метаязыкового сознания, как наивная оценка языковых явлений. Информаторы в рамках одного текста называют собственные лексические единицы параллельно с соответствующими формами персидского языка, литературного грузинского языка («мы говорим так, а вы говорите так») и формами, которые встречаются в говоре соседнего селения. Такие при- меры закладывают основы синонимичных параллелизмов, характерных для островного языкового коллектива.

Параллельные формы грузинского литературного языка появляются в процессе реинтеграции с ядерным культурным ареалом рядом с дру- гим, похожим явлением — синонимичным параллелизмом, возникшим в результате влияния персидского языка, когда диалектная и иноязычная формы засвидетельствованы параллельно. Эта последняя указывает на ассимиляцию и теоретически дает возможность прогнозировать в бли- жайшее время исчезновение диалектной формы, а первая свидетельствует о том, что в диалекте уже начат процесс укоренения литературной формы.

В ферейданском диалекте ГДК при помощи ссылок отражены все зафиксированные случаи параллелизмов.

Сопоставление текстового и лексикографического материалов, изуче- ние лексики, общей с литературным языком, позволяет выявить лакунар- ное лексическое пространство — довольно большой «незаполненный» лексический участок в диалекте.

В ферейданском словаре ГДК отражены также внутридиалектные лексические различия — лингвистические и лексические особенности, существующие между разными грузинскими деревнями в Ферейдане.

Активизация метаязыкового сознания, по-видимому, является основ- ным признаком островной диалектной личности и часто становится причиной возникновения псевдолитературных форм, построенных по морфонологическим «законам» диалекта. Это происходит в том случае, когда из двух кодов, возникших в результате контакта с литературным языком, диалектная языковая личность отдает предпочтение литературной форме, старается следовать норме, меняя тем самым свой языковой код. Часто такая попытка содействует укоренению ошибочной формы. Такого типа случаи в научной литературе принято называть сверхкомпенсацией. Толчок к сверхкомпенсации дает сверхкомпетенция, когда возникают гипердиалектные/псевдодиалектные формы [Радченко и др. 2004]. Тра- диционная диалектология не допускает выделения в диалекте подобных сверхкомпенсационных данных и не признает их релевантными при опи- сании диалектного страта.

Система лингвистического маркирования словаря ГДК дает возмож-

ность дифференцировано представить ферейданскую лексику. Здесь маркированы формы, заимствованные из другого языка (New), псевдо- литературные формы (Pseudo), недавно вошедшие из грузинского лите- ратурного языка.

Грузинский диалектный корпус содержит информацию, касающуюся не только системных лингвистических характеристик диалекта, но и текущих процессов. Корпус, который постоянно пополняется, дает нам возможность наблюдать динамику развития в ферейданском диалектном острове, с одной стороны, ассимиляционных процессов и, с другой сто- роны, — интеграционных процессов.

**Литература**

1. *Беридзе М.M., Бакурадзе Л.Д.* (2015), Грузинский диалектный остров в Иране. Контрастивные исследования и прикладная лингвистика: материалы Международной научной конференции, г. Минск, 29–30 октября 2014 г.: в 2 ч. Ч. 2, УО «Минский государственный лингвистический университет», с. 3–6.
2. *Иванцова Е.В.* (2002), Феномен диалектной языковой личности: автореф. дис. Томск: Изд-во Томского университета, available at: <http://volgota.com/lib/> ivanlih.html (дата обращения 17.07.2016).
3. *Попова З.Д., Стернин И.А.* (2007), Семантико-когнитивный анализ языка: монография. Воронеж: Изд-во «Истоки».
4. *Радченко О.А., Закуткин Н.А.* (2004), Диалектная картина мира как идиоэт- нический феномен / Вопросы языкознания № 6, с. 25–48. available at: [http://www.](http://www/) ruslang.ru/doc/voprosy/voprosy2004-6.pdf (дата обращения 10.06.2016).

**Literature**

1. *Beridze M.M., Bakuradze L.D*. (2015), Gruzinskiy dialektnyi ostrov v Irane. Kontrastivnye issledovaniya i prikladnaya lingvistika: materialy Mezhdunarodnoy nauchnoy konferentsii, g. Minsk, 29–30 oktyabrya 2014 g.: v 2 ch. Ch. 2, UO «Minskiy gosudarstvennyi lingvisticheskiy universitet», s. 3–6.
2. *Ivantsova E.V*. (2002), Fenomen dialektnoy yazykovoy lichnosti: avtoref. dis. Tomsk: Izd-vo Tomskogo universiteta, available at: <http://volgota.com/lib/ivanlih.html> (data obrashcheniya 17.07.2016).
3. *Popova Z.D., Sternin I.A*. (2007), Semantiko-kognitivnyi analiz yazyka: monograﬁya. Voronezh: Izd-vo «Istoki».
4. *Radchenko O.A., Zakutkin N.A*. (2004), Dialektnaya kartina mira kak idioet- nicheskiy fenomen / Voprosy yazykoznaniya № 6, s. 25–48. available at: [http://www.](http://www/) ruslang.ru/doc/voprosy/voprosy2004-6.pdf (data obrashcheniya 10.06.2016).

***М.М. Беридзе***

**КОНЦЕПТУАЛИЗАЦИЯ ГРУЗИНСКОЙ ЯЗЫКОВОЙ ДЕЙСТВИТЕЛЬНОСТИ В ГРУЗИНСКОМ ЯЗЫКЕ**

**(по корпусным и лексикографическим данным)**

**Аннотация.** Под термином «грузинская языковая действительность» мы подразумеваем грузинское лингвокультурное пространство — картвельские языки — грузинский, занский (мегрельский и лазский), сванский и их диалекты как идиоэтнические компоненты одной языковой действительности. В статье рас- сматривается вопрос концептуализации языковой действительности по данным грузинских корпусов и словарей.

**Ключевые слова.** Грузинская языковая действительность, грузинская языко- вая картина мира, картвельские языки, корпус, лексикография.

***M.M. Beridze***

**CONCEPTUALIZATION OF THE GEORGIAN LINGUISTIC REALITY IN THE GEORGIAN LANGUAGE**

***(*Based on Corpus and Lexicographic Data*)***

**Abstract.** By the term “Georgian Linguistic Reality” I imply the geographical spread of the Georgian lingua-cultural areal: Kartvelian languages — Georgian, Zan (Megrelian-Laz), Svan, and their dialects as idio-ethnic components of the single lin- guistic worldview. I present the state of the conceptualization of the Georgian linguistic picture based on Georgian corpora and dictionaries.

**Keywords.** Georgian linguistic reality, Georgian linguistic picture, Kartvelian languages, Corpora, Lexicography.

Концепт «язык» как составляющая концептосферы «человек» явля- ется одним из важнейших при моделировании национальной картины мира. Номинативная презентация концепта «язык» в грузинском языке осуществляется различными лексемами: словами, обозначающими языко- вую иерархию: язык, речь, слово, наречие, говор.; **глаголами говорения**: сказать, разговаривать, говорить, рассказывать и др.; **другими глаго- лами со значением вербальной коммуникации**: понимать, слушать, общаться.; **глаголами с эпистемической семантикой**: знать, владеть, изучать (или же их метафорическими вариантами **—** «ломает»); **глаго- лами**: означает, есть.

Из дефиниционных признаков концепта «язык» для моделирования интересующей нас сферы (грузинская языковая действительность) суще-

© БЕРИДЗЕ Марина Михайловна. — Beridze Marina. ТГУ Институт языкознания им. Арн. Чикобава, Грузия. — TSU Arn.Chikobava Institute of Linguistics, Georgia. E-mail: [marineberidze@yahoo.com](mailto:marineberidze@yahoo.com)

ственными являются два признака: **этничность** (// **региональность**) **—** неотъемлемая и двусторонняя связь языка (или какой-либо региональной языковой подсистемы) с говорящим на нем коллективом и **манифести- руемость в акте речи** — проявление языка в виде конкретного речевого процесса [Полиниченко 2004].

Текстовые источники без исключения представляют указанные воз- можности вербализации концепта на всех иерархических и хронологиче- ских ступенях языка (древне-, средне-, новогрузинский язык, диалекты грузинского языка) [Беридзе 2011: 1–34].

С этой точки зрения интересным оказалось исследование данных лексикографических ресурсов. Нами изучен вопрос, насколько реа- лизован дефиниционный признак **этничности/региональности** для моделирования концепта «язык» в грузинских словарях. Исследование показало, что при толковании **лексемы «язык»** в грузинских словарях почти не представлен признак этничности языка в главных структурных полях словарной статьи, однако в иллюстративном блоке словарей все- таки проскальзывает неотъемлемая связь языка с этничностью [ТСГЯ] [Абуладзе 1973], а при толковании имен прилагательных, производных от этнонимов, дефиниционный признак этничности/региональности также в основном представлен только в иллюстративном блоке (**грузинский** язык, **греческий** язык, **армянский** язык и т. д.). А в некоторых словарях вообще не активизировано этой возможности выявления концепта [Орбе- лиани 1991, 1993].

**Признак манифестации в речи** представлен главным образом упо- треблением имени концепта в сочетании с глаголами, обозначающими речевую деятельность в словосочетаниях типа говорить на каком-либо языке [Полиниченко 2004].

В грузинском языке, так же как и в других языках, часто встречается эллиптическая модель при выражении признака этничности языка, когда лексема, номинирующая язык (или наречие), опущена, и презентация кон- цепта происходит именем прилагательным, обозначающим этническое или региональное происхождение: *kartulad [A:Sg Ad] laparako*̣*bs по-грузински разговаривает.*

Эллиптическая модель концепта представлена двумя компонентами **—** именным (имя этнического/регионального происхождения) и глагольным. Соответственно, следует ожидать, что в словарях должна быть отражена связь с именем концепта, с одной стороны, при толковании имен, обо- значающих происхождение, а с другой стороны, — соответствующих глаголов.

Нужно сказать, что при толковании имени концепта — «язык», а также других составляющих данного лексико-семантического поля не имеется полного отражения концепта. Связь перечисленных глаголов с лексикой

номинативного поля данного концепта либо крайне редко встречается в информационных полях словарных статей, либо же совсем не засвиде- тельствована.

Это то, что касается вопроса отражения в грузинских лексикографиче- ских источниках концептуального «портрета» ядерной лексемы концепта

«язык».

Интересно, как обстоит дело в отношении тех единиц номинативного поля концепта, которые используются в языке с целью внутриязыковой дифференциации: это имена, обозначающие различные речевые разновид- ности (диалект/наречие, говор), и связанные с ними глаголы говорения и глаголы с эпистемичной семантикой, а также лексемы регионального происхождения.

Следует отметить, что дифференциация по региональному проис- хождению в грузинском дискурсе совпадает с лингвогеографической дифференциацией и представлена она формами, обозначающими регио- нальное происхождение, образованными по одной и той же модели: imeruli (имеретинский), mesxuri (месхский), svanuri (сванский) и др. Все эти еди- ницы объединяет один термин со значением этнического происхождения kartuli (грузинский).

В отличие от лексикографических источников, текстовой материал в полной мере репрезентирует концепт «язык» и его фрагмент — **грузин- ская языковая действительность**. Исследования возможностей вербаль- ной объективации концепта было осуществлено на базе существующих корпусных ресурсов [GNC, KAWAC, GDC]. Нами были изучены колока- ции лексем номинативного поля концепта. К сожалению, формат статьи не позволяет детально описать все аспекты исследования и тем более не дает возможности продемонстрировать иллюстративный материал. Ниже представлены только полученные выводы:

* + В текстовом материале строго дифференцирована иерархичность языка: имя концепта — **ena (язык)** — представляет верхнюю ступень этой иерархии, и с ним может быть связана лишь определенная группа глаголов, а с языковыми единицами более низкой иерархии (напр., наре- чие, говор) связаны глаголы другой группы.
  + С лингвистическими терминами, обозначающими языковую подси- стему (диалект, наречие, говор), в эллиптической модели **не используются** эпистемичные глаголы (семантического поля знания, компетенции): напр., знает имеретинский, изучает кахетинский, забыл гурийский и др.
  + В эллиптических моделях при терминах, обозначающих низкую языковую иерархию, **не встречаются** глаголы говорения (\*разговаривает по-гурийски, \*говорит по-кахетински, \*сказал по-имеретински, \*беседует по-гурийски, \*спросил по-гурийски, \*«ломает» по-гурийски).
    - С этими же терминами из глаголов говорения **встречаются** всего лишь глаголы ***u-kc-ev-s* и u-ki**̣**d-eb-s** (оба эти глагола обозначают диалект- ную речь): *gurul-ad, kartlur-ad, moxeur-ad, tušur-ad*… ***ukcevs*** (разговаривает на гурийском, картлийском, мохевском, тушинском и др. диалектах или говорит на литературном языке с акцентом этого диалекта).
    - Глаголы говорения могут **употребляться** при терминах низкой языковой иерархии только **в неэллиптической модели**, когда в высказы- вании обязательно представлен термин, обозначающий эту подсистему: разговаривает, говорит, беседует, изьясняется на гурийском, мохевском, кахетинском, имеретинском. диалекте (наречии).
    - Использование глаголов говорения ***ukcevs, uk***̣***idebs*** при терминах высокой языковой иерархии возможно только в том случае, когда человек, владеющий тем или иным языком разговаривает с акцентом (интонацией) другого языка: ***kartulad ukcevs, uk***̣***idebs*** — разговаривает на каком-либо языке с грузинским акцентом.
    - **Недопустимо** использование глагола «ukcevs» с терминами: сван- ский, мегрельский, лазский, исключая те случаи, когда речь идет о грузи- ноязычной коммуникации (или же коммуникации на другом языке) пред- ставителя этого языкового круга: ***svanurad, megrulad, ukcevs*** — говорит на грузинском (или каком-либо другом языке) с сванским, мегрельским, акцентом.

Как показал текстовой анализ, в грузинском языковом сознании концептуализована также картвельская диглоссия: ***svanurad laṗaraķobs***

«говорит по-свански» (= говорит на сванском языке) / ***svanurad ukcevs*** (= говорит на другом языке со сванским акцентом) — это признаки одной языковой личности, члена одного лингвокультурного единства. ЯЛ — это

«своего рода «семантический фоторобот» — личность словарная, этносе- мантическая» [Воркачев 2001]. Этот «семантический фоторобот» может репрезентировать общенациональные знания и ценностные приоритеты на двух языках: созданная им языковая картина мира объединена сово- купностью универсальных и идиоэтнических феноменов, подобно тому, как языковая картина мира любого представителя других этнографических регионов Грузии является языковым отражением универсального, обще- национального и регионального восприятия картины мира.

Проследим, как в ТСГЯ концептуализован этот срез национальной культурной картины мира.

Интерес для нас представляют прилагательные, образованные от грузинских этнонимов, как одна из важных возможностей вербализации концепта «язык». При толковании одного из значений прилагательного — ***kartuli*** *(грузинский)* — в ТСГЯ указывается имя концепта: ***kartuli ena*** *(грузинский язык).*

Что же касается прилагательных, обозначающих региональное про- исхождение, их словарные статьи структурированы неодинаково — часть этих толкований указывает на связь с номинативным полем концепта

«язык», а другая часть вообще не отражает такого отношения. Очевидно, что для составителей толкового словаря не стояло конкретной задачи концептуальной полноты.

В первом случае в поле толкования или иллюстративном материале есть указание на термин низкой иерархической ступени внутриязыковой дифференциации ***ki***̣***lo*** *(диалект, наречие):* ***kartluri*** *(картлийский),* ***tu*š*uri*** *(тушинский),* ***ingilouri*** *(ингилойский),* ***xevsuruli*** *(хевсурский),* ***fereydnuli*** *(ферейданский),* ***guruli*** *(гурийский) — диалект, наречие.*

Во второй части словарных статей, как уже было сказано выше, не отражается связи термина, обозначающего региональное происхожде- ние с семантическим полем языка: ***svanuri*** (сванский), ***lazuri*** (лазский), ***k***̣***axuri*** (кахетинский), ***ač***̣***aruli*** (аджарский). Более того, в словаре вообще не приводятся названия некоторых регионов, как и имен прилагательных, производных от этих названий.

При толковании имен прилагательных со значением регионального происхождения, с которыми связан картвельский языковой компонент: ***svanuri*** *(сванский),* ***lazuri*** *(лазский),* ***č***̣***anuri*** *(чанский),* ***megruli*** *(мегрель- ский)* связь с семантическим полем указана только при лексеме «чанский»: čanuri ***k***̣***ilo****(чанское наречие).*

Как видим, грузинские лексикографические ресурсы фрагментарно описывают интересующий нас срез грузинского лингвокультурного про- странства, и только корпусный материал дает возможность полноценного его моделирования. Также выяснилось, что корпусно репрезентированная концептуализация грузинской языковой действительности с точностью не совпадает с некоторыми научными концептуализациями и вычленяет в грузинской языковой действительности четыре языка — **грузинский, сван- ский, мегрельский, лазский** — и их диалекты. Проведенное исследование еще раз подтверждает, что корпусный материал является безальтернативным источником для изучения лингвокультурной проблематики. А также свиде- тельствует, что лексикографические недостатки при отражении языковой картины мира можно преодолеть, опираясь на системный корпусный подход.

**Литература**

1. *Абуладзе И.В.* (1973), Словарь древнегрузинского языка, Тб.
2. *Беридзе М.М.* (2011), Language Conceptualization of the Georgian ethno-social reality FOLIA CAUCASICA, Frankfurt am Main-Tbilisi, pp. 1–34.
3. *Воркачев С.Г.* (2001), Лингвокультурология, языковая личность, концепт: становление антропоцентрической парадигмы в языкознании // Филологические науки, № 1, с. 64–72.
4. *Орбелиани Сулхан-Саба* (1991, 1993), Leksiḳoni kartuli (Грузинский сло- варь), Т. 1–2.
5. *Полиниченко Д.Ю.* (2004), Естественный язык как лингвокультурный семио- тический концепт: автореф. дис. канд. филол. наук. Волгоград, 22 с.
6. ТСГЯ (1950–1964), Толковый словарь грузинского языка, Тб.
7. Грузинский диалектный корпус: [http://corpora.co](http://corpora.co/)
8. Грузинский национальный корпус: <http://gnc.gov.ge/gnc>
9. Веб-корпус грузинского языка: <http://corpus.leeds.ac.uk/internet.html>

**Literature**

1. *Abuladze I.V*. (1973), Slovar’ drevnegruzinskogo yazyka, Tb.
2. *Beridze M.M*. (2011), Language Conceptualization of the Georgian ethno-social reality FOLIA CAUCASICA, Frankfurt am Main-Tbilisi, pp. 1–34.
3. *Vorkachev S.G*. (2001), Lingvokul’turologiya, yazykovaya lichnost’, kontsept: stanovlenie antropotsentricheskoy paradigmy v yazykoznanii // Filologicheskie nauki,

№ 1, s. 64–72.

1. *Orbeliani Sulkhan-Saba* (1991, 1993), Leksiḳoni kartuli (Gruzinskiy slovar’), T. 1–2.
2. *Polinichenko D.Yu*. (2004), Yestestvennyi yazyk kak lingvokul’turnyi semio- ticheskiy kontsept: avtoref. dis. kand. ﬁlol. nauk. Volgograd, 22 s.
3. TSGYa (1950–1964), Tolkovyi slovar’ gruzinskogo yazyka, Tb.
4. Gruzinskiy dialektnyi korpus: [http://corpora.co](http://corpora.co/)
5. Gruzinskiy natsional’nyi korpus: <http://gnc.gov.ge/gnc>
6. Veb-korpus gruzinskogo yazyka: <http://corpus.leeds.ac.uk/internet.html>

***Е. Бертеме***

**ИСПОЛЬЗОВАНИЕ КОРПУСОВ ПРИ ОБУЧЕНИИ ПРАГМАТИЧЕСКОЙ СОСТАВЛЯЮЩЕЙ РЕЧЕВЫХ ФОРМУЛ**

**Аннотация.** В статье высказывается мнение о необходимости использования корпусов для обучения иностранным языкам. Русские формулы рассматрива- ются в сопоставлении с французскими, для последних предлагается дословный перевод и краткое толкование. Большая часть примеров взята из личного опыта автора в качестве преподавателя русского языка для иностранной, по большей мере франкоязычной, аудитории. На примере НКРЯ особое внимание уделяется формуле *Ни пуха ни пера*. Высказывается мнение о совместном использовании словарей и корпусов. Результаты исследования могут быть рассмотрены как вклад в преподавание иностранных языков при помощи корпусов.

**Ключевые слова.** Иллокутивный речевой акт, иностранный язык, коммуни- кация, контекст, фразеология, словарь.

***E. Berthemet***

**THE USE OF CORPORA IN LEARNING OF PRAGMATIC PART OF FORMULAE**

**Abstract.** In this paper, a peculiar type of idioms which are called formulae is considered. It is referred to only pragmatic part of formulae for the sake of conﬁning it to an acceptable number of pages. In order to illustrate our proposal Russian formulae are considered in comparison with French idioms. Literal translation and short deﬁnition are proposed. Although the paper is relevant to Russian language teaching it addresses issues related to formulae teaching in general. We consider how the Russian National Corpus can contribute to learn the pragmatic part of formulae. Our impression is that the potential of the RNC used to teach formulae has not yet fully been exploited. Pos- sibilities for combining dictionary resources with corpora are discussed.

**Keywords.** Сommunication, context, dictionary, foreign language, illocutionary speech act, phraseology.

### Проблематика

Известно, что для полноценного практического владения иностранным языком, помимо лексики, грамматики и произношения, студентам очень важно овладеть прагматической информацией. Использование таких речевых формул, как *Скатертью дорога! Нашел дурака!* и *Ишь ты!* является одним из признаков аутентичности общения. В данном исследо- вании, вслед за Барановым и Добровольским, под *речевыми формулами*

© БЕРТЕМЕ Елена. — Berthemet Elena. Внештатный член научного центра лингвистики в Сорбонне, Франция. — Associated member of Centre de Linguistique en Sorbonne, France. E-mail: [elenaberthemet@free.fr](mailto:elenaberthemet@free.fr)

понимаются «идиомы разных структурных типов (преимущественно законченные высказывания) с фиксированной иллокутивной силой или определяющие иллокутивные характеристики речевого высказывания» [Баранов, Добровольский, 2008: 81]. Речь идет об устойчивых выражениях, которые не придумываются каждый раз по мере необходимости и служат в целях коммуникации.

Форма, значение и употребление речевых формул зарождаются в определенном языковом сообществе, и поэтому не принадлежащий к нему человек может испытывать сложности в их изучении. Возможно, именно формулы, в большей мере, чем иная лексика, требуют понимания ситуации, в том числе социальной, а недопонимание приводит к ошибоч- ному употреблению. Для того, чтобы понять, а затем воспроизвести новую речевую формулу, иностранные студенты должны не только выучить ее саму, но и новый смысл, который содержит прагматическую составляю- щую. Иными словами, они должны четко представлять прототипическую ситуацию общения, в которой выражение может быть использовано.

В самом простом случае похожая ситуация в разных языках может тре- бовать употребления формул, основанных на схожих внутренних формах1, т. е. ситуация языка 1 (СЯ1) = ситуации языка 2 (СЯ2), а внутренняя форма языка 1 (ВФЯ1) = внутренней форме языка 2 (ВФЯ2). Так, например, во французском языке существуют эквиваленты формул *Каким ветром?* (фр. *Quel vent ?*) и *Какая муха тебя укусила?* (фр. *Quelle mouche t'a piqué ?*).

В случаях, когда схожая ситуация лексикализирована в обоих языках, но имеются расхождения во внутренней форме, т. е. СЯ1 = СЯ2, но ВФЯ1

≠ ВФЯ2, сложность изучения эквивалента возрастает. Речь идет о таких единицах, как *Мало не покажется!* (фр. *Tu vas voir ce que tu vas voir* ‘Уви- дишь, что ты увидишь’) и *Лёгок на помине* (фр. *Quand on parle dи loup [on en voit la queue]* ‘Стоит заговорить о волке [, виден его хвост]’).

Кроме этих двух случаев, относительно легких для изучения в том смысле, что выражение существует в родном языке и, следовательно, студенты естественным образом захотят сказать что-то в данной ситуации, существуют русские выражения, которые непонятны для иностранца, т. е. СЯ1 ≠ СЯ2, и, следовательно, ВФЯ1 > 0 (отсутствие идиоматического эквивалента). Такие формулы, как *Знай наших! Счастливо оставаться! С лёгким паром!* и *Носи/Кушай на здоровье !* не только сложно объяснить, но и найти эквивалент тяжело в связи с тем, что формулы не имеют какого- либо вербального отражения во французском языке.

Незнание формул может привести к тому, что общение не только не облегчается, но, напротив, затрудняется, ставя одного или нескольких участников в затруднительное положение. Так, например, употребление

1 Под внутренней формой, согласно Баранову и Добровольскому, имеется в виду «образ, фик- сированный в плане содержания фразеологизма <…>» [Баранов, Добровольский 2013 : 130].

следующих дословных формул может привести к забавным ситуациям:

1. фр. *Vous avez l’heure?* ‘У вас есть время?’, означающую ‘который час?’ нельзя перевести на рус. *У вас есть время?* и (2) фр. *Tu m’étonnes* ‘Ты меня удивляешь’ со значением ‘ничего удивительного, я полностью с тобой согласен’ практически является антонимом русскому *Ты меня удивляешь.* Противоположные значения имеют также высказывания (1) фр. *Chich* ‘Шиш’ ‘спорим (, что ты этого не сделаешь)’, т. е. собеседнику бросается вызов, в то время как русский *Шиш [с маслом]* означает отказ собеседнику в какой-либо просьбе и (2) фр. *Pas question* ‘Не вопрос’ выражает несо- гласие, а русское *Не вопрос* означает согласие.

Проблема в том, что, в отличие от морфологических и синтаксических ошибок, заметить прагматические ошибки в употреблении речевых фор- мул можно лишь в контексте, т. е. в диалоге, так как используются они преимущественно в устной речи. На практике воспроизведение ситуаций в учебной обстановке трудоемко и занимает большое количество времени. Обратимся к учебным словарям и попытаемся ответить на вопрос: какую информацию о прагматической составляющей в речевых формулах

могут найти студенты?

***Ни пуха ни пера* в печатных словарях**

Возьмем для примера русскую формулу *Ни пуха ни пера*. Грубо говоря, во французском языке ей могут соответствовать два эквивалента : (1) фр. *Merde* ‘Дерьмо’ и (2) фр. *Bonne chance* ‘Удачи’. Сложность состоит в том, что, во-первых, первый эквивалент стилистически окрашен. Во-вторых, на русскую формулу *Ни пуха ни пера* принято отвечать *К чёрту*, в то время как первый французский эквивалент, напротив, не требует ответа, а после второго принято использовать формулу фр. *Merci* ‘Спасибо’.

Большинство двуязычных изданий игнорируют формулы: либо они предлагают эквиваленты, которые с трудом являются таковыми, либо они отсутствуют. Одноязычные словари принимают во внимание и прагмати- ческую составляющую.

Ниже приводится словарная статья из *Фразеологического объясни- тельного словаря русского языка ФОСРЯ* [Баранов, Добровольский 2009], посвященная речевой формуле *Ни пуха ни пера* и иллюстрирующая данное утверждение:

Ритуализованная реакция пожелания успеха адресату в каком-либо деле (экзамене, конкурсе, соревновании, поединке, получении какого-л. ресурса и т. п.), вызываемая завершением коммуникации между адреса- том и говорящим и не предполагающая возобновления коммуникации до момента завершения этого дела, которая предполагает ответ «К черту». Порядок слов фиксирован.

[Примеры]

Во-первых, благодаря термину «ритуализированная» реакция, ино- странному студенту понятно, что речь идет о негласном социальном согла- шении, а слово «реакция» указывает на то, что формула интерактивна, иными словами, говорящий имеет представление о знаниях собеседника и пользуется ею для межличностного взаимодействия.

Во-вторых, указывается на конкретный иллокутивный речевой акт, а именно «пожелание успеха», при котором говорящий не просто утверж- дает что-то с целью повлиять на исход ситуации (апеллятивная функция в терминологии Якобсона), но и выражает свою поддержку (экспрессивная функция).

В-третьих, предлагаются конкретные примеры ситуаций, в которых принято произносить формулу: «экзамен, конкурс, соревнование…». Таким образом, французский студент может увидеть разницу в употребле- нии. Кроме того, подчеркнута немаловажная информация о том, что речь идет о «завершении коммуникации <…> до момента завершения дела», что также не является очевидным для иностранцев.

В-четвертых, пользователь предупрежден о том, что принято давать ответную формулу *К черту*. И, наконец, выражение «порядок слов фик- сирован» указывает на то, что фраза устойчива. Таким образом, прагма- тическая сторона данной формулы объяснена прямо в толковании, и в дополнение предлагаются примеры.

Сложность состоит в том, что иностранные студенты редко пользу- ются одноязычными толковыми словарями. Необходимо отметить, что словари, объясняющие прагматическую сторону формул с такой же чет- костью, как процитированный ФОСРЯ, являются скорее исключением. Кроме того, возможность проконсультировать этот словарь во Франции очень редка, поэтому для целенаправленного обучения формулам пред- ставляется необходимым обращение к корпусам. Необходимо также заме- тить, что традиционные словари физически ограничены, а корпусы — нет.

### Ни пуха ни пера в НКРЯ

Несомненным преимуществом корпусов является возможность пред- ставить речевые формулы в разнообразных контекстах. Таким образом, они позволяют объяснять их функционирование внутри реального дис- курса, т. е. в тесной связи с текстом и ситуацией, в которых они произ- несены. Однако, принимая во внимание то, что смысл формул может меняться не только в зависимости от контекста, но и от интонации и что формулы зачастую сопровождаются жестами и мимикой (например, *Тьфу тьфу тьфу!* и *Да брось!*), разработка мультимодальных корпусов, вклю- чающих речевые формулы, принесла бы еще большую пользу.

На запрос *Ни пуха ни пера* НКРЯ предлагает 121 вхождение в 99 доку- ментах (на 20 сентября 2016 г.). Безусловно, данное количество примеров

больше подходит лексикографу для достаточно точного описания ситуа- ции. Однако можно представить преподавателя, при условии наличия сво- бодного времени, который может вручную перебрать примеры и выбрать ситуации, в которых прагматическая сторона формул наиболее ярко выражена. Действительно, такие примеры, как (1) и (2), вряд ли помогут иностранцу представить ситуацию, в этом смысле они являются пустыми:

* 1. Воробьи стали нагонять нас, и, вероятней всего, они задали бы нам такую взбучку, после которой от нас с Малининым не осталось бы *ни пуха ни пера* [Валерий Медведев. Баранкин, будь человеком! (1957)]
  2. Все справки и рапорта шли через него, а потому он был «в абсо- лютном курсе». ― Завтра утром. ― Ну, чтоб со щитом. ― «*Ни пуха ни пера*» Сан Саныч не желал никому, дабы его не посылали к черту.

Более того, в примере (1) выражение не является идиомой, и, следо- вательно, оно не может способствовать пониманию фразеологического смысла, но студент вряд ли может понять это.

А такие результаты, как (3) и (4), вполне могут дать общее представ- ление о ситуации, в которой используется выражение:

* 1. Перед экзаменом на пожелание «*ни пуха ни пера*» ― послать к черту; посмотреть в зеркало, а также положить в туфлю под левую пятку монету в пять копеек или пять рублей (пятак). [Ю. В. Саенко. Суеверия современных студентов (2004) // «Вопросы психологии», 2004.08.10]
  2. ― Ну давай… Пожелай мне *ни пуха*… Если со мной что-то слу- чится, можешь оставить себе мою майку, ― тихо проговорил Ванька и встал. [Дмитрий Емец. Таня Гроттер и магический контрабас (2002)]

Так, в (3), помимо того, что объясняется ритуализированная реакция *Ни пуха ни пера* и ответная реакция посылания к черту, упоминается о других суевериях современных русских студентов, что, вероятно, будет интересно иностранным студентам. Например, информация о сравнении высшей оценки с монетой будет полезна французам, у которых двадцати- балльная система оценок и, таким образом, эта информация обогатит их знания о русской культуре.

Из сказанного следует, что речевые формулы необходимо рассма- тривать в комплексном подходе, т. е. с учетом их семантики, синтаксиса и прагматики.

Ситуации, в которых употребляются речевые формулы, совсем не описаны в двуязычных, наиболее популярных и экономных с точки зрения времени словарях. Наилучшими источниками прагматической информации являются одноязычные словари и корпусы. Оба ресурса, несмотря на безусловные преимущества по отношению к двуязычным словарям, могут быть проблематичными для прямого самостоятельного пользования студентами. Во-первых, из-за сложности используемой лек-

сики, а во-вторых, относительно корпусов, из-за обилия примеров, а также достаточно большого числа «пустых» контекстов. Кроме того, корпусы предлагают множество примеров, в которых прагматическая информация изложена имплицитно, и, следовательно, требует определенных усилий и временных затрат.

Можно было бы предположить, что для экономии времени студентов преподаватель собственноручно отбирает несколько значимых примеров. Но практически это очень сложно осуществить по причине нехватки времени, так же, как и обнаружить прагматические ошибки. Таким обра- зом, несмотря на временные затраты, корпусы являются несомненным источником разного рода информации, в том числе прагматической, и предоставляют дополнительные возможности для целенаправленного обучения идиомам.

**Литература**

1. *Баранов, А.Н., Добровольский Д.О*. (2008), Аспекты теории фразеологии, Москва: Знак.
2. *Баранов А.Н., Добровольский Д.О*. (2009), *Фразеологический объяснитель- ный словарь русского языка*. ФОСРЯ, Москва: ЭКСМО.
3. *Баранов А.Н., Добровольский Д.О*. (2013), Основы фразеологии, Москва: Флинта, Наука.

**Literature**

1. *Baranov, A.N., Dobrovol’skiy D.O*. (2008), Aspekty teorii frazeologii, Moskva: Znak.
2. *Baranov A.N., Dobrovol’skiy D.O*. (2009), Frazeologicheskiy ob’yasnitel’nyi slovar’ russkogo yazyka. FOSRYa, Moskva: EKSMO.
3. *Baranov A.N., Dobrovol’skiy D.O*. (2013), Osnovy frazeologii, Moskva: Flinta, Nauka.

***S. Birzer, H. Zinsmeister***

**THE UTILITY OF COLOUR MARKUP IN TEXTS FOR LEARNERS OF RUSSIAN**

**Abstract.** Cyrillic script features a number of easily confusable graphemes that tend to co-occur within derivational and inﬂectional morphology, e.g. *-ющий; -ание*, and pose a problem to learners. In an online survey we measured the effect of colour markup on word recognition tasks with different learner groups. Colour markup in general resulted in a higher rate of correct answers. Its effect on the reaction time is more complex and will be discussed in detail in the paper.

**Keywords.** Corpus linguistics, reading skills, colour markup, reaction times, graphemes, morphemes, statistics, experiment

### 1. Introduction

Easily confusable Cyrillic graphemes are a notorious problem for both L1 and L2 learners of Russian. For L1 learners, didactic literature [e.g. Bobkina 2012; Lysko 2013] recommends exercises enhancing phonemic awareness to improve writing skills — an approach that is hardly applicable for L2 learners. The didactic literature on Russian as L2 has recently addressed problematic issues for dyslexic students [cf. e.g. Mehlhorn 2012], but difﬁculties on the graphematic level are disregarded. The experimental data presented in this paper aims at this lacuna, as it tests the usefulness of grapheme and morpheme mark up for displaying texts.

The hypotheses for our test design are the following:

1. Beginners are mainly concerned with the distinction of graphemes in morphologically “plain” vocabulary. Consequently, we test whether different colouring of easily confusable graphemes improves the speed of word recognition in comparison to the same stimuli without mark up.
2. Advanced learners are confronted with morphologically complex vocabulary. Given the co-occurrence of easily confusable graphemes in sufﬁx morphology, we test the effect of different colouring of easily confusable graphemes vs. the different colouring of sufﬁxes vs. no mark up.

### Test design

The experiment was conducted online between February and March 2016 with the help of the online test design tool SoSciSurvey. The test was available in the two instruction languages German and English.

© BIRZER Sandra. Innsbruck University, Austria. E-mail: [Sandra.Birzer@uibk.ac.at](mailto:Sandra.Birzer@uibk.ac.at)

© ZINSMEISTER Heike. Universität Hamburg, Germany. E-mail: Heike.Zinsmeister@uni- hamburg.de

* 1. ***Design parameters***

In the ﬁrst part of the test participants are asked to give the following sociolinguistic information:

* Age
* Learner type
  + heritage language
  + foreign language
* Length of study
  + less than 6 months
  + more than 6 months
  + more than 1 year
* Acquisition of reading skills
  + at home
  + at school in Russia / in a Russian speaking country
  + in Russian classes at school in a non-Russian speaking country
  + in the course of university studies in a non-Russian speaking country
  + in a Russian educational insti- tution situated in a non-Russian speaking country
  + elsewhere (option to specify)
* Country of birth
  + German-speaking countries (Germany, Austria, Switzerland)
  + English-speaking countries (USA, Great Britain, Canada)
  + countries of the former Soviet Union
  + other (with the option to specify)
* Age of migration
* Type of university studies
  + Russian Philology
  + Russian Linguistics
  + Philology of a language other than Russian
  + Linguistics
  + other humanities
  + other studies
  + I do not attend / have not attended university
* Colour vision deﬁciency

We will now turn to describing the actual experiment. We decided to test only non-italic Cyrillic script for several reasons: It is the standard for printed Russian texts and for display of texts in the Russian National Corpus — an important point of reference, as we plan to use our test result for the construc- tion of a levelled study corpus for learners of Russian. Students, especially beginners, are familiar with non-italic script, but not necessarily with italics, and the strings of easily confusable graphemes differ in italic and non-italic script.

* 1. ***Beginners***

For the beginners we hypothesized that they are mainly concerned with the distinction of graphemes as such and that morphological issues and thus colour mark up of sufﬁxes is not relevant for them, since the basic vocabulary they know is morphologically simple. Consequently, we tested whether differ- ent colouring of easily confusable graphemes improved word recognition in comparison to the same stimuli in black type. As an improvement we deﬁned a higher number of correctly recognized words, in the ideal case to be accompa- nied by a lower reaction time. Consequently, all tests were designed to measure reaction time.

Vocabulary was required to be understandable for beginners and to contain easily confusable graphemes. To meet both criteria, we included several internationalisms such as *инструкция — instrukcija* ‘instruction’ in the test.

To restrict test wiseness, we employed various test formats. We took precau- tions so as not to make the order of answers predictable and the correct answers for all tasks were given only after the whole test had been completed. Since the recognition speed of words as such is a result of training, all tasks in the test and all items of each task were given in the same order to all respondents; no permu- tations of tasks and items were conducted so as not to affect the reaction times.

* + 1. ***(In)correct spelling of a word (“spelling”)1***

Respondents were presented one word at a time and asked whether the spelling was correct or not. The test contained both correctly and incorrectly spelled trigger words; the “spelling mistakes” were of different types (wrong order of correct graphemes; incorrect grapheme in the right place) and occurred in different places and with different graphemes.

* + 1. ***Jumbled letters (“jumbled\_letters”)***

Respondents were presented with four variants of the same word, three of them with jumbled graphemes, and asked to pick the variant with the correct order of graphemes.

* + 1. ***Which word is depicted in the picture? (“word\_assign”)***

Respondents were presented with one picture and two lexemes at one time and had to pick the lexeme that matched the picture best. Roughly speaking, the pair of lexemes formed a “graphematic minimal pair”, i.e. they differed in one grapheme, and the two differing graphemes constituted a pair of easily confusable graphemes.

* + 1. ***Which word in the sentence is spelled wrong? (“sentence”)***

Respondents were presented with one sentence at a time and asked to identify the incorrectly spelled word.

* + 1. ***Which column contains most incorrectly spelled words? (“column”)***

Respondents were presented with one table at a time; each table contained three columns with four words each.

* 1. ***Advanced learners***

The aim of the experiment was to ﬁnd out whether coloured markup as such improves word recognition and, if so, whether learners proﬁt more from the marking of single, easily confusable graphemes or from the marking of

1 The keyword in parentheses is used in the results section to refer to the task.

sufﬁxes.1 Consequently, the experiment for advanced learners had three ver- sions, namely black type only, secondly, colour marking of easily confusable graphemes (as with the beginners) and, thirdly, colour markup of derivational and inﬂectional morphology.

* + 1. ***Filling in gaps with a choice of word forms with the same root morpheme (“gaps”)***

Respondents were presented one sentence at a time containing a gap and were asked to ﬁll in the gap from a choice of four possible word forms. The answers were all formed with the same root, but represented different word classes and thus comprised different derivational and inﬂectional morphology. We expected the context to enable respondents to identify the word class that ﬁts the gap best, and that the colour markup (of morphemes more than of single graphemes) of the answers helps to identify their word class status faster and more reliably.

* + 1. ***Identiﬁcation of synonyms (“synonyms”)***

Respondents were confronted with one lexeme as stimulus and four other lexemes as potential synonyms. Additionally, respondents could choose the answer “don’t know”. Neither the stimulus nor the answers were embedded into a context. The synonym pairs either belonged to the same word class but represented different morphological patterns or belonged to different word classes with the same function (e.g. attributive function for both adjectives and participles). Again we hypothesized that colour markup enhances reaction time and the number of correct answers.

* + 1. ***Assignment of grammatical categories (“gram\_cats”)***

Many (semi)productive derivational morphemes contain easily confusable graphemes. Respondents were presented with one NP at a time consisting of a modiﬁer and a noun. One of these two NP elements featured morpheme(s) con- taining confusable graphemes. Respondents were asked to assign the respective element to one of the following categories: active participle, passive participle, nominalized verb or “other”. Additionally, the answer “don’t know” was avail- able. The integration of the tested elements into a NP provided the respondents with information on the functional status of the respective items.

Therefore, one can assume that in the course of the task, respondents make inferences on the functional status of items featuring certain morphemes, and, in a second step, on the interrelation between the morphological gestalt of items and their afﬁliation to grammatical categories and the functional status

1 Many (semi)productive derivational morphemes contain easily confusable graphemes. For example, {н / n}, easily confused with {и / i}, forms (part of) the derivational morpheme for denominal adjectives (e.g. кож-а / kož-a ‘leather’ — кож-ан-ый / kož-an-yj ‘made of leather’), is a marker for the passive participle (e.g. прочита-нн-ый / pročita-nn-yj ‘read’) and, ﬁnally, co- occurs with {и / i} in the sufﬁx -ни- / -ni- for the derivation of deverbal nouns (eg. рисова-ни-е

/ risova-ni-e ‘painting’).

of these categories. This effect of test wiseness is comparable to the effect of analyzing the matches of a corpus query, so we can hypothesize that the best colour markup is the one where respondents a) give most correct answers and

1. display the fastest reaction time.

### Results

277 participants completed the survey. We had to discard 10 of them: 3 due to attested colour vision deﬁciency and 7 because they did not provide the length of their Russian studies. The recorded reaction times varied a lot (up to almost 3 hours). We are interested in the spontaneous effect of colour markup, so we restricted the analysis to instances with a reaction time of 25 sec or less with the exception of the “column” task in the beginner setting (2.2.5) where we restricted reaction time to 60 sec (motivated by the overall reaction time distributions).

Table 1 shows the distribution of instances after cleaning the data that are the basis of the analysis.

*Table 1.* Distribution of the instances

|  |  |  |  |
| --- | --- | --- | --- |
|  | **black** | **colour** | **morph** |
| **beginner** | 406 | 352 | – |
| **advanced** | 2327 | 2071 | 1406 |

For the proportion of correct answers, there is a positive effect of length of study in the coloured setting of task 2.2.1 “spelling”. Respondents with more than 6 months studies provided correct answers with a relative frequency of 0,91 versus respondents with less than 6 months study with 0,75. The coloured setting also produced a higher proportion of correct answers than the black set- ting. The differences are not signiﬁcant though.

Advanced foreign speakers beneﬁt from colour markup see table 2. For beginners we observe only tendencies in some tasks, also for heritage speakers there are some positive tendencies but no signiﬁcant results. Advanced heritage speakers beneﬁt from morpheme marking in the “gaps” and “synonyms” tasks. Overall heritage speakers have a higher proportion of correct answers within 25 seconds than foreign language learners..

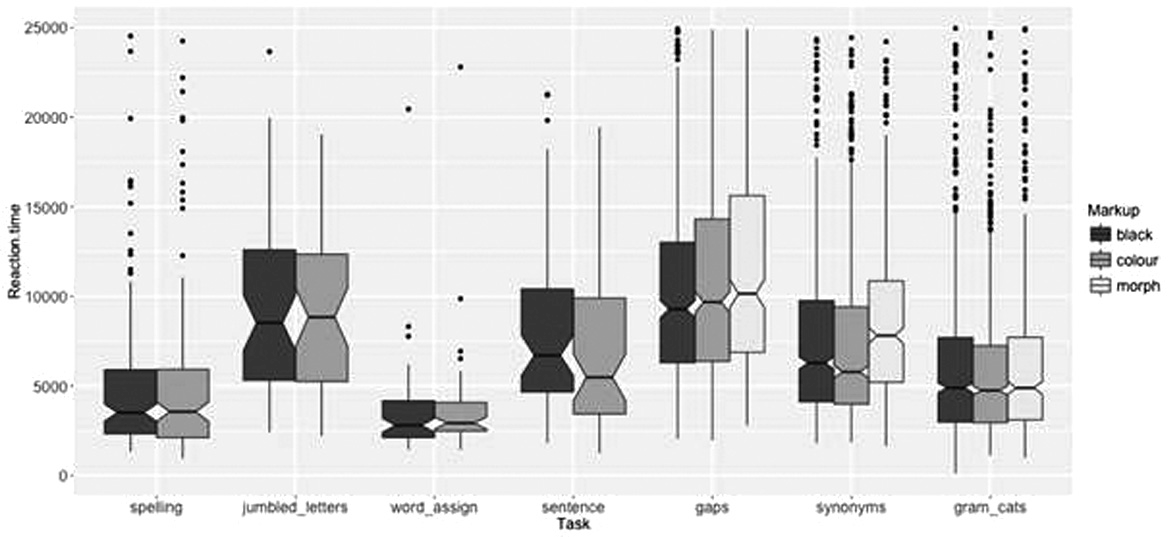
*Table 2.* **Average proportion of correct answers**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **all** | **black** | **colour** | **morph** |
| **foreign** | **beginner** | 0,77 | 0,83 | 0,71 | – |
|  | **advanced** | 0,83 | 0,82 | 0,861 | 0,81 |
| **heritage** | **beginner** | 0,93 | 0,94 | 0,93 | – |
|  | **advanced** | 0,83 | 0,85 | 0,83 | 0,822 |

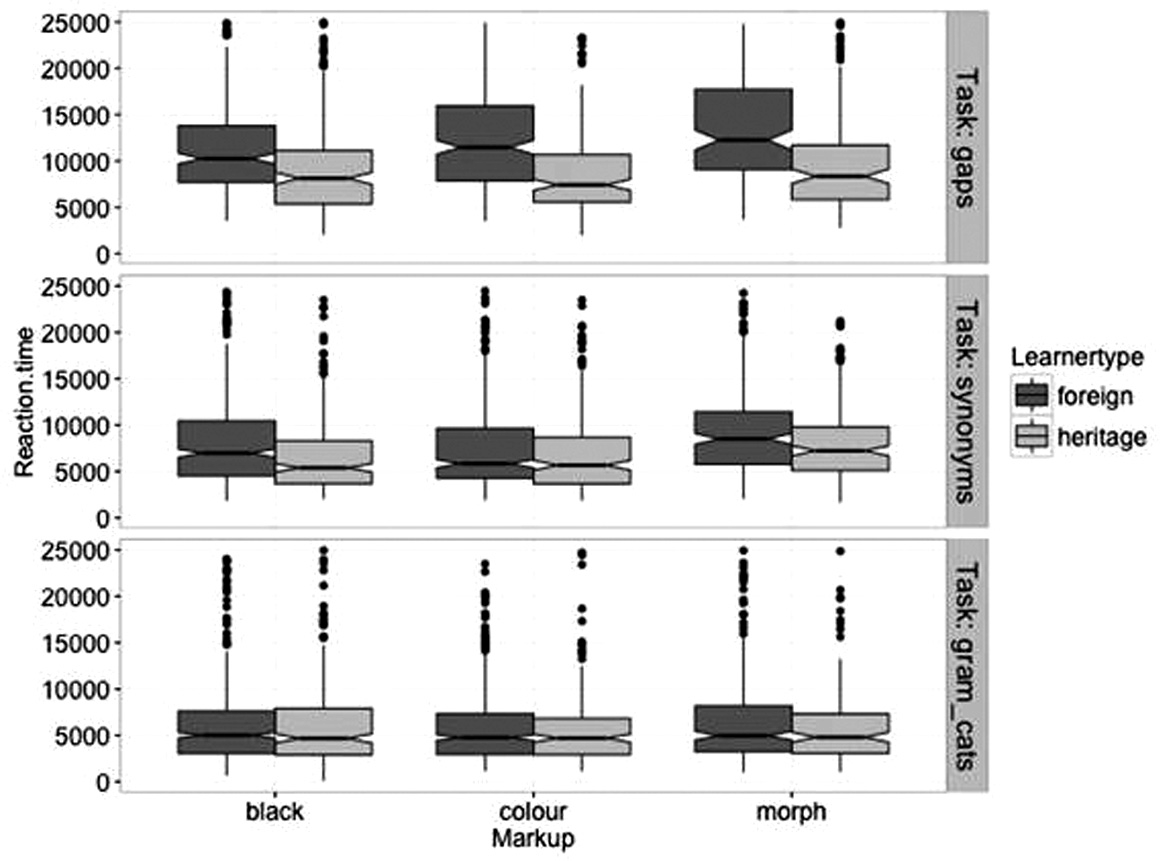
1 The average is based on ﬁve conditions, two of which tested signiﬁcant in a chi-square test in comparison to their black counterparts: gaps\_colour=0,89\*\*\*; gram:cats\_colour=0,86\*.

2 The average is based on: gaps\_morph=0.94; synonyms\_morph=0,89; gram\_cats\_morph=0,62.

Fig. 1 shows boxplots of the reaction times for different tasks given correct answers only. Beginners made faster correct decisions with coloured markup in the “sentence” task only. Advanced learners proﬁted in two out of three tasks from the colour markup. However, the colour marking of morphemes slowed down their reaction times.



*Fig. 1.* Boxplot of the reaction times (correct answers only)



*Fig. 2.* Detailed view: Boxplots of reaction times of advanced learners conditioned on learner type (correct answers only)

Fig. 2 shows the reaction times of the advanced learners conditioned on the markup, the task, and the learner type (foreign language learner and heritage speaker). The medians are marked; non-overlapping notches indicate relevant differences. In general, heritage speakers answered more quickly than foreign language learners. The colour markup reduced the reaction times of the foreign learners in the “synonyms” and “gram\_cats” tasks. For the heritage speakers colour helped in the “gaps” task.

To conclude, colour markup increased the overall proportion of correct answers. The effect on reaction times was mixed and requires further inspection of the individual tasks and individual learners. We also want to include additional independent variables such as the participants’ L1 (German vs. English) and where they acquired their reading skills. The markup of morphemes increased the reaction time which is a negative effect. On the other hand, it showed positive tendencies for the proportion of correct answers in the case of the heritage learners. The results motivate us to further analyse the effects on colour marking of particular linguistic features for different learner groups.

**Literature**

1. *Bobkina, O.G.* (2012), Problemy čtenija i pis’ma. Igry i upražnenija, napravlen- nye na korrekciju disgraﬁi i disleksii. Tavričanka /Primorskij kraj, available at: http:// nsportal.ru/nachalnaya-shkola/raznoe/2012/10/16/problemy-chteniya-i-pisma-igry-i- uprazhneniya-napravlennye-na (accessed 15.10.2016).
2. *Lysko, T.J.* (2013), Upražnenija dlja soveršenstvovanija techniki čtenija v načal’nych klassach. (Tradicionnaja sistema obučenija (1-3)). Ljantor, available at: <http://nsportal.ru/nachalnaya-shkola/chtenie/2013/01/19/uprazhneniya-dlya-soversh-> enstvovaniya-tekhniki-chteniya-v (accessed 15.10.2016).
3. *Mehlhorn G*. (2012), „Das kann ich eh nicht!“ Lernende mit Lese-Rechtschreib- Schwierigkeiten, Praxis Fremdsprachenunterricht, 1/2012, pp 4–6.
4. *Römer U.* (2008), Corpora and language teaching. In: Anke Lüdeling and Merja Kytö (eds.). Corpus Linguistics. An International Handbook (vol. 1). Berlin, pp. 112–130.

***С.Ю. Богданова***

**ВЫЯВЛЕНИЕ КОЛЛОКАТОВ КЛЮЧЕВОГО СЛОВА КАК МЕТОД КОРПУСНОЙ ЛИНГВИСТИКИ**

**Аннотация.** В современной лингвистике все чаще применяется корпусная методология при решении как сугубо лингвистических, так и прикладных задач. В докладе предпринимается попытка обобщить имеющиеся методы корпусной лингвистики, а также на конкретном примере представить один из методов, заключающийся в выявлении коллокатов ключевого слова с целью уточнения особенностей его функционирования.

**Ключевые слова.** Корпусная лингвистика, коллокат, метод.

***S.Yu. Bogdanova***

**REVEALING COLLOCATES OF THE KEYWORD AS A METHOD OF CORPUS LINGUISTICS**

**Abstract.** Corpus methodology has proved to be important in contemporary linguistics, including applied linguistics. However, the exact methods of corpus linguistics are not usually enumerated in the articles on the subject. The author endeavours to classify some of the most often used methods. One of them is revealing collocates of the keyword. For instance, corpus study of the most frequent collocates of synonymous words shows the main differences in their use.

**Keywords.** Corpus linguistics, collocate, method.

Корпусная лингвистика предоставляет как богатый текстовый мате- риал на большом количестве языков, распределенный по жанрам и перио- дам, так и удобный для пользователя интерфейс специальных поисковых систем. Она предлагает набор методов, которые могут применяться при решении различных исследовательских задач. Корпусная лингвистика, в отличие от других областей лингвистики, таких как социолингвистика, психолингвистика и др., в своем названии содержит указание не на то, *что* изучается, а на то, *какая методология* применяется в исследовании [Lindquist 2010: 1].

Методология — это система принципов и способов организации и построения теоретической и практической деятельности, а также учение об этой системе [ФЭС 1983]. В данной статье речь идет о методологии науки как систематическом анализе методов, применяемых для получе- ния научного знания, и тех общих принципах, которыми направляется научное исследование, а точнее, о частной методологии, исследующей методологические проблемы отдельных наук или их узких групп [Фило-

© БОГДАНОВА Светлана Юрьевна. — Bogdanova Svetlana. Иркутский государственный университет, Россия. — Irkutsk State University, Russia. E-mail: [rusjap@mail.ru](mailto:rusjap@mail.ru)

софия: Энциклопедический словарь 2004]. Как методология, современная корпусная лингвистика тесно связана с историей лингвистики как эмпи- рической науки [Lüdeling 2008: v].

Итак, корпусная лингвистика представляет собой частную методоло- гию, объединяющую множество взаимосвязанных методов, которые могут быть использованы учеными разных теоретических направлений. Эти методы широко представлены в различных разделах двухтомника “Corpus Linguistics”, посвященных отдельным аспектам корпусной лингвистики [Lüdeling 2008], однако их классификация до сих пор не предпринималась.

Корпусными называют методы, которые используются при работе с корпусами текстов. К ним относятся разнообразные методы сбора, аннотирования и осуществления поиска в корпусе, а также обработки корпусных данных [Lüdeling 2008: v]. Из данного определения вытекает несколько обособленных групп методов корпусной лингвистики:

1. методы, применяющиеся при создании корпусов текстов;
2. методы, применяющиеся в процессе поиска в корпусе (автоматизи- рованное извлечение информации);
3. методы статистической обработки полученных данных.

Далее обозначим некоторые из этих методов в соответствии с выде- ленными группами.

1. Методы сбора и аннотирования текстового материала являются соб- ственно методами корпусной лингвистики, разрабатываемыми со времени создания первых корпусов текстов, т. е. с 1960-х гг. (см. [Захаров 2013]). В основе аннотирования, в частности, по частям речи (part-of-speech tag- ging), лежали масштабные исследования в области автоматической обра- ботки естественного языка (Natural Language Processing — NLP), главной целью которых было создание искусственного интеллекта.
2. Методы, применяющиеся в процессе поиска в готовом корпусе текстов, в большинстве случаев определяются встроенными в корпусный менеджер поисковыми инструментами: поиск конкретных словоформ, построение конкордансов, метод извлечения ключевой лексики, выбор лексических и синтаксических коллокаций (повторяющихся, воспроизво- димых конструкций) и др. Традиционно корпусы текстов служат «банком» примеров для лингвистических исследований. Корпусная лингвистика предоставляет поисковый инструмент, который обычно позволяет найти убедительные примеры в конкретном корпусе.
3. Среди методов, применяемых в корпусной лингвистике и положи- тельно зарекомендовавших себя в разнообразных лингвистических иссле- дованиях, следует отметить такие важнейшие статистические методы, как частотный анализ, хи-квадрат тест, меры ассоциации (MI, log-likeligood. t-score и др.). Ученые отмечают, что статистический аппарат, применяемый в корпусах текстов, «позволяет пользователям ранжировать результаты

поиска по разным параметрам и задавать пороговые значения, что приво- дит к выдаче наиболее значимой информации» [Хохлова 2010: 3]. Стати- стический вывод позволяет лингвисту на основе отмеченных в отдельном подкорпусе характерных черт сделать обобщения, касающиеся всего языка в целом [Biber 1998: 9]. Существует связь между использованием вычис- лительных и, следовательно, алгоритмических и статистических методов, с одной стороны, и качественными изменениями в результатах научных наблюдений, вытекающими из данного подхода [Tognini Bonelli 2000: 205]. Следует отметить также методы других наук, опирающихся на кор- пусы текстов. Метод опоры на корпус при разработке словаря нашел применение в практике составления толкований значений, например, в Кембриджском международном словаре английского языка. Одним из современных методов анализа является создание и исследование корпуса параллельных текстов, жестко ограниченных тематикой, структурой и объ- емом, но различающихся родными языками их авторов и, как следствие, ракурсами описания одного и того же жестко заданного явления [Беляева 2009: 54]. Корпусные данные содержат важнейшую информацию для нескольких прикладных сфер, таких как обучение языку и лингвистиче- ские технологии машинного перевода, синтеза речи и т. д. [Svartvik 1992]. Так, многие приложения автоматической обработки естественного языка зависят от доступности большого объема текстовых данных. В частности, они используют статистические алгоритмы, обученные на электронных

корпусах текстов [Lüdeling 2008: xi].

В данной статье кратко опишем один из методов, отнесенных нами ко второй группе, а именно автоматическому поиску коллокатов ключевого слова, проведенному на материале корпусов BNC и COCA [(http://corpus.](http://corpus/) byu.edu/ дата обращения 15.09.2016), с целью их дальнейшего анализа. Поисковая система корпуса выявляет важные коллокаты ключевого слова благодаря их частотности и представляет их в виде списка, расположен- ного по уменьшению количества с/у, предоставляя также информацию об их индивидуальной частотности в корпусе и о мере статистической связности MI. На основании полученных количественных данных иссле- дователь может приступать к их качественному описанию, осмысливая и анализируя их. Так, коллокаты можно классифицировать по одному или более параметрам — грамматической позиции, части речи, семан- тической близости и др. Такая классификация, по мнению Е. Тонини Бонелли, в свою очередь, ведет к обобщениям, которые суммируют черты повторяющихся языковых событий. Эти наблюдения обычно позволяют сформулировать гипотезу, объясняющую выявленные зако- номерности. В действительности такая гипотеза представляет собой отнесение значения к той или иной формальной модели. Ожидается, что каждое отдельное значение слова или сочетания слов будет ассоцииро-

ваться с уникальной моделью слов или сочетания слов в их окружении [Tognini Bonelli 2000: 209].

Исследовательская задача заключалась в выявлении семантико- синтаксических особенностей английских темпоральных наречий со зна- чением мгновенности, образующих синонимический ряд. За основу был взят тезис о том, что сочетаемость слова раскрывает особенности семантики синонимических лексических единиц, а также в ряде случаев содержит в себе потенциал для развития их значений. Для разграничения сфер употребления синонимичных наречий применялись разные методы, среди которых особую состоятельность продемонстрировал метод сравне- ния их коллокаций. Сначала в корпусе BNC коллокации выявлялись в диа- пазоне от – 4 до + 4 слов от ключевого слова — темпорального наречия *immediately*. Внимание привлекло наличие коллокатов, принадлежащих к определенной части речи — именам прилагательным, из которых самым частотным было прилагательное *obvious*. При обращении к языковым примерам с данным коллокатом выяснилось, что во всех случаях он непо- средственно примыкает к наречию *immediately* справа и в подавляющем большинстве случаев образует конструкцию *to be + immediately + obvious*: *It was* ***immediately obvious*** *that Wheeler had chosen the earliest possible date to ﬁt his hypothesis* (BNC). Помимо *obvious* (95 с/у)*,* в списке непосред- ственно примыкающих правых коллокатов наречия *immediately* оказалось достаточно большое количество прилагательных (17 из 100 наиболее частотных коллокатов), зафиксированных в той же конструкции в преди- кативной функции. Самыми частотными из них являются *apparent* (88 с/у) и *available* (76 с/у), например: *The reason was* ***immediately apparent*** (BNC)*; As you know, few ﬁles dealing with intelligence matters are* ***immediately avail- able*** (BNC). Детальный анализ примеров позволил уточнить лексическое наполнение конструкции в части употребления в качестве глагола-связки, наряду с глаголом *be*, глаголов *feel, make, go, become*: *He could change his appearance, but not enough to* ***become*** *immediately unrecognizable* (BNC).

Изучая распределение данной конструкции по регистрам, мы пришли

к выводу, что как в британском, так и в американском вариантах англий- ского языка, представленных в корпусах BNC и COCA, чаще всего она встречается в регистре academic, а реже всего — в регистре spoken. Отмечая различия употребления конструкции *to be + immediately + adjec- tive* в этих двух вариантах, следует указать на то, что общее количество предложений с коллокатами-прилагательными из первых 100 коллокатов наречия *immediately* в британском варианте английского языка в 2,5 раза больше, чем в американском (при сравнении нормированных подсчетов). Также обращает на себя внимание отсутствие в американском варианте коллокации *immediately available*, третьей по частотности для данной конструкции в британском варианте. Наиболее частотные коллокаты-

прилагательные располагаются следующим образом: *apparent* (220 с/у),

*clear* (201 с/у), *obvious* (141 с/у).

Применение того же метода показало, что темпоральное наречие со значением мгновенности *instantly* также достаточно частотно в кон- струкциях с прилагательными в предикативной функции. В числе первых 100 примыкающих правых коллокатов наречия *instantly* в BNC отмечено 25 прилагательных, а в COCA — 23, хотя ни одно из них в данной кон- струкции не является частотным, за исключением *recognizable* в COCA (161 с/у) и *recognisable*/*recognizable* в BNC (32 с/у и 15 с/у соответственно). Подобную сочетаемость в единичных случаях демонстрируют тем- поральные наречия *instantaneously* и *at once* из данного синонимического ряда, что говорит об устойчивости данной конструкции в современном английском языке и необходимости ее включения в грамматические спра- вочники и учебные пособия. Что касается решения задачи разграничения синонимов, она решена частично: по критерию сочетаемости с прилага- тельными в предикативной функции темпоральные наречия со значением мгновенности делятся на две группы: 1) *immediately, instantly, instanta- neously, at once*; 2) *forthwith, right away, straight away, promptly*. Для даль- нейшего разграничения необходимо сосредоточить внимание на других

методах исследования или использовать другие критерии.

Исследование, проведенное с использованием корпусного инстру- мента для выявления коллокаций, позволило определить характерные для наречий *immediately* и *instantly* грамматические конструкции с прилага- тельными в предикативной функции, прежде не описанные в грамматиках английского языка, и их семантическое содержание. Это расширило суще- ствующее представление о функционировании английских темпоральных наречий, и в дальнейшем были сделаны выводы о концептуализации временного промежутка, актуализируемого в современном английском языке как наречием *immediately*, так и другими наречиями со значением мгновенности, образующими синонимический ряд (см. [Богданова 2015]).

**Литература**

* 1. *Беляева Л.Н.* (2009), Методы корпусной лингвистики при изучении гло- бального английского языка: этнолингвистический аспект // Иностранные языки в дистанционном обучении: материалы III Международной научно-практической конференции (Пермь, 23–25 апреля 2009 г.), с. 52–59.
  2. *Богданова С.Ю., Коробейникова Н.Н.* (2015), Концептуализация временн*ы*х отношений в высказываниях с темпоральным наречием *right away* // Вестник MГЛУ: Серия Языкознание и литературоведение. М.: ФГБОУ ВПО МГЛУ. Вып. 3 (714), с. 9–22.
  3. *Захаров В.П., Богданова С.Ю.* (2013), Корпусная лингвистика: учебник для студентов гуманитарных вузов. 2-е изд. СПб.: СПбГУ.
  4. Философия: Энциклопедический словарь (2004), М.: Гардарики.
  5. Философский энциклопедический словарь (1983), М.: Советская энцикло- педия.
  6. *Хохлова М.В*. (2010), Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов): авторефер. дис. канд. филол. наук. Санкт-Петербург.
  7. *Biber D., Conrad S., Reppen R.* (1998), Corpus Linguistics: Investigating Lan- guage Structure and Use. Cambridge University Press.
  8. *Lindquist H.* (2010), Corpus Linguistics and the Description of English. Edin- burgh: Edinburgh University Press.
  9. *Lüdeling A., Kytö M.* (eds.), Corpus Linguistics. An International Handbook (2008), Volume 1. Berlin & New York: Walter de Gruyter.
  10. *Svartvik J.* (1992), Directions in Corpus Linguistics // Proceedings of Nobel Symposium 92, Stockholm, 4–8 August 1991. Berlin: Mouton de Cruyter.
  11. *Tognini Bonelli E.* (2000), Corpus Classroom Currency. Darbai ir Dienos, pp. 205–243.

**Literature**

1. *Belyaeva L.N.* (2009), Metody korpusnoy lingvistiki pri izuchenii global’nogo angliyskogo yazyka: etnolingvisticheskiy aspekt // Inostrannye yazyki v distantsionnom obuchenii: materialy III Mezhdunarodnoy nauchno-prakticheskoy konferentsii (Perm’, 23–25 aprelya 2009 g.), s. 52–59.
2. *Bogdanova S.Yu., Korobeynikova N.N*. (2015), Kontseptualizatsiya vremen- nykh otnosheniy v vyskazyvaniyakh s temporal’nym narechiem right away // Vestnik MGLU: Seriya Yazykoznanie i literaturovedenie. M.: FGBOU VPO MGLU. Vyp. 3 (714), s. 9–22.
3. *Zakharov V.P., Bogdanova S.Yu*. (2013), Korpusnaya lingvistika: uchebnik dlya studentov gumanitarnykh vuzov. 2-e izd. SPb.: SPbGU.
4. Filosoﬁya: Entsiklopedicheskiy slovar’ (2004), M.: Gardariki.
5. Filosofskiy entsiklopedicheskiy slovar’ (1983), M.: Sovetskaya entsiklopediya.
6. *Khokhlova M.V.* (2010), Issledovanie leksiko-sintaksicheskoy sochetaemosti v russkom yazyke s pomoshch’yu statisticheskikh metodov (na baze korpusov tekstov): avtorefer. dis. kand. ﬁlol. nauk. Sankt-Peterburg.
7. *Biber D., Conrad S., Reppen R.* (1998), Corpus Linguistics: Investigating Lan- guage Structure and Use. Cambridge University Press.
8. *Lindquist H.* (2010), Corpus Linguistics and the Description of English. Edin- burgh: Edinburgh University Press.
9. *Lüdeling A., Kytö M.* (eds.), Corpus Linguistics. An International Handbook (2008), Volume 1. Berlin & New York: Walter de Gruyter.
10. *Svartvik J.* (1992), Directions in Corpus Linguistics // Proceedings of Nobel Symposium 92, Stockholm, 4–8 August 1991. Berlin: Mouton de Cruyter.
11. *Tognini Bonelli E.* (2000), Corpus Classroom Currency. Darbai ir Dienos, pp. 205–243.

***П.Л. Гроховский, М.О. Михайлова, A.C. Крамскова***

**КОРПУСНО-ОРИЕНТИРОВАННЫЕ МЕТОДЫ ИЗУЧЕНИЯ НАУЧНОГО ФУНКЦИОНАЛЬНОГО СТИЛЯ**

**В ТИБЕТСКОМ ЯЗЫКЕ**1

**Аннотация.** Проведенное исследование направлено на изучение научного функционального стиля на материале научных тибетских текстов разных перио- дов. В ходе научно-исследовательского проекта был составлен корпус объемом 35 тыс. словоупотреблений, который затем был проанализирован с помощью морфологических, лексических и синтаксических квантитативных индексов. Опыт подобного стилостатистического исследования может представлять интерес как для тибетологов, так и для специалистов в прикладной лингвистике.

**Ключевые слова.** Тибетский язык, корпус, квантитативные индексы, научный стиль, функциональная лингвистика.

***P. L. Grokhovskiy. M.O. Mikhailov. A.C. Kramskova***

**CORPUS-ORIENTED METHODS OF STUDYING SCIENTIFIC WRITING IN TIBETAN**

**Abstract.** The research aims at investigating Tibetan scientiﬁc writing formed by Tibetan texts from different historical periods. In course of the study a 35 000 token corpus was created, which was then analyzed with the help of morphological, lexical and syntactic quantitative indices. Such an experience of stylostatistic research may be of interest not only for tibetologists, but for specialists in general and applied linguistics as well.

**Keywords.** Tibetan, corpus, quantitative indexes, scientiﬁc writing, functional linguistics.

### Основной текст

В качестве направления в лингвистике функциональная стилистика оформилась уже к середине XX в. в работах И.А. Бодуэна де Куртенэ, Е.Д. Поливанова, М.М. Бахтина, В.В. Виноградова, Г.О. Винокура.

1 Исследование с использованием методов корпусной лингвистики выполнено при фи- нансовой поддержке РГНФ в рамках научно-исследовательского проекта «Программные средства автоматической обработки текста на современном тибетском языке (морфоло- гический уровень)» (16-04-12016). Авторы также благодарят СПбГУ за финансовую поддержку в рамках научно-исследовательского проекта 2.38.293.2014 «Тибетская пись- менная традиция и современность», позволившую провести содержательное и стилисти- ческое исследование текстов корпуса.

© ГРОХОВСКИЙ Павел Леонович. — Grokhovskiy Pavel. Санкт-Петербургский государственный университет, Россия. — Saint-Petersburg State University, Russia. E-mail: [p.grokhovskiy@spbu.ru](mailto:p.grokhovskiy@spbu.ru)

© МИХАЙЛОВА Мария Олеговна. — Mikhailova Maria. Санкт-Петербургский государствен- ный университет, Россия. — Saint-Petersburg State University, Russia. E-mail: [ariyatan@gmail.com](mailto:ariyatan@gmail.com)

© КРАМСКОВА Анна Сергеевна. — Kramskova Anna. Санкт-Петербургский государственный университет, Россия. — Saint-Petersburg State University, Russia. E-mail: [tibetomongol@gmail.com](mailto:tibetomongol@gmail.com)

За научным стилем как результатом функционирования языка в сфере научного знания априори был признан набор универсальных характе- ристик — отвлеченность, обобщенность, логичность, точность, тер- минологическая насыщенность, развернутость повествования. Однако с развитием теории и метода лингвистической типологии возник интерес к более глубокому изучению отличий между функциональными стилями в разных языках.

Научный функциональный стиль в тибетском языке представляет собой самобытное языковое явление, так как является результатом не про- стого калькирования системы из языка с более высокой научной культу- рой, но синтеза транслируемых идей с собственной давней1 традицией научного знания. Однако проведенное исследование важно не только с тибетологической точки зрения, но и как интересный опыт анализа данных путем применения методов прикладной лингвистики к текстам, организованным в виде корпуса.

Сложность изучения функциональной стилистики тибетского языка обусловлена труднодоступностью необходимых источников, отсутствием крупных корпусов тибетского языка, сложностью оптического распознава- ния тибетского письма и малым числом специалистов, занятых в области. Основным прорывом в данном направлении стало объединение сфор- мулированного еще Дж. Гринбергом метода индексного анализа с техни-

ческими возможностями электронных корпусов [Гринберг 1963].

Проведенное исследование основывалось на применении оптималь- ного набора индексов к корпусу из семи основных текстов общим объемом порядка 35 тыс. словоупотреблений на общую тематику (тибетская пись- менность):

* + трех научных статей из современного тибетского научного журнала

«Тибетология» («Исследование происхождения тибетского письма» 1987 г., «Небольшое замечание по поводу тибетской письменности» 1987 г., «Изложение происхождения и особенностей тибетского письма» 1993 г.)

* + и четырех классических текстов разных эпох, которые в Тибете считались научными в рамках традиционной школы мысли («Волшеб- ный светильник — трактат о письме на тысячу тем и трактат о ручках» в двух частях XI–XII в., «Сокровищница драгоценных сутр и тантр» XVIII–XIX в., «Энциклопедия познаваемого» XIX в., учебник для вузов

«Правила написания тибетской письменности — Зрелище, желанное для всех», 1982 г.).

1 Воспринятыми из Индии еще в VII–VIII вв. традиционными индо-тибетскими науками (*тиб*. rig gnas, *санскр*. vidyā), среди них: грамматика, логика, ремесло, медицина, буддий- ская философия; астрология, поэтика, просодика, синонимия, драматургия.

### Подготовительный этап исследования

Отсканированные копии всех используемых текстов на предвари- тельном этапе проходили через систему оптического распознавания, дальнейшая обработка текста была частично автоматизирована: полу- ченный тибетский текст автоматически транслитерировался, разбивался на токены, проводилась частеречная разметка и лемматизация (с выводом всех возможных вариантов), результаты всех операций были проверены и при необходимости скорректированы.

Тибетский — слоговой язык с графическим обозначением разбиения на слоги, но не на слова. В таких случаях используется либо сегментация по словарю, либо требующая большого и специфически размеченного тре- нировочного материала статистическая система сегментации, основанная на рейтингах встречаемости конкретных слогов в начале, середине и конце слова [Nianwen Xue, 2002]. Для обработки данного корпуса использовался словарь, автоматически составленный на материале корпуса тибетских текстов общего содержания (подготовлен в рамках проекта ФФЛИ C-47

«Базовый корпус тибетского классического языка с русским переводом и лексической базой данных» в 2012–13 гг.) и корпуса тибетских грамма- тических текстов (подготовлен в рамках проекта РФФИ «Пилотная версия электронного корпуса тибетских грамматических сочинений»13-06-00621 А) в 2013–15 гг.), а также словаря глаголов британского тибетолога Н. Хилла1 [Гроховский 2014].

Используемый скрипт (на языке Python) для автоматической токениза- ции, лемматизации и разметки тибетского текста был создан при подготовке корпуса грамматических текстов; он использует словарь и небольшое коли- чество правил. Например, он корректно выделяет единицы, отсутствующие в словаре, если они образованы от найденных в словаре основ с помощью аффиксов из заданного списка — масдары, причастия, порядковые числи- тельные и т. п. Другое необходимое правило заключается в способности выделять грамматические показатели и знаки препинания, которые никак не отделены графически от предшествующего токена.

Сегментацию несколько усложнили западноевропейские знаки пунк- туации, не характерные для ранее обработанных тибетских текстов: необходимо было настроить их автоматическое распознавание системой. В целом полученные результаты похожи на результаты обработки других тибетских корпусов. Так, неразмеченными остались 7–10% токе- нов — для прошлого корпуса эта цифра была 10–15%. Такой невысокий рост в полноте разметки при значительном пополнении словаря свидетель- ствует о примерно одинаковой доле редких слов в текстах. Возможно, это

одна из черт тибетского письменного языка.

1 https://github.com/tibetan-nlp/tibetan-verbs-database (Дата обращения 20.09.2016)

При выведении всех возможных вариантов тегов частеречной разметки для значимых слов предлагалось 1–2 варианта, тогда как для служебных единиц обычно выдавалось 4–6 вариантов разметки (из-за характерной для тибетского языка омонимии односложных слов). Подобную неодно- значность в дальнейшем можно будет отчасти снять введением правил, так как грамматические нормы значительно ограничивают сочетаемость служебных элементов.

Наконец, была проведена разбивка на предложения с опорой на пра- вило (с обязательным условием перепроверки):

а) финитная форма глагола в сочетании с двойным или одинарным *шэ1*

означает конец предложения;

б) даже если использована нефинитная форма глагола перед одиноч- ным шадом, но дальше начинается совершенно иной смысловой блок (например, неоформленный заголовок подраздела), он также считается отдельным предложением.

### Основной этап исследования

Для основного этапа исследования на Python был написан модуль, состоящий из скриптов для разных математических манипуляций — под- счета встречаемости элементов с определенными критериями (длиной, тегом, леммой, находящихся в определенном положении), их соотношения между собой или с какими-то другими цифрами, сортировки по заданным параметрам (например, частоте или длине), вывода примеров. В сочетании с уже имевшимся модулем, открывавшем файлы корпуса в удобном для работы формате и осуществлявшим базовые операции с корпусом, а также ресурсами имеющихся для Python библиотек наподобие NumPy и возмож- ностью прямо из консоли записать готовый результат в нужном формате в текстовый или табличный файл, получилась чрезвычайно эффективная и гибкая система, позволяющая извлечь любые данные, обработать их и записать максимально наглядно.

Кроме того, модуль содержит скрипт, позволяющий переводить гото-

вые размеченные тексты из рабочего формата в форматы, используемые распространенными программами для работы с корпусами. Сейчас он содержит методы перевода в .xml (с xml-разметкой) и в формат тексто- вого файла, в котором слово и теги записываются через косую черту (его использует, например, модуль NLTK).

После обращения к опыту других исследователей, работавших в индексном методе, и собственных размышлений, был разработан опти- мальный список индексов и проведены необходимые вычисления [Ягу- нова 2014].

1 *тиб*. shad — специальный пунктуационный знак, обозначающий синтаксическую паузу.

Показатели глагольности и субстантивности (отношения глаголов и существительных к общему числу слов) оказались типичны для стан- дартного научного стиля и практически одинаковы для всех текстов (в среднем 0,25 и 0,46 соответственно).

Вычисление индекса местоименности выявило результат, обратный тому, который обычно показывают тексты научного стиля в других язы- ках — высокие показатели употребления местоимений в текстах совре- менного литературного языка, что объясняется сближением последнего с разговорной нормой и, наоборот, господством в классических сочине- ниях принципа «эстетики тождества».

Большая доля отглагольных имен в современных текстах также отра- жает изменившуюся систему языка, в которой повысилась роль номина- тивных форм глагола.

Такой показатель, как индекс связности (совместная доля дискур- сивных слов, союзов и предлогов), указывает на более высокую степень семантико-грамматического единства в современных текстах, что также не характерно для научного стиля в европейских языках.

Индекс терминологической насыщенности рассчитывался исходя из списка общеупотребительных для всех текстов терминов (как общенауч- ных — «*исследование*», «*причина*», — так и узкоспециальных — («*мате- риализм*», «*предложение*»). Более высокие показатели характерны для современных научных статей.

Частотность цитирования рассчитывалась по специальному показа- телю — *тиб*. zhes, производному от глагола «говорить», которым в тибет- ском языке оформляется прямая речь, — намного выше он у современных текстов.

Традиционная нунктуация используется в современных текстах про- гнозируемо реже, поскольку средняя длина предложения в них больше, а также применяются современные знаки препинания (кавычки, скобки, сноски).

Таким образом, применение корпусно-ориентированных методов анализа позволило классифицировать обработанные тексты и сделать отдельные выводы относительно характера научного функционального стиля в тибетском языке, обладающего особенностями, традиционно приписываемыми научному стилю, но с учетом специфики тибетского литературного языка, а также относительно динамики развития тибетского научного функционального стиля.

Дальнейшую работу планируется вести в направлении как расшире- ния исследовательской базы (новые тексты), так и совершенствования методологических подходов (новые индексы, дополнительная метараз- метка).

**Литература**

1. *Nianwen Xue* (2002), Combining Classiﬁers for Chinese Word Segmentation. Proceedings of the ﬁrst SIGHAN workshop on Chinese language processing, Strouds- burg, vol. 18, pp. 1–7.
2. *Гринберг Дж.* (1963), Квантитативный подход к морфологической типоло- гии языков. Новое в лингвистике. Вып. 3, с. 60–94.
3. *Гроховский П.Л., Захаров В.П., Попов А.М., Смирнова М.О., Хохлова М.В*. (2014), Лексическая база корпуса тибетских грамматических сочинений. Тех- нологии информационного общества в науке, образовании и культуре: сборник научных статей. Материалы XVII Всероссийской объединенной конференции

«Интернет и современное общество» IMS-2014. СПб., c. 277–282.

1. *Ягунова Е., Поспелова А.* (2014), Опыт применения стилевых и жанровых характеристик для описания стилевых особенностей коллекций текстов. Новые информационные технологии в автоматизированных системах: материалы семнад- цатого научно-практического семинара. М.: ИПМ им. М.В. Келдыша, c. 347–357.

**Literature**

1. *Nianwen Xue* (2002), Combining Classiﬁers for Chinese Word Segmentation. Proceedings of the ﬁrst SIGHAN workshop on Chinese language processing, Strouds- burg, vol. 18, pp. 1–7.
2. *Grinberg Dzh*. (1963), Kvantitativnyi podkhod k morfologicheskoy tipologii yazykov. Novoe v lingvistike. Vyp. 3, s. 60–94.
3. *Grokhovskiy P.L., Zakharov V.P., Popov A.M., Smirnova M.O., Khokhlova M.V*. (2014), Leksicheskaya baza korpusa tibetskikh grammaticheskikh sochineniy. Tekhnologii informatsionnogo obshchestva v nauke, obrazovanii i kul’ture: sbornik nauchnykh statey. Materialy XVII Vserossiyskoy ob’yedinennoy konferentsii «Internet i sovremennoe obshchestvo» IMS-2014. SPb., s. 277–282.
4. *Yagunova E., Pospelova A*. (2014), Opyt primeneniya stilevykh i zhanrovykh kharakteristik dlya opisaniya stilevykh osobennostey kollektsiy tekstov. Novye infor- matsionnye tekhnologii v avtomatizirovannykh sistemakh: materialy semnadtsatogo nauchno-prakticheskogo seminara. M.: IPM im. M.V. Keldysha, s. 347–357.

***В.П. Захаров, В. Бенко, М. В. Хохлова***

**КТО ИЩЕТ, ТОТ НАЙДЕТ. ПОИСКОВАЯ ФУНКЦИОНАЛЬНОСТЬ КОРПУСОВ**

**РУССКОГО ЯЗЫКА**

**Аннотация.** Лингвисты, изучающие русский язык, стоят перед выбором не только того, какой из имеющихся корпусов представляет лучшую выборку данных языка, но и перед вопросом, какой из корпусных менеджеров обладает достаточными поисковыми и статистическими функциями для осуществления нужных исследовательских задач. В статье обсуждается функциональность кор- пусного менеджера, являющегося интегральной частью проекта Национального корпуса русского языка.

**Ключевые слова.** Корпуса русского языка, функциональность корпусных менеджеров, языки запросов, конкорданс.

***V.P. Zakharov, V. Benko, M.V. Khokhlova***

**SEARCH FUNCTIONALITY**

**OF RUSSIAN LANGUAGE CORPORA**

**Abstract.** Linguists studying the Russian language face a choice not only which of existing corpora of Russian is the best selection of language data, but also which of corpus managers have sufﬁcient search functions for the implementation of relevant research problems. The article discusses the functionality of the corpus manager which is an integral part of the Russian National Corpus.

**Keywords.** Corpora of the Russian language, corpus manager functionality, query languages, concordance.

### Введение

Можно сказать, что все современные лингвистические исследования и работы по составлению словарей и грамматик так или иначе ориенти- рованы на использование представительных корпусов текстов. Развитие современных интеллектуальных программных систем, предназначенных для обработки текстов на естественном языке, также требует большой экспериментальной лингвистической базы. Спрос на корпусные данные совпал с появлением соответствующих технических возможностей.

© ЗАХАРОВ Виктор Павлович. — Zakharov Victor. Санкт-Петербургский государственный университет, Россия. — Saint-Petersburg State University, Russia. E-mail: [v.](mailto:v.zakharov@spbu.ru)[zakharov@spbu.ru](mailto:zakharov@spbu.ru)

© БЕНКО Владимир. — Benko Vladimír. Институт языкознания им. Л. Штура, Словакия. — Ľ. Štúr Institute of Linguistics, Slovak Academy of Sciences, Slovakia. E-mail: vladob@juls. savba.sk

© ХОХЛОВА Мария Владимировна. — Khokhlova Maria. Санкт-Петербургский госу- дарственный университет, Россия. — Saint-Petersburg State University, Russia. E-mail: [m.khokhlova@spbu.ru](mailto:m.khokhlova@spbu.ru)

Помимо возможности получения обширного текстового материала, релевантного запросу (конкорданс), корпусы могут использоваться для получения разнообразных справок и статистических данных о языковых и речевых единицах. Репрезентативный массив языковых данных за определенный период позволяет изучать динамику процессов изменения лексического состава языка, проводить анализ лексико-грамматических характеристик в разных жанрах и у разных авторов и т. д.

В мире корпусная лингвистика как особое направление сложилась к началу 1990-х гг. Безусловно, ключевым моментом развития корпусной лингвистики в России стало создание Национального корпуса русского языка (2004) (http://ruscorpora.ru). О востребованности корпуса свиде- тельствуют многочисленные исследования и публикации, подготовленные на его основе. Однако корпуса русского языка сегодня создаются не только

«в недрах» НКРЯ и не только в России.

Задача данной статьи — дать краткий обзор поисковой функциональ- ности корпусов русского языка. Обзор корпусов русского языка различного типа дан в работе [Захаров 2015].

Для поиска в корпусе используются специализированные поисковые системы — корпусные менеджеры. В общем случае современный корпус- ный менеджер должен:

* + строить конкорданс;
  + искать контексты не только по отдельным словам, но и словосоче- таниям;
  + сортировать списки по нескольким критериям, выбираемым поль- зователем;
  + давать возможность отображать найденные словоформы в расши- ренном контексте;
  + давать статистическую информацию по отдельным элементам корпуса;
  + отображать леммы, морфологические характеристики словоформ и метаданные (библиографические, типологические), что зависит от глу- бины разметки корпуса;
  + сохранять и распечатывать результаты;
  + работать как с отдельными файлами, так и с корпусами, неограни- ченными по размеру;
  + быстро обрабатывать запросы и выдавать результаты;
  + быть легким (интуитивно понятным) в использовании как для опыт- ного, так и для начинающего пользователя.

Разные корпусные службы обслуживаются разными корпусными менеджерами (CQP, CQPweb, NoSketch Engine, KonText, Sketch Engine и др.) и предоставляют пользователю разный по объему и исполнению набор функций. Оговоримся, что сформировать сводный перечень функ- ций современных корпусных менеджеров не так просто. Еще сложнее

привести их к одному «знаменателю». Сложности понятны: одни и те же функции в разных службах могут называться по-разному. В разных служ- бах эти функции могут решаться в разных объемах и на выходе выдавать похожие, но не идентичные результаты.

В данной статье мы рассмотрим поисковые возможности корпусных менеджеров на примере Национального корпуса русского языка (НКРЯ). Корпус доступен в сети с 2004 г. и отвечает критерию репрезентативности и другим требованиям, предъявляемым к современным корпусам [Нацио- нальный корпус… 2005; Национальный корпус… 2009]. НКРЯ в настоя- щее время, помимо основного корпуса, включает следующие подкорпусы: газетный, диалектный, синтаксический, обучающий, поэтический, устный, мультимедийный, исторический и параллельные. Все подкорпуса обслу- живаются одним и тем же корпусным менеджером с модификациями во входном интерфейсе, которые определяются особенностями разметки подкорпусов. В числе целого ряда особенностей, отличающих НКРЯ от других корпусов и, прежде всего, от других национальных корпусов, отметим в первую очередь три: наличие мультимедийного подкорпуса [(http://ruscorpora.ru/search-murco.html),](http://ruscorpora.ru/search-murco.html)) семантической разметки (http:// ruscorpora.ru/corpora-sem.html) и статистического инструмента проведения диахронических исследований «Графики» [(http://ruscorpora.ru/ngram.html).](http://ruscorpora.ru/ngram.html)) В качестве корпусного менеджера НКРЯ используется поисковая система Яндекс, специально адаптированная под задачи НКРЯ [Аброс- кин 2009]. Поисковый аппарат НКРЯ хорошо документирован (см. пошаговая инструкция по поиску [http://ruscorpora.ru/instruction-main.](http://ruscorpora.ru/instruction-main)

pdf и контекстные меню).

Основной результат, выдаваемый из НКРЯ пользователю, — это кон- корданс. Кроме того, НКРЯ выдает разнообразную статистику, как сопут- ствующую запросу, так и «аккумулированную». С точки зрения последней НКРЯ дает одну из наиболее полных статистических картин распределе- ния текстов по элементам метаданных, предоставляя сведения о распре- делении текстов по подкорпусам, по видам текстов, по жанрам, по сфере функционирования, по тематике, по хронологическим периодам. Все эти распределения можно посмотреть в терминах числа текстов, предложений, числа и процентной доли словоупотреблений по данной характеристике по отношению к общему объему. Эти характеристики НКРЯ суммарно выложены непосредственно на сайте (см. статистику <http://ruscorpora.ru/> corpora-stat.html). Также можно получить частотные списки словоформ и n-грамм [(http://ruscorpora.ru/corpora-freq.html).](http://ruscorpora.ru/corpora-freq.html))

Корпусной менеджер НКРЯ обладает достаточно большими и гибкими

поисковыми возможностями, которые основываются как на поисковой силе Яндекс-сервера, так и на богатой лингвистической и металингви- стической разметке НКРЯ.

Все функции корпусного менеджера можно разделить на три крупных части: язык запросов к корпусу (табл.), средства управления выходным интерфейсом и сервисные функции.

*Таблица*. Язык запросов корпусного менеджера НКРЯ

|  |  |  |
| --- | --- | --- |
| **Функции** | **Наличие** | **Комментарий** |
| Способ задания запроса: форма | + | Внутри системы используется специальный язык |
| Способ задания запроса: язык регулярных выражений | – |  |
| Поисковые элементы: |  |  |
| слова (леммы) | + |  |
| словоформы | + | Через точный поиск или путем задания граммем |
| грамматические признаки | + |  |
| словосочетания | + |  |
| метаданные | + | Путем предварительного создания подкорпуса |
| семантические признаки | + | По частям речи |
| Поиск по n-граммам | + | Выдается статистика для каждого члена словоизменительной парадигмы |
| Средства морфологической нормализации/денормализации текстовых элементов запроса | + |  |
| Усечение | + | Все 3 типа (слева, внутри, справа), как по лемме, так и по словоформе |
| Операторы расстояния и позиции (словосочетания) | + | Между соседними словами рас- стояние считается равным 1 и т. д. Отрицательное число в поле  «Расстояние» соответствует обратному порядку слов |
| Операторы позиции в тексте | + | В начале предложения, в конце предложения, перед (после) любым или перед конкретным знаком препинания |
| Повтор поискового элемента в тексте | + | Элемент может быть как лексемой, так и граммемой |
| Учет границ предложения | частично | через дополнительные признаки (знаки препинания и позиции  в предложении) |
| Булевы операторы (конъюнкция, дизъюнкция, отрицание) | + | Дизъюнкция и отрицание |
| Сложные логические запросы | – |  |
| Подсказка при вводе запросов | + | Лексемы |
| Поиск по морфемам | + | Включая корень. С учетом или без учета чередования |

**Средства управления выходным интерфейсом**. Основной результат поиска в корпусе — это конкорданс, выдаваемый в двух формах: каждая единица выдачи — цельное предложение (по умолчанию) и KWIC, единица выдачи на запрос — строка определенной длины. Есть возможность включить режим выдачи «С ударениями». Имеется возможность задавать количество примеров на одной веб-странице выдачи. Для KWIC- выдачи можно задать количество слов в контексте. В ручном режиме для выбранного слова можно просмотреть данные морфологической и семантической разметки. От стандартной выходной формы можно перейти к выдаче статистических данных по метаатрибутам. Имеется функция «Распределение по годам», которая показывает диахронический график частоты употребления заданного поискового запроса по годам, начиная с 1800 г. Та же функция существует как отдельный поисковый режим («График») с более богатыми возможностями.

**Сервисные функции**. Лингвистов-исследователей сегодня все в боль-

шей степени интересуют функции по дополнительной, можно сказать, интеллектуальной обработке корпусных данных. Сервисные функции и функциональность корпусных служб во многом базируются на обра- ботке статистических данных. Статистику в корпусных службах условно можно поделить на внешнюю и внутреннюю.

Под внешней статистикой мы понимаем данные, которые выдаются непосредственно пользователю как отчет, как справка. В НКРЯ эта ста- тистика представлена достаточно широко в виде справочных разделов на сайте. НКРЯ среди всех корпусных систем дает одну из наиболее полных статистических картин распределения текстов по элементам метаданных. НКРЯ предоставляет сведения о распределении текстов по подкорпусам, по видам текстов, по жанрам, по сфере функционирования, по тематике, по хронологическим периодам. Все эти распределения можно посмотреть в терминах числа текстов, предложений, числа и процент- ной доли словоупотреблений по данной характеристике по отношению к общему объему.

Внутренняя статистика — это те данные, которые используются для

«внутренних» подсчетов и выдачи новых данных, непосредственно в кор- пусе не заложенных, например, количественное вычисление устойчивости сочетаний в тексте, парадигматическая семантическая кластеризация лексических единиц, выявление опорных ключевых слов текста и т. п., т. е. для получения собственно лингвистических результатов.

К сожалению, в НКРЯ эти сервисные функции практически не пред- ставлены. Но это уже не имеет непосредственного отношения к функциям поиска. Что же касается поисковой функциональности НКРЯ, то его корпусной менеджер предоставляет достаточно богатый и гибкий набор возможностей для поиска.

**Литература**

1. *Аброскин А. А* (2009), Поиск по корпусу: проблемы и методы их решения // Национальный корпус русского языка: 2006–2008. Новые результаты и перспек- тивы. СПб.: Нестор-История, с. 277–282.
2. *Захаров В.П.* (2015), Корпуса русского языка. Труды Института русского языка им. В.В. Виноградова. T. 6, М., с. 20–64.
3. Национальный корпус русского языка: 2003–2005. Сборник статей. М.: Индрик, 2005.
4. Национальный корпус русского языка (2006–2008). Новые результаты и перспективы. СПб.: Нестор-История, 2009.

**Literature**

1. *Abroskin A.A*. (2009), Poisk po korpusu: problemy i metody ikh resheniya // Natsional’nyi korpus russkogo yazyka: 2006–2008. Novye rezul’taty i perspektivy. SPb.: Nestor-Istoriya, s. 277–282.
2. *Zakharov V.P*. (2015), Korpusa russkogo yazyka. Trudy Instituta russkogo yazyka im. V.V. Vinogradova. T. 6, M., s. 20–64.
3. Natsional’nyi korpus russkogo yazyka: 2003–2005. Sbornik statey. M.: Indrik, 2005.
4. Natsional’nyi korpus russkogo yazyka (2006–2008). Novye rezul’taty i per- spektivy. SPb.: Nestor-Istoriya, 2009.

***К.Ю. Игнатов***

**«ОКСФОРДСКИЙ СЛОВАРЬ ЦИТАТ» КАК КОРПУС: РАЗМЕТКА ПО ЖАНРАМ ВЫСКАЗЫВАНИЙ**

**Аннотация.** В статье представлены результаты сплошного анализа «Оксфорд- ского словаря цитат» как корпуса, в котором каждая цитата из одной словарной статьи рассматривается как текст. Для разметки корпуса использованы указанные в словаре сведения (автор, национальная принадлежность, дата высказывания, язык и т. д.). В аннотацию корпуса введен дополнительный параметр — жанр высказывания; также уточнены словарные данные. Приводятся обобщенные количественные результаты анализа и формулируются возможные применения исследования.

**Ключевые слова. «**Оксфордский словарь цитат», цитата, разметка корпуса, статистический анализ, жанр высказывания.

***K.Y. Ignatov***

***OXFORD DICTIONARY OF QUOTATIONS* AS A CORPUS: ANNOTATING THE GENRES OF QUOTATIONS**

**Abstract.** The paper presents the results of a comprehensive analysis of the *Oxford Dictionary of Quotations* in which the dictionary is treated as a corpus, with each quota- tion from a dictionary entry representing a separate text. The corpus is tagged with the information that the dictionary contains (authors, their nationality, date of the quota- tions, their language, etc.). The relevant information in the dictionary is updated, and a new tag, namely, quotation genre is added to the annotation of the corpus. Generalised qualitative results are provided, and suggestions of their interpretation and use are put forward.

**Keywords.** “Oxford Dictionary of Quotations”, quotation, corpus tagging, statisti- cal analysis, utterance genre.

Задача корпусного подхода к словарю крылатых выражений была вызвана необходимостью определения «системы отсчета» для анализа интертекстуальных ссылок как показателя авторского стиля в художе- ственном произведении. Интертекст понимается как совокупность прямых отсылок и неявных аллюзий к внетекстовой реальности художествен- ного произведения: географическим названиям, литературным реми- нисценциям, историческим фактам, политическим именам и событиям, социальным явлениям, объектам культуры и бытовым реалиям. Важной характеристикой авторского стиля может быть не только способ органи- зации и функции отсылки, но и ее референт. Для объективного анализа

© ИГНАТОВ Кирилл Юрьевич. — Ignatov Kirill. Московский государственный университет им. М.В. Ломоносова, Россия. — Moscow State Lomonosov University, Russia. E-mail: [kirill72@yahoo.com](mailto:kirill72@yahoo.com)

стиля одного или сопоставительного анализа нескольких авторов нужна некая «нейтральная» система референтов, характеризующая языковую жизнь общества в целом.

Для решения задачи был выбран «Оксфордский словарь цитат» как наиболее авторитетный источник информации о часто используемых кры- латых выражениях в английском языке. Словарь, впервые увидевший свет в 1941 г., успешно выдержал семь изданий, причем принципы формиро- вания каждого из них во много отражали меняющийся мир англоязычной культуры. Изначально цитаты отбирались экспертами; сегодня словарь составляется на основе работы с современными печатными текстами и текстами интернет, а также с привлечением Оксфордского националь- ного корпуса английского языка. Расширялся и спектр справочной инфор- мации, содержащейся в словаре. Современный словарь предназначен для использования в качестве реферативного источника данных о цитатах, используемых в современной речи и прошлом (в отношении источников, которые не утратили релевантности по сей день, прежде всего англоязыч- ных изданий классической литературы).

Данное исследование проводилось на основе последнего издания

[Oxford Dictionary of Quotations 2009]. В нем содержатся сведения о национальной принадлежности автора, годах жизни, основных видах деятельности, дате создания произведения речи и параллели (темати- ческие, исторические, биографические и т. д.) с содержанием других словарных статей.

В лексикографии отмечалась значимость подобного словаря как источника культуроведческой информации об англоговорящем мире. Так, отечественные исследователи пишут, что «словарь цитат является одним из источников социокультурной, лингвопрагматической информации о цитатах, средством формирования лингвокультурной компетенции» [Гриднева 2008, 20], указывая также, что для более полного использования потенциала словаря необходимо расширить его систему дополнительной информации [Гриднева 2008, 21]. Решению этой задачи также поможет корпусный подход к словарю с соответствующей переработкой системы аннотации.

К самому важному изменению относится добавление разметки по параметру «жанр высказывания» в аннотации. Было выделено 12 тексто- вых жанров: нехудожественные письменные тексты, художественная проза, драма, поэзия, христианские гимны, песни (включая оперы и мюзиклы), кино- и телефильмы, тексты средств массовой информации (включая телепрограммы, но информационные, а не развлекательные), устные тексты, мемуары и заметки, переписка и предсмертные высказыва- ния. Такое разбиение выкристаллизовалось в результате непосредственной работы со словарем.

К группе нехудожественных письменных текстов относятся книги, брошюры и статьи в специализированных изданиях, опубликованные дис- сертации, официальные договоры и документы (включая дипломатическую переписку), официальные судебные материалы и решения, а также данные под присягой показания и прошения об отставке или памфлеты, даже если они составлены в виде письма (например, «Письмо шерифу Бристоля» Эдмунда Бёрка) или опубликованы в прессе (как было с британским мини- стром по вопросам занятости населения Норманом Фаулером, который своим письмом, опубликованным в газете «Гардиан», ввел в оборот выра- жение «проводить больше времени с семьей» как повод для отставки).

В тексты художественной прозы включены рассказы, новеллы, романы и сказки. К драмам относятся все пьесы в стихах и прозе, включая те, которые впоследствии были переработаны в киносценарии (например, произведения Ноэла Кауарда). Группа «Поэзия» состоит из стихотворе- ний, эпиграмм, баллад, историй и рассказов в стихах, а также псалмов (за исключением текста Библии).

Отдельную группу составляют христианские гимны. К песням отно- сятся также и тексты мюзиклов, либретто опер и оперетт. К фильмам относятся все цитаты из сценариев (точнее — монтажных листов) кино- и телевизионных фильмов, даже если изначально они были написаны для постановки в театре, но все же стали популярными после экранизации. В эту же группу включены развлекательные телевизионные программы (например, «Летающий цирк Монти Пайтона»).

К текстам средств массовой информации относятся цитаты их новост- ных, информационных и публицистических программ телевидения, радио, общественно-политических и развлекательных газет и журналов, включая интервью, пресс-конференции, опубликованные в газетах и журналах письма. К устным высказываниям относятся речи, лекции, проповеди, устные выступления и обращения, которые, хотя и были впоследствии зафик- сированы в печатном или ином виде (например, записи скетчей Вуди Аллена, выпущенные на пластинках), но изначально предназначались для восприятия на слух. Также в эту группу включены собственно устные высказывания, которые были письменно зафиксированы их адресатами

и слушателями.

К мемуарам и заметкам отнесены дневники, воспоминания, завеща- ния, разрозненные заметки и выписки. К переписке относятся письма, переданные по телеграфу сообщения (например, первое сообщение Сэмю- эла Морзе) и телеграммы, в том числе обращения к группам лиц, напри- мер, партиям, съездам, народам мира и т. д. (как случилось с письмом к Г.Е. Зиновьева к Компартии Великобритании, опубликованном в газете

«Таймс»). К предсмертным высказываниям относятся последние прижиз- ненные слова авторов, так и зафиксированные в словаре.

Жанр произведений всегда уточнялся по энциклопедиям, поскольку во многих случаях он неочевиден из приводимого названия произведения. В случае затрудненной атрибуции жанра цитаты в качестве ориентира использовались указания словаря, которые одновременно проверялись по специализированным энциклопедическим источникам. Например, цитаты, связанные с Питером Пеном, литературным героем Джеймся Барри, могут быть отнесены как к романам автора, так и к созданным на их основе пьесам, мюзиклам и фильмам.

Также каждой словарной статье сопоставлено количество цитат дан- ного источника (в соответствии с организацией словаря это может быть не только отдельный автор, но и неавторский текст или текст с коллектив- ным автором, например Библия, Упанишады, Тора и т. д.).

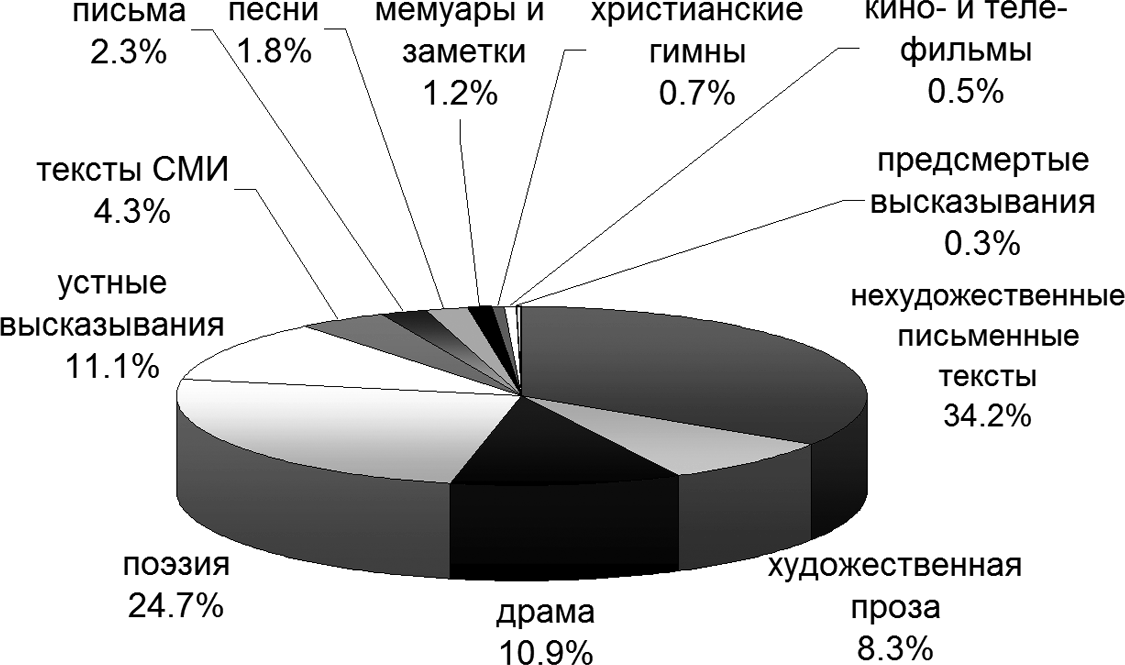
Кроме того, дополнительно введена усредненная дата цитат в каждой словарной статье для каждого источника, в том числе коллективных или анонимных авторов. Такая разметка позволит выделять наиболее продук- тивные исторические периоды с точки зрения «провербиальности» как для групп авторов, национальных образований, так и для жанров. Было принято решение учитывать объективную дату создания высказывания, а не время актуализации цитаты в языке. Например, творчество Мацуо Басё относится к XVII в., но крылатые выражения из хайку японского поэта стали популярны только после их перевода на английский язык уже в XX в. Тем не менее в аннотации корпуса указывается средняя дата создания высказываний — 1670 г.

Как показала работа, необходимо было провести уточнение указанной в словаре хронологии и определения рода деятельности, а также нацио- нальной принадлежности авторов. Сохраняя содержащиеся в словаре данные, также предпринята попытка их уточнения (указания ‘feminist’ ‘human right activist’ ‘abolitionist’ и т. д. объединены общим указанием ‘activist’) и определенной унификации (например, ‘dramatist’ и ‘playwright’ определяют один и тот же вид деятельности).

В результате сплошного анализа словаря выделено 3 454 словарные статьи, в которых содержатся 20 193 цитаты (статьи, лишь отсылающие читателя к другим статьям, не учитывались).

Статистические результаты, характеризующие словарь как корпус с точки зрения жанров высказываний, представлены на рис.

Основное возможное применение размеченного таким образом кор- пуса — исследование крылатых выражений английского языка не только по тематике высказывания, что делалось и ранее с помощью индекса в печатном издании или путем использования его электронной версии, но и учет жанровых особенностей высказываний. Например, можно поставить вопрос о хронологических закономерностях в распределении соотношений проза / поэзия, художественная / нехудожественная литера-



*Рис.* Распределение текстов цитат по жанрам

тура, устная / письменная речь и т. д. в отдельные исторические периоды. Более того, распространяя подобный подход на предыдущие издания словаря (а разметка проведена в седьмом издании), можно выявить диа- хронические изменения в культурной значимости отдельных авторов, жанров и, собственно, цитат, подобно тому, что в случае Национального корпуса русского языка исследователи называют «микроисторической» лингвистикой, «в центре внимания которой находятся не глобальные изменения в истории языка, а изменения менее масштабные, занимающие десятилетия» [Плунгян 2005: 306]. С другой стороны, подобно тому как в учебных корпусах «накопление текстов одного и того же автора (авторов) в течение некоторого времени ... позволяет представить процесс овладения языком в динамике» [Камшилова 2014: 104], составление диахрониче- ского корпуса из всех изданий словаря цитат даст богатый материал как для лексикографов, так и для лингвистов и литературоведов, поскольку позволит проследить усвоение или отторжение лингвокультурой того или иного высказывания или автора.

**Литература**

* 1. *Гриднева Н.А.* (2008), Словарь цитат в парадигме современной английской лексикографии (на материале английских словарей цитат): автореф. дис. канд.

филол. наук. Самара, 31 с.

* 1. *Камшилова О.Н.* (2014), LC-технологии в исследовании и освоении языка / Известия Российского государственного педагогического университета им. А.И. Герцена. № 168, с. 103–110.
  2. *Плунгян В.А.* (2005), Зачем мы делаем Национальный корпус русского языка? / Отечественные записки. № 2, с. 296–308.
  3. *Oxford Dictionary of Quotations.* (ed. by Elizabeth Knowles) (2009), 7th ed. Oxford: OUP.

**Literature**

1. *Gridneva N.A.* (2008), Slovar’ tsitat v paradigme sovremennoy angliyskoy lek- sikograﬁi (na materiale angliyskikh slovarey tsitat): avtoref. dis. kand. ﬁlol. nauk.

Samara, 31 s.

1. *Kamshilova O.N*. (2014), LC-tekhnologii v issledovanii i osvoenii yazyka / Izvestiya Rossiyskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gert- sena. № 168, s. 103–110.
2. *Plungyan V.A*. (2005), Zachem my delaem Natsional’nyi korpus russkogo yazyka? / Otechestvennye zapiski. № 2, s. 296–308.
3. *Oxford Dictionary of Quotations.* (ed. by Elizabeth Knowles) (2009), 7th ed. Oxford: OUP.

***О.Н. Камшилова, Е.Д. Смульская***

**ПЛАНИРОВАНИЕ ЛОНГИТЮДНОГО ИССЛЕДОВАНИЯ: БОЛЬШИЕ ПРОБЛЕМЫ МАЛЕНЬКОГО КОРПУСА**

**Аннотация.** В докладе рассматриваются возможности применения лонги- тюдного метода в лингвистических исследованиях, в частности строительство и назначение лонгитюдных учебных корпусов текстов. Анализируются проблемы валидности лонгитюдного корпусного исследования на примере планирования и разработки лонгитюдного корпуса текстов петербургских школьников, изучаю- щих английский язык.

**Ключевые слова.** Лонгитюд, лонгитюдный учебный корпус текстов, валид- ность.

***O.N. Kamshilova, E.D. Smulskaya***

**PLANNING A LONGITUDINAL STUDY: BIG PROBLEMS OF A SMALL CORPUS**

**Abstract.** The paper concerns possibilities of longitudinal linguistic studies and longitudinal learner corpora in particular. It describes planning and developing a longitudinal corpus project aimed at Learner English study and analyses its diagnostic value problems.

**Keywords.** Longitude, longitudinal learner corpus, diagnostic value.

Исследование интеръязыка в учебном корпусе SPbEFL привело к обна- ружению частых случаев контаминации в области сказуемого, а именно: смешения двух базовых структур — высокочастотной для корпуса струк- туры с именным сказуемым S <BE> C и структур с простым и составным глагольным сказуемым SV и SVO [Камшилова 2012, 45]:

*I’m totally agree with them (SPb Essay)*

*But, in any case, it’s depend on person (SPb Essay)*

*…it’s depend on how much people have … (CA Essay) It’s can make human get rich (CA Essay)*

Тот факт, что эта ошибка характерна для носителей разных родных языков, позволяет предположить, что возможные причины ее появле- ния следует связывать как с индивидуальным когнитивным типом обу- чающегося, так и с интерференцией некоторых методических практик, закрепляющих «союз» личного местоимения и связки на начальных эта- пах овладения английским языком. Подобное предположение возможно

© КАМШИЛОВА Ольга Николаевна. — Kamshilova Olga. Российский государственный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [onkamshilova@gmail.com](mailto:onkamshilova@gmail.com)

© СМУЛЬСКАЯ Екатерина Дмитриевна. — Smulskaya Ekaterina. Российский государствен- ный педагогический университет им. А.И. Герцена, Россия. — Herzen University, Russia. E-mail: [Kgulaya@yandex.ru](mailto:Kgulaya@yandex.ru)

проверить долговременным наблюдением, позволяющим фиксировать с помощью регулярных срезов процесс овладения английским языком отдельным субъектом или группой обучаемых. Таким путем весьма вероятно установить, когда эта ошибка появляется в речи обучаемых и насколько стабильна она на протяжении обучения. Другими словами, речь идет о так называемом лонгитюдном исследовании, которое в прак- тике строительства учебных корпусов относится к одному из новых приоритетных направлений.

Известно, что лонгитюд как научный метод сложился в рамках пси- хологических и социологических наук. Он опирается на собственную систему понятий и набор аналитических методов и инструментов, обес- печивающих планирование, проведение исследования и последующую обработку данных.

Существенным для лонгитюдного исследования является наличие гипотез развития («каузальных гипотез»), которые проверяются в ходе наблюдения за экспериментальной группой (выборочной совокупностью) и получают оценку. Классический лонгитюд опирается на данные, кото- рые собираются для одних и тех же выборок испытуемых. В некоторых случаях лонгитюдным (кросс-секционный лонгитюд) признают исследо- вание отдельных срезов, собранных на разных выборках, но признанных взаимозаменяемыми [Корнилов 2011: 105].

Использование лонгитюда в лингвистике связано с психолингвистиче- скими исследованиями детской речи. Здесь применяются оба вида лонги- тюдных исследований: как кросс-секционный анализ, аналогом которого можно считать известные исследования и публикации К.И. Чуковского, так и классический лонгитюд, реализованный на минимальной выборке — таковыми можно считать исследования дневниковых записей речи детей (см., например, исследования В. и К. Штерн, А.В. Гвоздева, труды лабо- ратории детской речи под руководством С.Н. Цейтлин и др. [Елисеева 2015; Поспелова 2015]).

Интерес к лонгитюдным исследованиям становится приоритетным и в корпусной лингвистике. Как прообраз лонгитюдных корпусов текстов в 1980-х гг. возникает международная система обмена данными по детской речи CHILDES (Children Language Database Exchange System). [Зырянова 2012; Казаковская 2012]. Крупнейший европейский центр корпусной лингвистики в Лувене (Бельгия), который исследует процессы освоения языков, в том числе иностранных, еще в 2008 г. инициировал проект *Lon- gitudinal Database of Learner English* (LONGDALE)1. Этот проект призван преодолеть ограничения кросс-секционного анализа уже существующих учебных корпусов текстов (Learner corpora).

1 [https://www.uclouvain.be/en-cecl-longdale.html](http://www.uclouvain.be/en-cecl-longdale.html)

Лувенский лонгитюд, респонденты которого — студенты европей- ских вузов, с начальным уровнем владения английским языком не ниже В1, планируется на основе проспективного панельного дизайна, который предполагает, что сбор данных «проводится “здесь и сейчас” для каждого временного отрезка» [Корнилов 2011:105]. Корпус должен включать два типа данных: собственно тексты, устные и письменные, а также экспери- ментальные панельные данные, включающие метаданные о респондентах, их прогрессе в изучении языка, данные о среде обучения и ее влиянии на академический успех, в том числе результаты экспериментальных тестов на срезах [Altendorf, Gerckens 2012].

Создание собственного лонгитюдного корпуса для проверки гипо- тезы о причинах и времени появления контаминации в области сказуе- мого у русских школьников, изучающих английский язык, выявило ряд серьезных проблем, которые могут поставить под сомнение валидность самого исследования, так как не всегда удается учесть влияние на ход исследования таких факторов, как возраст респондентов и их иноязычная компетентность.

Планирование лонгитюдного исследования предполагает учет типичных и характерных для лонгитюда угроз валидности. Рассмотрим те моменты, которые реально вмешиваются в проведение нашего иссле- дования, и которые предстоит преодолеть по его ходу.

* 1. ***Определение выборочной совокупности и потеря респондентов*** Проблема потери респондентов относится к числу характерных для лонгитюда. Определение численного состава выборки в лонгитюдном исследовании не имеет строгих границ. Самые масштабные лонгитюды охватывали группы в несколько тысяч человек (геронтологический лон- гитюд Г. Томэ в Германии [Thomae 1993]). В лингвистике количество участников может колебаться от одного человека, что характерно для исследований дневниковых записей детей, см., например, [Казаковская

2012], до ста [Altendorf, Gerckens 2012].

При выборе респондентов учитывались доступность площадки для проведения эксперимента в течение всего исследования и владение ино- странным языком на уровне, позволяющем порождать небольшие тексты описательного характера. В итоге в качестве площадки была выбрана одна из петербургских школ с углубленным изучением английского языка, пре- подавание которого осуществляется по одной дидактической программе, по крайней мере в 5–7 классах. Требованию достаточного для наших целей владения языком соответствуют ученики 6 класса. Для того, чтобы иметь возможность кросс-секционного измерения, к участию в исследовании были привлечены ученики 7 класса. Таким образом были определены две выборочные совокупности по 23 человека в каждой. Пять проведенных

срезов свидетельствуют о незначительном колебании в числе респонден- тов, однако есть основания опасаться, что состав выборки будет меняться (пропуски по здоровью, смена школы и пр.). Таким образом, есть опасение потерь в отношении чистоты выборочной совокупности, но сохраняется надежда, что собранный материал позволит выделить индивидуальных респондентов для проведения целевых исследований (case studies).

* 1. ***Планирование временных срезов и промежутков между ними***

Количество срезов и длина промежутков между ними при плани- ровании лонгитюда определяется предполагаемой формой развития, на выявление которой направлен лонгитюд. В лингвистических лонги- тюдах, в отличие от психологических или социологических, которые могут длиться 20 лет и более, период исследования определяется ранними годами речевого развития («от двух до пяти») и доступными дневнико- выми записями (например, записи на протяжении 16 месяцев, с 1.5 до 2.8 [Казаковская 2012:130]). Или, как в случае LONGDALE, продолжитель- ность лонгитюда (не менее 3-х лет) диктуется академическими рамками. Планируемое исследование предполагается вести в течение 2 учебных лет, с ежемесячными продольными срезами. Таким образом, количество запланированных срезов равно 18. Адекватность выбранного периода определяется школьной программой, предусматривающей раз в месяц контрольный урок по развитию навыков письменной речи. Именно в рамках этого урока эксперимент меньше всего вмешивается в учебный процесс. Однако о равномерности выбранных промежутков и выполнении всех запланированных срезов говорить не приходится в силу причин объ- ективных (каникулы, конец четверти, плановые школьные мероприятия) и субъективных (волевые решения некоторых учителей). Таким образом, оценка выявляемых изменений может привести к некорректным выводам. Однако это обстоятельство может оказаться не столь релевантным в нашем случае, поскольку проект корпуса не предполагает предъявления респон- дентам контрольных тестов или анкет, которые должны влиять на модели развития, если экспериментатор управляет ими (например, при контроле обучения, проводимого во время лонгитюда). С другой стороны, нельзя не учитывать при последующей оценке контролируемого материала время его сбора, так как внешние факторы могут некоторым образом повлиять

на конечный продукт.

* 1. ***Планирование объема корпуса и структурная валидность***

Наибольшую проблему и возможную угрозу валидности начатому лон- гитюдному исследованию представляет процесс накопления объема кор- пуса. Изначально объем корпуса предполагался небольшим, что в нашем случае объясняется направленностью корпуса, характером респондентов и техническими возможностями.

Планируя ежемесячно получать от каждого респондента текст задан- ного жанра (описание картинки) и размера (не менее 10 предложений), мы столкнулись с тем, что уже со второго этапа наши респонденты частично перестали выполнять поставленные перед ними задачи. Так, если после первого среза мы получили в выборке шестиклассников 11 текстов длинной в 10 предложений и ни одного текста короче 7 предложений, то на втором срезе 10 предложений было зафиксировано в 4 текстах, на третьем — только в 3, зато значительно возросло число текстов, в которых было от 5 до 2 пред- ложений. Та же картина наблюдается и в ответах респондентов из выборки семиклассников. Кроме того, качество многих текстов свидетельствует о формальном и небрежном выполнении задания, что ставит под сомнение релевантность этих текстов для дальнейшего исследования. В результате нарушается идентичность условий порождения и жанра проверяемого и оцениваемого речевого материала, что мы рассматриваем, возможно, не вполне корректно, как аналог нарушения идентичности проверяемых в равные промежутки времени конструктов в психологических лонгитюдах, т. е. как нарушение критерия структурной валидности.

Однако, если следовать опыту LONGDALE, в лонгитюдный корпус

могут и должны входить тексты разных жанров и величины, порожденные в разных условиях. Исследование на основе всей совокупности текстов можно считать релевантным при условии соблюдения требований к выбо- рочной совокупности. Поэтому мы планируем пополнять объем корпуса другими жанрами и формами текстов, в том числе устными.

Проведенные пять срезов лонгитюдного исследования речи петербург- ских школьников, изучающих английский язык, позволяют оценить пер- вые результаты и скорректировать дальнейшую работу по сбору данных. Следует быть готовым к объективным потерям в выборочной сово- купности и предусмотреть возможность отбора данных для наиболее постоянной группы и индивидуальных случаев для проведения целевых

исследований (case studies).

Необходимо допустить и планировать изменения промежутков между продольными срезами, поскольку в нашем случае не предполагается управление моделями развития со стороны экспериментаторов (создате- лей корпуса), поэтому колебания сроков срезов не могут в значительной степени влиять на контрольный материал. Однако нельзя не учитывать воздействия фактора времени сбора материала (период после каникул, конец четверти и пр.) на его качество.

Наконец, предварительные оценки текстового материала требуют кор- ректировки решений о жанрах и формах текста, а также поисков стимулов для мотивации респондентов с учетом их возрастных характеристик. Без этого трудно достичь объема корпуса, обеспечивающего доказательность выводов.

**Литература**

1. *Елисеева М.Б.* (2015), Союзные средства в речи ребенка от двух до трех лет: лонгитюдное исследование // Вестник Северного (Арктического) федерального университета. Серия: Гуманитарные и социальные науки, № 2, c. 74–82.
2. *Зырянова Е.В.* (2012), Система CHILDES как метод сбора материалов и изучения детской речи // Известия Российского государственного педагогиче- ского университета им. А.И. Герцена, № 35 (76), ч. 2: (Педагогика, психология, теория и методика обучения), c. 113–118.
3. *Казаковская В.В.* (2012), Речь взрослого и усвоение ребенком прилагатель- ных: анализ русского лонгитюдного корпуса данных // Вопросы психолингви- стики, № 2, c. 128–135.
4. *Камшилова О.Н.* (2012), Учебный корпус текстов: потенциал, состав, структура. СПб.
5. *Корнилов С.А.* (2011), Лонгитюдные исследования: теория и методы // Экс- периментальная психология, Т. 4, № 4, c. 101–116.
6. *Поспелова С.Н.* (2015), Направления и методы изучения детской речи в оте- чественной и зарубежной лингвистике // Молодой ученый. № 10, c. 1432–1434.
7. *Altendorf, U., Gerckens, C.* (2012), Presenting LONGDALE-GE — a multi- methodological longitudinal corpus at the crossroads of English linguistics // Learner Language, Learner Corpora. Oulun yliopisto. 5–6.10.2012, p. 15.
8. *Thomae, H.* (1993), The Bonn Gerontologic Longitudional Study (BOLSA) // Zeitschrift fur Gerontologie. Vol. 26 (3), pp. 142–150.

**Literature**

1. *Yeliseyeva M.B.* (2015), Soyuznye sredstva v rechi rebenka ot dvukh do trekh let: longityudnoe issledovanie // Vestnik Severnogo (Arkticheskogo) federal’nogo universiteta. Seriya: Gumanitarnye i sotsial’nye nauki, № 2, s. 74–82.
2. *Zyryanova E.V.* (2012), Sistema CHILDES kak metod sbora materialov i izucheniya detskoy rechi // Izvestiya Rossiyskogo gosudarstvennogo pedagog- icheskogo universiteta im. A.I. Gertsena, № 35 (76), ch. 2: (Pedagogika, psikhologiya, teoriya i metodika obucheniya), s. 113–118.
3. *Kazakovskaya V.V.* (2012), Rech’ vzroslogo i usvoenie rebenkom prilagatel’nykh: analiz russkogo longityudnogo korpusa dannykh // Voprosy psikholingvistiki, № 2, s. 128–135.
4. *Kamshilova O.N*. (2012), Uchebnyi korpus tekstov: potentsial, sostav, struktura. SPb.
5. *Kornilov S.A.* (2011), Longityudnye issledovaniya: teoriya i metody // Eksperimental’naya psikhologiya, T. 4, № 4, s. 101–116.
6. *Pospelova S.N*. (2015), Napravleniya i metody izucheniya detskoy rechi v otechestvennoy i zarubezhnoy lingvistike // Molodoy uchenyi. № 10, s. 1432–1434.
7. *Altendorf, U., Gerckens, C.* (2012), Presenting LONGDALE-GE — a multi- methodological longitudinal corpus at the crossroads of English linguistics // Learner Language, Learner Corpora. Oulun yliopisto. 5–6.10.2012, p. 15.
8. *Thomae, H.* (1993), The Bonn Gerontologic Longitudional Study (BOLSA) // Zeitschrift fur Gerontologie. Vol. 26 (3), pp. 142–150.

***Г.Е. Кедрова, С.Б. Потемкин***

**РАЗРАБОТКА КОРПУСА НАРОДНЫХ СКАЗОК**

**Аннотация.** Сказки, как существеннейшая часть фольклора, являются, пожа- луй, наиболее устойчивой формой культурных традиций нации. В статье описан процесс подготовки текстового корпуса на основе собрания сказок А.Н. Афана- сьева. К обработке исходных текстов привлекались студенты филологического факультета МГУ, что также способствовало их профессиональной подготовке. Результатом работы является полнотекстовый корпус с возможностью поиска по различным критериям — словоупотреблениям, словосочетаниям, леммам, совместной встречаемости слов в тексте одной сказки, построения конкордансов, частотных словарей и пр.

**Ключевые слова.** Фольклор, лемматизация, архаизмы, словарь, компьютер.

***G.E. Kedrova, S.B. Potemkin***

**DEVELOPMENT OF THE FOLK TALES CORPUS**

**Absrtact.** Fairy tales, as an essential part of the folklore, are perhaps the most stable form of the cultural traditions of the nation. This article describes the process of preparing a text corpus on the basis of A.N.Afanaseva tales collection. The students of the Philological Faculty of the Moscow State University, were involved in processing the source texts. This work also contributed to their professional training. The result is a full-text corpus searchable according to various criteria — words, phrases, lemmas, co-occurrence of words in a fairy tale, enables concordances constructing, frequency dictionaries, etc.

**Keywords.** Folklore, lemma, archaic forms, lexicon, computer.

### Введение

Записи, собранные выдающимся русским фольклористом Алексан- дром Николаевичем Афанасьевым, — это замечательное и уникальное явление в русской литературе, и даже шире — в русской культуре. «Здесь Русский дух, здесь, Русью пахнет!» — Эти пушкинские слова полностью подходят для характеристики сказок Афанасьева. Опыт и мудрость нации приходит к нам через века со сказкой. Коллекция знаменитого Александра Николаевича (1826–1871) отвечали потребностям образованного обще- ства прикоснуться к подлинному национальному духу, национальной культуре после засилья западных влияний в XVIII — начале XIX в. Для сборника в восьми книжках «Народные русские сказки» [Афанасьев 1855–1863] было отобрано более 600 текстов. С точки зрения научного

© КЕДРОВА Галина Евгеньевна. — Kedrova Galina. МГУ им. М.В. Ломоносова, Россия. —

* 1. Lomonosov Moscow State University, Russia. E-mail: [kedr@philol.msu.ru](mailto:kedr@philol.msu.ru)

© ПОТЕМКИН Сергей Борисович. — Potemkin Sergey. МГУ им. М.В. Ломоносова, Рос- сия. — M.V. Lomonosov Moscow State University, Russia. E-mail: [potemkin@philol.msu.ru](mailto:potemkin@philol.msu.ru)

качества собрание Афанасьева намного превосходят коллекцию сказок братьев Гримм. Афанасьев не позволял себе вносить какие-либо измене- ния или литературную обработку, к которым прибегали браться Гримм. Издание Афанасьева также содержит варианты, в отличие от собрания Гримм. Филологическое исследование сказок имеет достаточно длитель- ную историю, в рамках которой сосуществуют несколько теоретических подходов, основные положения которых создавались на протяжении двух предшествующих веков. В частности, отдавая дань уважения собирателям фольклора XVIII–XIX вв., ученые ХХ в. решились отвергнуть некото- рые, весьма популярные прежде, истолкования фольклора, например,

«мифологическую школу» и «метеорологическую теорию», ярчайшим представителем которых был А.Н. Афанасьев [Афанасьев 1865–1869]. В любом материале, сказке, былине, пословице он находил «отголоски мифа о грозе, тучах, о борьбе мрака со светом». Несмотря на огромный авторитет В.Я. Проппа, некоторые его положения также вызывают возра- жения. Так, представление о принадлежности Бабы-Яги к царству мертвых не находит подтверждения при анализе текстов сказок.

Мы считаем, что эти и другие наблюдения и утверждения могут быть проверены и подтверждены либо опровергнуты в ходе исследования фактического материала. Для этих целей в ЦНИТ ГО филологического факультета МГУ им. М.В. Ломоносова разрабатывается Корпус фольклор- ных текстов (народных сказок). В основу разрабатываемого корпуса легло последнее полное академическое издание сборника сказок [Афанасьев 1984]. Это издание явилось основой для электронного ресурса1, который используется для составления представленного корпуса фольклорных текстов.

### Подготовка фольклорного корпуса

Процесс подготовки электронного ресурса к оформлению его в соот- ветствии с корпусными требованиями был многоступенчатым. Обработка коллекции включает в себя несколько этапов, как чисто технических, так и требующих привлечения специалистов-филологов или студентов про- фильного вуза, имеющих определенную филологическую подготовку. В разрабатываемом корпусе на этом этапе привлекались студенты- бакалавры, которые закончили 2 курса филологического факультета МГУ им. Ломоносова. Они выполняли обработку материалов сборника сказок во время летней компьютерной практики в Лаборатории обеспеченного компьютером обучения по специально разработанным для этой цели про- токолам. Основные стадии обработки на этом этапе были следующие. Первоначально текст каждой сказки был скопирован с вышеупомянутого сайта в отдельный файл в формате UTF8, чтобы сохранить лингвистиче-

1 https://ru.wikisource.org/wiki/Народные\_русские\_сказки\_(Афанасьев)

ски релевантную специальную разметку (символы с диакритическими знаками, ударения и т. д.). Затем текст сказки был разделен на отдельные слова, при этом каждый знак препинания принимался также за отдель- ное слово. Кортежи слов, заканчивающиеся точкой, вопросительным, восклицательным знаком, а также двоеточием и точкой с запятой рас- сматриваются как предложения. Морфологический анализ каждой слово- формы, включал лемматизацию, определение POS (часть речи), разметку и определение грамматических характеристик (род, число, падеж для существительных и прилагательных, безличных форм глагола, время, число, род для глаголов в личной форме и т. п.). В частности, для предлога указывается набор падежей существительного, которые могут управляться этим предлогом. Также проводился частичный синтаксический анализ, в том числе устанавливались связи существительное — прилагательное, предлог — существительное, глагол — существительное и т. д. Кроме того, было проведено частичное разрешение анафорических ссылок вида местоимение — существительное. Эта техническая работа была проведена в автоматическом режиме [Потемкин 2015] с использованием словаря А. А. Зализняка [Зализняк 2009].

### Несловарные формы

Хорошо известно, что на этапе первичной обработки текстового массива сказок всякий исследователь сталкивается с определенными трудностями, обусловленными типом и жанром материала. В частности, многие архаизмы, диалектизмы, иностранные заимствования (в основном из белорусского и украинского), отсутствующие в словаре Зализняка, были нами выделены и помещены в отдельный список, насчитывающий около 4000 слов.

Каждое слово в этом списке было привязано к сказке, в которой оно встречается. Впоследствии студенту предлагалась для обработки часть этого списка, состоящая из 50 слов. Пример списка приведен в табл.

*Таблица.* Часть списка слов, не зафиксированных в Словаре Зализняка

|  |  |  |
| --- | --- | --- |
| **№ словоформы** | **Словоформа** | **№№ сказок** |
| 3659 | грешон | 1, |
| 3411 | ясни | 10,11, |
| … |  |  |

Студенту предлагалось составить словарную статью для каждого слова из списка, содержащую следующие пункты:

№ словоформы (напр., 3659 для слова «*грешон*»); саму словоформу;

№ сказки (1); лемма («*грешной*») — с указанием источника, а именно

№ словаря из Списка рекомендованных словарей [Кедрова, Потемкин 2016] (11, так как лемма взята из Национального корпуса русского языка (НКРЯ1); грамматические признаки в нотации НКРЯ (прил, м, ед, кратк) — с указанием источника; толкование — с указанием источника (словаря); полную словоизменительную парадигму. Толкование должно относиться только к той семантике лексемы, которая подразумевалась в тексте сказки. При выполнении задания студент находит нужное слово в тексте соот- ветствующей сказки и изучает контекст его вхождения. Например, слово грешон встречается в сказке № 1 в следующем контексте: «… у тебя семь- десят жен, ты завсегда грешон.» Из контекста сказки можно определить часть речи и грамматические признаки рассматриваемого слова. В дан- ном случае это краткое прилагательное, мужской род, ед. число, лемма: *грешной / грешный*. Затем в Списке рекомендуемых словарей и корпусов проводился поиск подтверждения или опровержения гипотезы. В дан- ном случае соответствующая словоформа есть в НКРЯ в аналогичном контексте: *«Но в этом разе я очень грешон!»* Там же приводится лемма

и грамматические признаки словоформы прил, м, ед, кратк.

Заметим, что довольно часто в НКРЯ встречаются явные ошибки, напр. для слова № 3467 *ишла*, которое встретилось в сказке в контексте:

«*Ишла лисичка-сестричка по дорожке, Несла скалочку…*» НКРЯ при- водит следующий контекст: «*Так я прожил до девятнадцати лет и был здоров столь ужасно, что со мною стали обмороки, и кровь носом ишла*.» В обоих случаях слово *ишла* — это глагол. Тем не менее это слово в НКРЯ характеризуется следующим образом: существительное, неодушевлен- ное, женский род, единственное число, т. е. часть речи определяется как существительное.

В случае подобной ошибки студент посылает сообщение об ошибке в НКРЯ и вместо этой информации самостоятельно вводит правильную. Затем составляется флективная парадигма и дается словарная дефиниция. Все слова из списка обрабатываются таким же образом. Выполнение работы способствует не только подготовке фольклорного корпуса сказок и исправлению метаданных, содержащихся в НКРЯ, но также оказывается полезным практическим упражнением при обучении студентов-филологов, что отражено в студенческих отчетах о прохождении практической части учебного плана подготовки специалистов по профилю «Филология».

Предполагается, что обработанный подобным образом компьютерный массив позволяет решать достаточно широкий спектр лингвистических, литературоведческих и культурологических задач. В качестве примера такого рода исследования можно привести разработанные нами процедуры и алгоритмы формирования индекса сюжетов народных сказок.

1 <http://www.ruscorpora.ru/> — Дата доступа: 13.09.2016.

### Заключение

Рассмотренные в статье процедуры и алгоритмы разработки Корпуса фольклорных текстов (народных сказок) служат для подготовки инстру- мента филологического анализа собрания сказок, собранных А.Н. Афа- насьевым. С помощью Корпуса можно изучать реальный язык конца XIX в. (время составления Собрания сказок). Подготовка корпуса с при- влечением студентов-филологов оказала положительное влияние на их навыки, а также помогла повысить уровень овладения ими современными информационными технологиями.

Опыт, приобретенный в ходе решения подобных практических задач, может быть использован для разработки учебных программ по специаль- ности филология, теоретическая и прикладная лингвистика, культуроло- гия, история и другим гуманитарным дисциплинам.

**Литература**

* + 1. *Афанасьев А.Н.* (1855–1863), Народные русские сказки. М. Вып. I–VIII.
    2. *Афанасьев А.Н.* (1865–1869), Поэтические воззрения славян на природу. Опыт сравнительного изучения славянских преданий и верований, в связи с мифи- ческими сказаниями других родственных народов. М. Т. 1–3.
    3. *Афанасьев А.Н.* (2015), Электронное собрание сказок: <http://feb-web.ru/feb/> skazki/default.asp?/feb/skazki/texts/af0/af0.html
    4. *Бараг Л.Г. и др.* (1979), Сравнительный указатель сюжетов: Восточносла- вянская сказка. Л. 437 с.
    5. *Зализняк А.А.* (2009), Грамматический словарь русского языка: словоизме- нение: около 110000 слов / А.А. Зализняк. 6-е изд., стер. М.: АСТ-ПРЕСС, 795 c.
    6. *Кедрова Г.Е., Потемкин С.Б.* (2007), Использование корпуса параллельных текстов для пополнения специализированного двуязычного словаря // Труды и материалы III Международного Конгресса исследователей русского языка «Рус- ский язык: исторические судьбы и современность», с. 627–628.
    7. *Кедрова Г.Е., Потемкин С.Б.* (2016), Инструкция по выполнению лет- ней компьютерной практики 2016: [http://www.philol.msu.ru/~serge/Afanasev/](http://www.philol.msu.ru/%7Eserge/Afanasev/) ИНСТРУКЦИЯ-2015.doc (дата доступа: 13.09.2016).
    8. *Потемкин С.Б.* (2015), Авторский корпус и словарь языка Антона Чехова // Труды международной конференции «Корпусная лингвистика — 2015». СПб., с. 382–389.

**Literature**

1. *Afanas’yev A.N.* (1855–1863), Narodnye russkie skazki. M. Vyp. I–VIII.
2. *Afanas’yev A.N*. (1865–1869), Poeticheskie vozzreniya slavyan na prirodu. Opyt sravnitel’nogo izucheniya slavyanskikh predaniy i verovaniy, v svyazi s miﬁcheskimi skazaniyami drugikh rodstvennykh narodov. M. T. 1–3.
3. *Afanas’yev A.N*. (2015), Elektronnoe sobranie skazok: <http://feb-web.ru/feb/> skazki/default.asp?/feb/skazki/texts/af0/af0.html
4. *Barag L.G.* i dr. (1979), Sravnitel’nyi ukazatel’ syuzhetov: Vostochnoslavyan- skaya skazka. L. 437 s.
5. *Zaliznyak A.A.* (2009), Grammaticheskiy slovar’ russkogo yazyka: slovoiz- menenie: okolo 110000 slov / A.A. Zaliznyak. 6-e izd., ster. M.: AST-PRESS, 795 s.
6. *Kedrova G.E., Potemkin S.B.* (2007), Ispol’zovanie korpusa parallel’nykh tekstov dlya popolneniya spetsializirovannogo dvuyazychnogo slovarya // Trudy i materialy III Mezhdunarodnogo Kongressa issledovateley russkogo yazyka «Russkiy yazyk: istoricheskie sud’by i sovremennost’», s. 627–628.
7. *Kedrova G.E., Potemkin S.B*. (2016), Instruktsiya po vypolneniyu letney komp’yuternoy praktiki 2016: [http://www.philol.msu.ru/~serge/Afanasev/INSTRUKT-](http://www.philol.msu.ru/%7Eserge/Afanasev/INSTRUKT-) sIYa-2015.doc (data dostupa: 13.09.2016).
8. *Potemkin S.B.* (2015), Avtorskiy korpus i slovar’ yazyka Antona Chekhova // Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2015». SPb., s. 382– 389.

***Н.Н. Коробейникова***

**КОРПУСНО-ОРИЕНТИРОВАННЫЙ ПОДХОД**

**К ИССЛЕДОВАНИЮ СИНОНИМИЧЕСКОГО РЯДА ТЕМПОРАЛЬНЫХ НАРЕЧИЙ**

**Аннотация.** Данная статья посвящена изучению лексической сочетаемости английских темпоральных наречий с семантикой мгновенности, образующих синонимический ряд*.* В результате проведенного корпусного анализа были установлены сочетаемостные схождения, а также сугубо специфические сферы лексической комбинаторности исследуемых наречий, что позволило более точно дифференцировать темпоральные единицы в синонимическом ряду.

**Ключевые слова.** Темпоральные наречия, синонимический ряд, коллокаты, лексическая сочетаемость.

***N.N. Korobeinikova***

**CORPUS-ORIENTED APPROACH TO THE STUDY OF SYNONYMOUS TEMPORAL ADVERBS**

**Abstract.** The purpose of this article is to consider lexical compatibility of the Eng- lish synonymous temporal adverbs with the meaning of instantaneousness. The corpus analysis of empirical data has allowed to reveal the similar and different aspects in the lexical compatibility of the synonyms.

**Keywords.** Temporal adverbs, synonyms, collocates, lexical compatibility.

Синонимический ряд представляется как «смысловой фокус», сосре- доточивающий внимание на определенном фрагменте действительности и заложенный «в когнитивную модель говорящего» [Хантакова 2006: 96]. Каждая единица синонимического ряда, актуализируя определенную часть содержания, тесно связана с другими синонимами и одновременно противопоставлена им как часть целому [там же].

Исследователь английской синонимики В. Г. Вилюман считает, что синонимы показывают как схождения, так и расхождения по трем следую- щим признакам: значение, дистрибуция, сочетаемость [Вилюман 1980: 30]. Ю. Д. Апресян представляет более подробную классификацию возможных различий между синонимами, выделяя грамматические, стилистические, эмоционально-экспрессивные, жанровые, социальные, профессиональ- ные, территориальные и сочетаемостные [Апресян 1995: 216].

Данная статья посвящена изучению лексической сочетаемости семи английских темпоральных наречий с семантикой мгновенности, обра- зующих синонимический ряд: *immediately, instantly, instantaneously, right*

© КОРОБЕЙНИКОВА Наталья Николаевна. — Korobeinikova Natalia. Евразийский линг- вистический институт в г. Иркутске — филиал ФГБОУ ВО МГЛУ (МГЛУ ЕАЛИ), Россия. Eurasian Linguistic Institute, afﬁliate of MSLU, Russia. E-mail: [nati071](mailto:nati0711@mail.ru)[1@mail.ru](mailto:1@mail.ru)

*away, straight away, at once* и *forthwith.* Цель проводимого в рамках статьи исследования — выявить общее и различное в лексической сочетаемо- сти темпоральных наречий изучаемого синонимического ряда на основе анализа их наиболее частотных коллокатов. Эмпирической базой для исследования послужили данные двух корпусов: Британского националь- ного корпуса (British National Corpus — BNC) и Корпуса современного американского английского (Corpus of Contemporary American — COCA). Выбор BNC и COCA обусловлен тем, что наиболее полные и актуальные данные о семантической структуре и функциональном потенциале лексем могут быть выявлены при обращении к корпусам, в которых оперативно представляются текущие процессы и новые тенденции в различных ком- муникативных сферах системы «живого» языка [Баринова 2015: 100].

Для того чтобы определить сочетаемостный потенциал темпораль- ных наречий, мы обратились к статистическому инструментарию, пред- ставленному в BNC и COCA, а именно к опции определения коллокатов слова. Коллокаты (от *англ*. collocates) — это слова, встречающиеся в непо- средственной близости к ключевому слову, но не обязательно связанные с ним грамматически [Lindquist 2010: 73]. В корпусах BNC и COCA были рассмотрены коллокаты, находящиеся на расстоянии до четырех слов влево и вправо от каждого из наречий, поскольку, согласно Д. Байберу, словам не обязательно непосредственно примыкать друг к другу, чтобы их можно было ассоциировать друг с другом [Biber 1998]. Необходимо также отметить, что не частотные в отдельности, коллокаты могут быть объединены в тематические группы, относящиеся к той или иной семан- тической области.

Руководствуясь такими корпусными данными, как статистический

критерий абсолютной частотности (т. е. общей частоты встречаемости словоупотребления в корпусе) и мера MI (Mutual Information — «коэф- фициент взаимной информации или зависимости, который сравнивает зависимые контекстно-связанные частоты с независимыми, как если бы слова появлялись в тексте совершенно случайно» [Хохлова 2010: 12]), мы установили, классифицировали и сопоставили коллокаты темпоральных наречий.

В результате проведенного анализа было обнаружено, что б*о*льшая часть коллокатов исследуемых наречий принадлежит таким тематическим группам, как «ментальная сфера», «передача информации» и «фазовость действия».

Согласно корпусным данным группа коллокатов «ментальная сфера» является доминирующей для наречий *right away* и *straight away*, зани- мает вторую позицию среди тематических групп темпоральных наречий *at once* и *instantly* и третью — среди тематических групп *immediately*. В рамках данной тематической группы представленные выше наре-

чия сочетаются с такими лексемами, как *know*, *notice*, *recognize*, *guess*, *spot*, *realize*, например: *I* ***knew right away*** *he was after Julie, and I could tell by the way she looked at him that she was interested* [COCA]*; He had* ***noticed straight away*** *that Angel One was unarmed, having obviously discarded the gun he had ﬁred in the corridor earlier* [BNC]*; Elizabethan audiences would presumably have* ***spotted at once*** *that he was ‘putting on the style’ with his phrases cobbled together out of distant memories of Kyd, Marlowe, and Greene* [BNC]*; Coming to the gate, he looked in at the cot- tage and* ***instantly recognized*** *the fourth photograph in Heather’s collection* [BNC]*; From the dress of both the man and the woman she* ***realized imme- diately*** *that they were not only strangers, but class* [BNC]*.*

Однако стоит обратить внимание, что среди коллокатов *instantaneously* и *forthwith* в корпусах не были зафиксированы лексемы, относящиеся к тематической группе «ментальная сфера».

Достаточно большое количество коллокатов, входящих в тематиче- скую группу «передача информации», имеют пять темпоральных наречий из семи, исключением являются *instantly* и *instantaneously*. Совокупность лексических единиц этой группы занимает второе место по частотности среди тематических групп наречия *right away*, включая в себя следующие коллокаты: *answer*, *reply*, *notify*: *For the ﬁrst time in many years Europa didn’t* ***answer right away****; Hadrian didn’t* ***reply right away****, staring into the single candle and massaging his callused palms; The Caretaker was* ***noti- ﬁed right away*** [COCA].

Данная тематическая группа находится на третьем месте по частоте сочетаний с наречиями *straight away, forthwith* и *at once* и на пятом — среди тематических групп *immediately*. Наряду с глаголами, употребляемыми совместно с *right away*, в сочетаниях с четырьмя остальными наречиями были также обнаружены такие лексические единицы, как *inform* (*«…» they should* ***immediately inform*** *their ﬁrm of their employment plans and take no part in the audit of their future employer «…»*), *report* (*He wrote her a brief note merely saying he had had to go home owing to illness and had found, on his return, that he was to* ***report forthwith*** *to the War Ofﬁce*) и *respond* (*‘It is when you expect me to accept a pronouncement about men offered as fact and founded on ignorance,’ he* ***responded at once***)*.*

Лексемы, входящие в тематическую группу «фазовость действия», также имеют одни из наиболее высоких показателей частотности упо- требления среди коллокатов наречий *right away* и *straight away*. В данной тематической группе *right away* и *straight away* главным образом совместно встречаются с фазовыми глаголами начала действия *start* и *begin*, напри- мер: *He would arrive at her door and they would* ***begin right away****, sometimes before he took his coat off; Learning from our meandering in the previous sec- tion we shall* ***start straight away*** *with Gauss’s law* [BNC].

Среди коллокатов темпоральных наречий *instantly, instantaneously* и *at once* не были обнаружены лексические единицы с семантикой фазовости. В ходе исследования также обратили на себя внимание достаточно высокие сочетаемостные показатели употребления наречия неполноты действия *almost* совместно с анализируемыми темпоральными лексемами. *Almost* маркирует существование некоторой временной паузы между моментом принятия решения о совершении какого-либо действия и его непосредственным осуществлением. Из исследуемого нами синоними- ческого ряда с *almost* наиболее часто встречаются четыре темпоральных наречия, а именно *immediately*, *instantly*, *instantaneously* и *at once*. Следует отметить, что *almost* является одним из наиболее частотных коллокатов этих наречий. В частности, согласно данным BNC данная лексема зани- мает первую позицию среди коллокатов наречий *instantaneously* (*If one imagines a machine that can travel from junction to junction* ***almost instanta- neously****, ﬁnding a path to the exit would take only seconds*) и *at once* (*I reg- istered just after my eighteenth birthday and had my medical* ***almost at once***), вторую — в группе частотных сочетаний с *immediately* (*It seems that it was taken from the car park next to her ofﬁce but abandoned* ***almost immediately***) и третью — с *instantly* (*With the question circling around in her head, and certain she would lie awake in confused alarm, she fell asleep* ***almost instantly***).

При этом для темпоральных наречий *right away*, *straight away* и *forthwith*

совместное употребление с *almost* нехарактерно.

Рассмотрев тематические группы в той или иной мере свойственные большинству темпоральных наречий, перейдем к освещению лексических сочетаний, характерных для отдельных единиц исследуемого синоними- ческого ряда.

Для *instantly* и *instantaneously* характерно употребление совместно с коллокатами, обозначающими прекращение жизни (*die*, *kill*). Отсутствие этих лексических единиц среди коллокатов остальных наречий, возможно, обусловлено тем, что только у *instantly* и *instantaneously* доминирующим значением является «в течение короткого промежутка времени», которое выражает высокую скорость выполнения действия за предельно мини- мальный временн*о*й промежуток. В данном случае акцент делается на том, что смерть наступает мгновенно, например: *He* ***died instantly*** *from head injuries caused by either the fall, or the impact of a train coming the other way; He would certainly have been* ***killed instantly*** *had he still been at it; The pathologist at Princess Margaret Hospital who carried out the post mortems said all three men would have* ***died instantaneously****, describing the crash as ‘non-surviveable’; If this emission takes place it triggers the breaking of a vial of poison gas which* ***instantaneously kills*** *the cat* [BNC]*.*

*Immediately* и *at once* достаточно часто функционируют в сочетании

с лексемой *serve*, манифестирующей в большинстве случаев безотлага-

тельную подачу приготовленного блюда: *If the top isn’t browned, ﬁnish the dish under the grill and* ***serve immediately*** *with a green salad; Transfer to a serving platter and* ***serve at once*** *with the dipping sauces* [BNC]*.*

Наряду с этим отличительной особенностью *immediately* является совместное сочетание с лексемами, выражающими определенную после- довательность во времени или пространстве (*after*, *before*, *follow*, *preced- ing*, *precede*, *afterwards*, *prior*), что не свойственно другим наречиям из рассматриваемого нами ряда, например: *In the period* ***immediately after*** *the Second World War sweeping changes affected the staff in Scotland;‘The potted shrimps were put on the table by the parlourmaid* ***immediately before*** *the guests entered the dining-room,’ said Milsom* [BNC]*.*

Для темпорального наречия *right away*, в отличие от других рассма- триваемых наречий, свойственно сочетание с лексемой *like*, выражающей симпатию, положительную оценку, например: *She couldn’t explain why, but she* ***liked*** *this reporter* ***right away****; When offered a show by Dick Clark, he* ***liked*** *the idea* ***right away*** [COCA].

Темпоральное наречие *forthwith* главным образом присутствует в кол- локациях, не частотных для остальных наречий, что, по нашему мнению, может быть вызвано узкой спецификой его функционирования в англий- ском языке — судебно-правовой сферой. Самой многочисленной по коли- честву коллокатов является тематическая группа «приказ/требование», выделенная на основании присутствия в контекстах с наречием *forthwith* таких лексических единиц, как *to order, order, to require, request, to demand, demand, petition*, например: “<*…> the court shall not make an* ***order forthwith***

*<…>; Article 12 of the Convention* ***requires****, prima facie, the mandatory return*

***forthwith*** *of a child who has been wrongfully removed <…>*” [BNC]*.*

Таким образом, в результате проведенного анализа темпоральных наречий с точки зрения их лексической сочетаемости были установлены, систематизированы и сопоставлены тематические группы их коллокатов. Отмеченные выше сочетаемостные схождения в таких тематических группах, как «ментальная сфера», «передача информации», «фазовость действия» и «неполнота действия», а также наличие сугубо специфичных сфер лексической комбинаторности являются, на наш взгляд, одними из эффективных маркеров, позволяющих наряду с другими идентификато- рами более точно дифференцировать исследуемые темпоральные единицы в синонимическом ряду.

**Литература**

* 1. *Апресян Ю.Д.* (1995), Избранные труды, том I. Лексическая семантика: 2-е изд., испр. и доп. М.: Школа «Языки русской культуры», Издательская фирма

«Восточная литература» РАН.

* 1. *Баринова А.О., Гриценко Е.С.* (2015), Этничность в англоязычной картине мира (на материале лексикографии и корпусов): монография. М.: ФЛИНТА: Наука.
  2. *Вилюман В.Г.* (1980), Английская синонимика. М.: Высшая школа.
  3. *Хантакова В.М.* (2006), Теория синонимии: опыт интегрального анализа: монография. Иркутск: Иркутский государственный лингвистический университет.
  4. *Хохлова М.В.* (2010), Исследование лексико-синтаксической сочетаемости в русском языке с помощью статистических методов (на базе корпусов текстов): автореф. дис. канд. филол. наук. Санкт-Петербург.
  5. *Biber D.* (1998), Corpus Linguistics. Investigating language structure and use. Cambridge: Cambridge University Press.
  6. British National Corpus [Electronic resource]: <http://corpus.byu.edu/bnc/>(дата обращения 28.08.2016).
  7. Corpus of Contemporary American English [Electronic resource]: [http://corpus.](http://corpus/) byu.edu/coca/ (дата обращения 27.08.2016).
  8. *Lindquist H.* (2010), Corpus Linguistics and the Description of English. Edinburgh: Edinburgh University Press.

**Literature**

1. *Apresyan Yu.D.* (1995), Izbrannye trudy, tom I. Leksicheskaya semantika: 2-e izd., ispr. i dop. M.: Shkola «Yazyki russkoy kul’tury», Izdatel’skaya ﬁrma «Vostoch- naya literatura» RAN.
2. *Barinova A.O., Gritsenko E.S*. (2015), Etnichnost’ v angloyazychnoy kartine mira (na materiale leksikograﬁi i korpusov): monograﬁya. M.: FLINTA: Nauka.
3. *Vilyuman V.G*. (1980), Angliyskaya sinonimika. M.: Vysshaya shkola.
4. *Khantakova V.M*. (2006), Teoriya sinonimii: opyt integral’nogo analiza: monograﬁya. Irkutsk: Irkutskiy gosudarstvennyi lingvisticheskiy universitet.
5. *Khokhlova M.V.* (2010), Issledovanie leksiko-sintaksicheskoy sochetaemosti v russkom yazyke s pomoshch’yu statisticheskikh metodov (na baze korpusov tekstov): avtoref. dis. kand. ﬁlol. nauk. Sankt-Peterburg.
6. *Biber D.* (1998), Corpus Linguistics. Investigating language structure and use. Cambridge: Cambridge University Press.
7. British National Corpus [Electronic resource]: <http://corpus.byu.edu/bnc/>(data obrashcheniya 28.08.2016).
8. Corpus of Contemporary American English [Electronic resource]: [http://corpus.](http://corpus/) byu.edu/coca/ (data obrashcheniya 27.08.2016).
9. *Lindquist H.* (2010), Corpus Linguistics and the Description of English. Edin- burgh: Edinburgh University Press.

***M. Milojković***

**CORPUS-DERIVED SUBTEXT**

**IN RUSSIAN-ENGLISH TRANSLATION**

**Abstract.** In conventional translation contexts of situation are always derived from replacing of the *vocabulary* of the source language with that of the target language. Recently it has been discovered that not only the semantic prosody of lexical items, but also the most frequent lexical collocates of *grammar strings* (termed corpus-derived subtext, see Louw and Milojkovic [2016] have a bearing on a text’s meaning. As an illustration, the paper will compare the subtext of Falen’s [Pushkin 2011] translation of ‘Передо мной явилась ты’ to the subtext of the original, and discuss practical issues to do with the evaluation of existing translations and producing new ones.

**Keywords.** Corpus stylistics, Contextual Prosodic Theory, semantic prosody, corpus-derived subtext, Louw, Pushkin/

### Introduction

A ﬁne example of very gifted Russian-English translation is that of Alexan- der Pushkin by James E. Falen. In his *Note on the Translations* to Pushkin’s *Selected Lyric Poetry* Falen writes: ’Pushkin’s work is notoriously elusive for the translator and on many occasions as I worked on his poems, I had to abandon my efforts, ﬁnding many of the lyrics utterly untransferable into English, at least by me.’ It is all the more interesting to investigate the poems whose translations by Falen met his own very high standards and found their way into print. This paper will take up investigating his translation of the stanza by Pushkin whose subtext has already been researched in great detail [Louw and Milojkovic 2016: 213-239]: *Я помню чудное мгновенье:*

*//* ***Передо мной явилась ты, //*** *Как мимолетное виденье, // Kак гений чистой красоты.*

A short recapitulation of Louw’s Contextual Prosodic Theory is in order. It offers a theoretical framework for describing the linguistic mechanisms which bring about *foregrounding*, or linguistic deviation which can be explored for literary effect. Foregrounding can be explained by deviations from semantic prosodies of lexical words, as well as by deviations from the semantic auras of grammar strings. Regarding the latter, it has been found that the role of grammar strings may be shown through their most frequent lexical collocates — quasi- propositional variables [QPVs], which constitute the string’s *subtext*. Grammar strings possess both semantic auras and lexical subtext, both of which interact with context clues and add subliminal layers of meaning to that which is accessible to our intuition.

© MILOJKOVIĆ Marija. Faculty of Philology, University of Belgrade, Serbia. E-mail: marija. [milojkovic@ﬁl.bg.ac.rs](mailto:milojkovic@ﬁl.bg.ac.rs)

Given that the subtext of Pushkin’s original line, quoted in bold above, took 26 pages in the volume, it would be overambitios to recount it here and convince the reader. Instead, the paper will give the briefest possible overview of findings and go on to give a detailed description of the subtext in the translated line offered by Falen.

### Subtext in Pushkin — a brief overview of existing data

This section will focus on the description of the subtext in Pushkin’s origi- nal text (the reader is referred to Louw and Milojkovic [2016: 213-139] for a detailed analysis). Brieﬂy, “перед” (“before”) and “явиться” (“appear”), co- selected in this order with up to ﬁve words intervening, yielded 127 contexts in the RNC (Russian National Corpus). Taking into account Firth’s [1957: 182] taxonomy of the context of situation, in 6 contexts a woman appears before a man. These are the states of affairs: (a) the man is attracted by the woman in all 6 contexts; (b) an unexpected transition is present in all 6 contexts; nega- tive collocates are present in 5 contexts out of 6. The one exception is Natasha before Bolkonsky in *War and Peace*: she comes to ask the dying Bolkonski to forgive her past betrayal. She has the author’s every sympathy, but the very need for forgiveness points at her wrongdoing. All in all, these contexts of situation are surrounded by a negative aura.

The lexical subtext of Pushkin’s chosen grammar string was also inves- tigated. The line **ты+\*сь+перед\*** (“you \*ed before”; the word order needed to be changed as the original one is so poetic that there were no matches; this maneuver highlights the usefulness of the co-selection exercise above, where the preposition and the verb are in the right order) points at negativity in the RNC, except for line 3, a paraphrase of Pushkin:

А чем **ты умывалась перед** нанесением крема?

«Давай поиграем“, — предлагаешь **ты, прохаживаясь перед** ним в соблазнительном... Помню один чудесный момент, когда **ты появилась передо** мною, как будто...

Помнишь, **ты извинилась перед** гостем, вышла за...

— Чем же **ты провинилась перед** волшебницей, что она... так испугала моего садовника, зачем **ты появилась перед** ним?

It is clear that the subtext (as deﬁned above, the lexical items appear- ing within this grammar string in the reference corpus) is that of guilt and provocation. When a still shorter searchline, **\*ась перед** (“\*ed before”), was investigated in the RNC, 1404 contexts appeared, the ﬁrst 200 of which were studied. The table below presents a pertinent overview of the ﬁndings, citing characteristic examples of QPVs:

*Table.* The lexical collocates (QPVs) of **\*ась перед** (“\*ed before”)

|  |  |  |  |
| --- | --- | --- | --- |
| **QPVs in the meaning of ‘stopped’** | **QPVs sharing the mean- ing of ‘to be’ or ‘appeared’** | **QPVs describing loss of face or control** | **QPVs describing attractiveness** |
| oстановилась (26 verbs) останавливалась (3 verbs)  3/29 — positive  (if the woman stops before a building/ landscape) | 21 verbs (появилась оказалась находилась ) 3/21 — positive (if the woman is  an actress/performer) | 29 verbs (извинилась склонилась, разделась) | 14 verbs (вертелась разделась) |

The full implications of the QPVs, described in the headings and com- ments in the table, are more convincingly presented in the 2016 volume. Here it only remains to summarise that the subtext of the grammar string indicates that the woman appearing before a man is not positively viewed, unless she is an actress; moreover, the ﬁndings extracted from wider contexts show that the subtext of **\*ась перед** (“\*ed before”) also indicates a possible dangerous transition.

### Subtext in Falen’s translation — how close to the line of best ﬁt?

Let us now take up the research question of this paper, namely, whether Falen’s translation of Pushkin’s lines contains subtext which could be said to correspond to that of the original: *I still recall a wondrous vision: //****That day when I beheld your face****,// A ﬂeeting moment’s apparition// Of perfect beauty and of grace*. The translator has kept Pushkin’s rhythm and rhyming pattern, which, of course, restricted his lexical choices. The grammatical string he opted for in the second line does not contain either the verb ‘appear’ or the preposition ‘before’; instead, his chosen string is **that \* when I \* your \***. Let us investigate the string’s subtext.

The search in the BNC yielded no matches, while the COCA gave 6 concordance lines, all but one occurring only once. The string of words that occurred more than once — three times — was **that’s when I saw your**. Strictly speaking, this is a different grammatical string from the one used by Falen, as it has a grammar word in place of the ﬁrst wildcarded slot, and I would have probably ignored it had I been given more material to work with. What made me stop at this string of words was that all the three contexts seemed morbid: two involved death and one described the bullying of a younger child by an older one, which, judging by the context, could have ended in suffocation (the reader can check the contexts in COCA [Davies 2008]). The string **that’s when I saw your** was also checked in Google Books — Fiction (accessed in 2013, nowadays no longer available online). In the 10 available contexts, death is a present motif in 4, arson in one, theft in one, adultery in one. The string is a transition to rescue or safety in 6 contexts; it is a transition to discovering a

morbid issue in ﬁve contexts; to adultery in one, to romance (presumably) in one. Then I searched for **that is when I saw your** and discovered 7 contexts: one of mortal danger; one in which a Catholic priest recounts to an 18-year- old his meeting with the boy’s mother when she abandoned the child; “that is when I saw your message” — says the man who is not in love to the woman who is; “that is when I saw your likeness in him” — romance (“him” is the son of both); unintended rape under a magic spell; rape and incest; a murdered preacher mentioned in the wider context. In short, death is present in 2 contexts, romance in 2, birth in 2, rape in 2, incest in one. In 4 contexts the string is a transition to rescue, a saving moment.

The shorter string, **that is when I saw** will be of interest mainly because of the consistent state of affairs. Out of 20 contexts, violence, imminent danger or death are present in 11 contexts (4, 5, 6, 11, 13, 14, 16, 17, 18, 19, 20); faith

and religion in 5 (1, 2, 3, 7, 10); romance (wedding) in one (incidentally, the book is called *I Married a Preacher with Deep Dark Secrets*).

The string **that was when I saw your** yielded 2 contexts in the Google Books — Fiction corpus: one describing the grief of a man for the disappeared wife and the dead child, and one of a murder.

Next, since the Google Books corpus (Davies 2011) does not process strings of more than 5 words, I looked at the contexts of those of the ﬁrst 100 strings yielded by the searchline **that \* when I \*** that could be — and were — followed by ‘your’ in the corpus. The strings arrived at by the search of **that \* when I \* your** yield two lists of quasi-propositional variables. The ﬁrst list usually denotes a moment in the past, either through “is” or “was” or through lexis like “night”, “moment” or “happened”. In the place of the second wildcard there was found the verb “to be” as “is” or “was”, verbs of cognition or perception (“saw”, “heard”, “knew”, “realized”, “met”, “found”). The states of affairs created here are those of signiﬁcant transitions in the past, many involving morbid or romantic contexts. Romance is more present in the strings that do not contain “is” or “was” in the ﬁrst wildcarded slot and the lexical word “saw” in the second slot; death, morbidity or danger are present less. Religious contexts found in the contexts of “that is when I saw” were not found here.

On the whole it would seem that the subtext of Falen’s line “That day

when I beheld your face” corresponds to Pushkin’s subtext in “Передо мной явилась ты” in its evaluative ambiguity. In Pushkin the woman is liable to censure which is prevented by lexical collocates. Falen keeps all the lexical collocates implying transience and the supernatural nature of the vision, while his subtext in the language has been used either as a romantic transition or a dangerous one, and always signiﬁcant. This is often a moment of rescue, but sometimes of realization of irreparable harm. In the contexts of romance and man-woman contexts generally there is no implication of the woman’s guilt, or unlawful sexual provocation — Falen’s verb does not have a grammatical

feminine ending, but would it have changed anything? Crucially, if the string contains “saw” — a superordinate of “beheld” — the prospect of danger rises sharply, but the potential of saving the situation or salvation through faith also appears. However, “saw” does not necessarily chunk the same states of affairs as “beheld”. What is the contribution of “beheld” to the quality of Falen’s translation? What follows is the ﬁrst twenty random contexts yielded by ‘beheld your’ in Google Books — Fiction (repeated contexts were ignored): love and torture (*Notre Dame de Paris*); religious discourse (7 contexts); a vision of treachery described by Herodotus; context of a parent losing a child; romance (in sickness and with morbid consequences, from *The Monk*); an epistle (from Elegant Epistles, 1814); “I knew your illness and amaz’d beheld your sudden cure” (from Horace’s Satires, Book III); *Mystical Poems of Rumi* (“would you say that I beheld your face?” — erotic descriptions of faith in God); romance (*Humphrey Clinker*); a letter to a wife (“I beheld your likeness in your girls”); “this young man has beheld your foresight and your kindliness” — on tutoring the young (from *Emile: Or; on Education*, by Jean-Jacques Rousseau); context of jealousy in love (“beheld your deeds”); context of romantic love (“beheld your eyes”). In short, “beheld your” is used in elevated discourse, in older texts or their translations. Religion is the topic of 9 contexts out of 20, love of 6 contexts (with morbid elements in 2). In this, “beheld” resonates with contexts of religious worship in the Russian subtext of Pushkin’s line (not mentioned in this paper), as well as with contexts where the woman appears before the man carrying the aura of sexual attractiveness.

Generally speaking, Falen’s translation of Pushkin’s line does transfer many

of Pushkin’s subtextual implications into English, most of all through the very ambiguity of Falen’s chosen grammatical string — what is not love, could turn out to be death. “Beheld” co-selected with “your” contributes to the ambiguity by adding elements of religious discourse, also present in Pushkin’s subtext, and by strengthening the element of romance, which becomes complicated by unrequited love, torture, morbidity, the absence of the wife or jealousy — indeed, out of the 6 love contexts only in 2 do we ﬁnd no complications except the usual one of wooing. What Falen’s subtext does not involve at all is the implication that the woman may be at fault. Love in Falen’s subtext is often the tale of the ﬁrst meeting, sometimes of disappointments, complications, part- ings and even rape — but the woman is not the one whose appearing, despite her attractiveness, is thought wrong or inappropriate. If in Pushkin’s subtext the appearance of an attractive woman carries something wrong with it that we could read either as guilt or as provocation, for Falen’s subtextual contexts there is such rule. Thus, Falen’s subtext could be said to be as ambiguous as Pushkin’s, but — more adoring.

Had Falen’s line run ‘When you appeared before me’, the overall negativity

would have been preserved in **\*ed before**, with QPVs “appeared” (298), “hap-

pened” (160) and “died” (116) in the BNC; in the COCA they are “happened” (994), “died” (579), and “testiﬁed” (547). As for **appeared before**, apart from the decidedly prevalent prosody of appearing before courts and committees, it creates states of affairs similar to Russian ones that are chunked by “перед” (‘before’) and “явиться” (‘appear’). Still, the study of this collocation in the 9 available *ﬁction* contexts in the BNC shows that in English a woman is at fault *unless* she is attractive — which is at odds with Russian. The subtext of either guilt or sexual provocation persists in grammatical strings of both languages, but as for states of affairs chunked by the verb and the preposition, oddity and attraction do not mingle. Therefore, it seems to me, on the basis of these albeit insufﬁcient 9 contexts, that Falen’s subtext, in its aura of fatal outcomes, might be more convincingly ambiguous than the one created by **\*ed before**.

**Literature**

1. *Firth, J.R.* (1957), Papers in Linguistics 1934–1951. Oxford: OUP.
2. *Louw W.E. and Milojkovic M.* (2016), Corpus Stylistics as Contextual Prosodic Theory and Subtext. Amsterdam: John Benjamins.
3. *Pushkin A.* (2011), Selected Lyric Poetry. Northwestern University Press.
4. *Davies, M*. (2008), *The Corpus of Contemporary American English: 450 million words, 1990-present*: [http://corpus.byu.edu/coca/;](http://corpus.byu.edu/coca/%3B)
5. *Davies M*. (2011), *Google Books Corpus*: <http://googlebooks.byu.edu/>

***О.А. Митрофанова***

**ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ПАРАЛЛЕЛЬНЫХ И СОПОСТАВИМЫХ ТЕКСТОВ В ПОСТРОЕНИИ ТЕМАТИЧЕСКИХ МОДЕЛЕЙ КОРПУСОВ1**

**Аннотация.** В статье рассматриваются методы обработки параллельных и сопоставимых корпусов текстов в ходе тематического моделирования. Обсуж- даются процедуры оценки сопоставимости текстов в корпусе, отбора пар слов — кандидатов в переводные эквиваленты на основе двуязычных параллельных и сопоставимых текстов, формирования кластеров лексем — кандидатов в пере- водные эквиваленты, построения многоязычных тематических моделей корпусов. **Ключевые слова.** Корпусы текстов, тематическое моделирование, параллель-

ные и сопоставимые тексты.

***O.A. Mitrofanova***

**POSSIBILITIES OF PARALLEL AND COMPARABLE TEXTS IN THE DEVELOPMENT OF TOPIC MODELS FOR CORPORA**

**Abstract.** The paper presents discussion of the approaches to processing parallel and comparable text corpora in course of topic modelling. We focus our attention on the evaluation of text comparability, selection of possible translation equivalents from bilingual parallel and comparable texts, clustering of lemmata as possible translations, development of polylingual topic models.

**Keywords.** Text corpora, topic modelling, parallel and comparable texts.

### Введение

Параллельные двуязычные и многоязычные корпусы текстов являются ценнейшими источниками данных для систем машинного перевода, для поддержки и развития их лексикографических блоков. Особенно важны качественно подготовленные параллельные корпусы в терминоведческой практике, в ходе создания терминологических переводных словарей, при синхронизации терминов отдельных предметных областей [Беляева 2015]. Вместе с тем создание параллельных корпусов — это трудоемкий и дли- тельный процесс, требующий скрупулезной работы по поиску пар пере- водных текстов, их предобработке, выравниванию, морфосинтаксической

1 Исследование выполняется в рамках проектов «Разработка лингвистического комплек- са для автоматического семантического анализа русскоязычных корпусов текстов с при- менением статистических методов» (проект РФФИ № 16-06-00529) и «Квантитативные лингвистические параметры определения стилевых характеристик и предметной области текстов» (НИР СПбГУ, шифр ИАС НИД 30.38.305.2014).

© МИТРОФАНОВА Ольга Александровна. — Mitrofanova Olga. Санкт-Петербургский государственный университет, Россия. — St. Petersburg State University, Russia. E-mail: [o.mitrofanova@spbu.ru](mailto:o.mitrofanova@spbu.ru)

аннотации и т. д. По этой причине при решении ряда задач компьютерной лингвистики допустимо обращение к сопоставимым текстам, подготовка которых может быть проведена в более сжатые сроки и с меньшими затратами. Основные требования, предъявляемые к сопоставимым кор- пусам — это содержательная близость входящих в них текстов, единство выборочной схемы, сбалансированность блоков корпуса для разных языков. В зависимости от того насколько последовательно соблюдаются эти требования, можно говорить о «шумных» параллельных, строго сопо- ставимых и слабо сопоставимых текстах [Джепа 2015].

Не так давно на вооружении компьютерных лингвистов появились тематические модели — новые статистические инструменты исследования структурной организации текстов. Тематические модели призваны объяс- нять взаимозависимость множества документов корпуса, множества слов в документах и набора тем, отражающих содержание документов. В ходе тематического моделирования документам и словам корпуса ставится в соответствие семейство вероятностных распределений на множестве тем [Daud et al. 2010]. Известны многие разновидности тематических моделей: латентно-семантический анализ (LSA), вероятностный латентно- семантический анализ (pLSA), латентное размещение Дирихле (LDA) и ряд других. В силу малых ограничений и высокой производительности алгоритм LDA и его модификации оказываются востребованными, в том числе и при работе с параллельными и сопоставимыми корпусами текстов.

### Определение сопоставимости корпусов текстов с помощью тематических моделей

В исследовательской практике используются разнообразные методы формирования сопоставимых корпусов (выделение ключевых слов и их переводных эквивалентов для автоматического поиска близких по тематике документов, близость частотных распределений терминов, являющихся переводными эквивалентами, в парах текстов и т. д.), а также критерии оценки сопоставимости текстов (общность сюжета, близость сюжета, общий аспект сюжета, общая терминология, общие конструкции с именованными сущностями и т. д.).

Одним из наиболее привлекательных решений является тематическое моделирование с опорой на двуязычную модель BiLDA (bilingual LDA), которое позволяет связывать близкие по смыслу документы в двуязычном корпусе с точки зрения распределения общих тем. BiLDA — двуязычная модификация вероятностной тематической модели LDA. В BiLDA учи- тывается, что исходный корпус сформирован из пар параллельных или сопоставимых текстов с выравниванием фрагментов.

Так, в уже имеющемся двуязычном корпусе можно вычислить рас- пределение тем по документам и на основе этих статистических данных

отобрать для сопоставимого корпуса пары наиболее близких друг к другу текстов.

Так, в работе [Zede Zhu et al. 2013] представлены результаты, полу- ченные с помощью модели BiLDA, свидетельствующие о сопоставимости текстов в двуязычных китайско-английских корпусах, один из которых сформирован из новостных источников (объем свыше 3 тыс. пар доку- ментов), а второй — из статей Википедии (объем свыше 8 тыс. пар доку- ментов). Авторы указывают, что из множества критериев сопоставимости текстов оценивались прежде всего общность сюжета и близость сюжетов пары текстов.

Альтернативой использования модели BiLDA может быть построение одноязычной тематической модели для текстов на языке-источнике с даль- нейшим автоматическим переводом на язык-цель. В исследовании [Priss 2012] задача отбора сопоставимых английско-чешских пар новостных текстов решается именно таким образом. В качестве дополнительного критерия фильтрации текстов выступает индексирование конструкций с именованными сущностями в парах документов.

### Отбор пар слов — кандидатов в переводные эквиваленты на основе параллельных и сопоставимых текстов

Современные статистические методы анализа двуязычных и много- язычных корпусов текстов делают возможным автоматическое выделение пар слов или конструкций, являющихся вероятными переводными экви- валентами. Экспериментально подтверждено, что тематические модели являются дополнительным источником информации, необходимым для улучшения качества поиска переводов в тех условиях, когда недоступны словари систем машинного перевода.

В исследовании [Vulic, De Smet, Moens 2011] предполагается, что пара лемм является парой возможных переводных эквивалентов, если они регу- лярно встречаются в темах, общих для текстов языка-источника и языка- цели. Для доказательства данной гипотезы используется BiLDA. Авторы [Vulic, Moens 2012] проводили эксперименты с итальянско-английским корпусом текстов из Википедии объемом свыше 18 тыс. пар документов, а также с голландско-английским корпусом слушаний Европарламента объемом свыше 6 тыс. пар документов и голландско-английским корпу- сом текстов из Википедии объемом свыше 7 тыс. пар документов, причем каждая из пар выровненных документов имеет общую тематику. Темати- ческая модель строилась только на лемматизированных существительных, встречающихся в корпусе, при этом тесты проводились для частотных существительных.

Для вычисления наиболее вероятных кандидатов в переводные экви-

валенты использовались три критерия: (1) расстояние Кульбака-Лейблера,

оценивающее близость распределений слов-кандидатов по темам;

1. метод стимулов, определяющий вероятность того, что слова языка-цели будут выявлены в семантически соотносимых темах как реакции на слова из языка-источника; (3) коэффициент *TF*×*ITF* (term frequency — inverse topic frequency), рассчитываемый применительно к векторам слов в про- странстве тем. Помимо этого, авторы проводили эксперименты с комбина- циями данных критериев. Критерии (2) и (3), используемые отдельно и в комбинации, оказались более надежными по сравнению с критерием (1). Следует отметить, что во всех сериях экспериментов модель BiLDA позволяла получать для пары языков кластеры семантически связанных слов. Эти кластеры включали в себя как синонимы и их вероятные пере- водные эквиваленты, так и пары слов, связанные более широкими отно- шениями, что может быть полезно в тех случаях, когда перевод слова

необходимо восстанавливать по контекстной информации.

### Многоязычные тематические модели для корпусов параллельных и сопоставимых текстов

Тематические модели оказались продуктивными при работе не только с двуязычными, но и с многоязычными корпусами текстов. Проведению экспериментов с многоязычными данными способствовало появление корпуса слушаний Европарламента, корпуса документов Евросоюза, тематически сопоставимых текстов Википедии, и некоторых других источников. Однако методы, хорошо зарекомендовавшие себя в двуязыч- ной ситуации, оказываются недостаточно мощными при работе со столь сложным образом организованными ресурсами. В частности, при добав- лении параметров для новых языков в двуязычную тематическую модель увеличивается ее вычислительная сложность. По этой причине в иссле- дованиях предлагаются особые варианты тематического моделирования, одним из которых является многоязычная тематическая модель PLTM (polylingual topic model).

Модель PLTL является расширенной версией LDA. Особенность PLTL состоит в том, что единицей описания в ней является не отдельный доку- мент со своим собственным распределением на темах, а множество слабо- эквивалентных документов на разных языках, описываемое семейством распределений на темах. В итоге модель PLTM при заданных параметрах эксперимента порождает не один набор тем для корпуса, а *L* множеств лингвоспецифичных тем.

В работе [Mimno et al. 2009] представлены результаты серии экспе- риментов по построению многоязычных тематических моделей корпуса параллельных текстов слушаний Европарламента на 11 языках и корпуса сопоставимых текстов Википедии на 12 языках, общим объемом свыше 120 млн с/у. Сопоставительный анализ состава тем, сформированных для

корпуса параллельных текстов, подтвердил, что в темы входят преиму- щественно переводные эквиваленты. Важно, что по наполнению тем можно судить о специфике употребления как полнознаменательных, так и служебных слов в разных языках, а также особых словообразовательных классов слов (например, сокращений, многоосновных слов) и т. д. Экс- перименты с сопоставимыми текстами из Википедии, общим объемом порядка 150 млн с/у, подтверждают результаты работы с параллельными текстами, однако дополнительно позволяют определить меру расхождения между языками на уровне словаря (особенно это касается лексических способов обозначения реалий) и выявить различия в приоритетности тем для анализируемых языков.

### Исследовательские перспективы

Тематическое моделирование параллельных и сопоставимых корпу- сов текстов — это молодое направление, обещающее принести значимые научные результаты. По мнению экспертов в этой области, использование сопоставимых корпусов наряду с параллельными текстами может облег- чить этап подготовки текстовых данных в ходе инструментального иссле- дования различных языков. Очевидно, что тематические модели являются дополнительным ресурсом для систем машинного перевода и в некоторых случаях могут являться прототипом многоязычного машинного словаря. Хочется выразить надежду на то, что в ближайшее время появятся данные о привлечении русскоязычных корпусов текстов в построении

двуязычных и многоязычных тематических моделей корпусов.

**Литература**

1. *Беляева Л.Н.* (2015), Словари систем машинного перевода и параллельные корпуса текстов: проблемы корреляции // Труды международной конференции

«Корпусная лингвистика — 2015», с. 102–110.

1. *Джепа Т.Л.* (2015), Критерии сопоставимости двуязычного корпуса тек- стов // Труды международной конференции «Корпусная лингвистика — 2015», с. 181–188.
2. *Daud A., Li J., Zhou L., Muhammad F.* (2010), Knowledge discovery through directed probabilistic topic models: a survey // Frontiers of Computer Science in China. Vol. 4. № 2, pр. 280–301.
3. *Mimno D., Wallach H.M., Naradowsky J., Smith D.A., McCallum A.* (2009), Polylingual Topic Models // Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. Singapore, 6–7 August 2009, pр. 880–889.
4. *Preiss J.* (2012), Identifying Comparable Corpora Using LDA // 2012 Confer- ence of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Montréal, Canada, June 3–8, pр. 558–562.
5. *Vulić I., De Smetand W., Moens M.-F.* (2011), Identifying Word Translations from Comparable Corpora Using Latent Topic Models // Proceedings of the 49th Annual

Meeting of the Association for Computational Linguistics: short papers. Portland, Oregon, June 19–24, рp. 479–484.

1. *Vulić I. Moens M.-F.* (2012), Detecting Highly Conﬁdent Word Translations from Comparable Corpora without Any Prior Knowledge // Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. Avignon, France, April 23–27, рp. 449–459.
2. *Zede Zhu, Miao Li, Lei Chen, Zhenxin Yang.* (2013), Building Comparable Cor- pora Based on Bilingual LDA Model // Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Soﬁa, Bulgaria, August 4–9, рp. 278–282.

**Literature**

1. *Belyaeva L.N.* (2015), Slovari sistem mashinnogo perevoda i parallel’nye kor- pusa tekstov: problemy korrelyatsii // Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2015», s. 102–110.
2. *Dzhepa T.L.* (2015), Kriterii sopostavimosti dvuyazychnogo korpusa tekstov // Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2015», s. 181–188.
3. *Daud A., Li J., Zhou L., Muhammad F.* (2010), Knowledge discovery through directed probabilistic topic models: a survey // Frontiers of Computer Science in China. Vol. 4. № 2, pр. 280–301.
4. *Mimno D., Wallach H.M., Naradowsky J., Smith D.A., McCallum A.* (2009), Polylingual Topic Models // Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. Singapore, 6–7 August 2009, pр. 880–889.
5. *Preiss J.* (2012), Identifying Comparable Corpora Using LDA // 2012 Confer- ence of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Montréal, Canada, June 3–8, pр. 558–562.
6. *Vulić I., De Smetand W., Moens M.-F.* (2011), Identifying Word Translations from Comparable Corpora Using Latent Topic Models // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: short papers. Portland, Oregon, June 19–24, рp. 479–484.
7. *Vulić I. Moens M.-F.* (2012), Detecting Highly Conﬁdent Word Translations from Comparable Corpora without Any Prior Knowledge // Proceedings of the 13th Con- ference of the European Chapter of the Association for Computational Linguistics. Avignon, France, April 23–27, рp. 449–459.
8. *Zede Zhu, Miao Li, Lei Chen, Zhenxin Yang.* (2013), Building Comparable Cor- pora Based on Bilingual LDA Model // Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Soﬁa, Bulgaria, August 4–9, рp. 278–282.

***В.Н. Стриганов, Г.С. Трефилова***

**ПРОГРАММА ОБРАБОТКИ АНГЛОЯЗЫЧНЫХ ТЕКСТОВ КОРПУСА SPBEFL LC**

**Аннотация.** В статье затрагиваются проблемы автоматической морфологи- ческой разметки и разметки ошибок в учебном корпусе с помощью программы обработки англоязычных текстов корпуса SPbEFL LC.

**Ключевые слова.** Корпус ученических текстов, автоматическая морфологи- ческая разметка, поиск ошибок, разметка ошибок, шаблон.

***V.N. Stryganov, G.S. Treﬁlova***

## SPBEFL LC TEXTS PROCESSING PROGRAM

**Abstract.** This paper deals with automatic morphological annotation and error tagging in a learner corpus using SPbEFL LC Texts Processing Program.

**Keywords.** Learner corpus, morphological annotation, error mining, error tagging, pattern.

**Специфика задач программ обработки текстов в учебных корпусах** Любой существующий корпус является корпусом только в том случае, если он снабжен лингвистической и/или нелингвистической разметкой, которая позволяет осуществлять поиск и извлечение информации, необ-

ходимой пользователю.

Учебный корпус — это база для извлечения информации об ошибках в речи на иностранном и втором языке. Отличительная особенность такого корпуса — большой объем ненормализованных текстов, что вызывает проблемы с автоматической и ручной разметкой корпуса, поэтому задачи разметки актуальны для разработчиков корпуса ученических текстов SPbEFL LC [Камшилова 2015]. Работа с корпусом не предполагает норма- лизации текстов, поэтому предварительная обработка не включает в себя ни исправление ошибок, ни лемматизацию.

Решение проблем разметки позволит использовать корпус в первую очередь для выявления типичных ошибок обучающихся при изучении иностранного языка, для совершенствования дидактических материалов и методики преподавания английского языка, преодоления недостатков учебных материалов прошлого поколения.

Поэтому одной из приоритетных задач разработчиков корпуса SPbEFL LC является подбор или создание инструмента автоматической

© СТРИГАНОВ Виктор Николаевич. — Stryganov Viktor. Вологодский государственный университет, Россия. — Vologda State University, Russia. E-mail: [vitek161](mailto:vitek161147@yandex.ru)[147@yandex.ru](mailto:147@yandex.ru)

© ТРЕФИЛОВА Галина Сергеевна. — Treﬁlova Galina. Российский государственный педа- гогический университет им А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [tinkey@yandex.ru](mailto:tinkey@yandex.ru)

морфологической разметки и определение процедуры автоматического извлечения ошибок, их дальнейшей классификации и разметки.

Проблему разметки ошибок разработчики учебных корпусов решают по-разному, в зависимости от специфики корпуса. На сегодняшний день в свободном доступе нет универсальной программы обработки ненормали- зованных текстов, которой могли бы пользоваться разработчики учебных корпусов для морфологической разметки и разметки ошибок (в рамках хотя бы ограниченного списка). Для этого требуется создание специальной программы с понятным интерфейсом и минимальным набором процедур автоматической обработки ненормализованных текстов.

### Морфологический анализ и разработка шаблонов для извлечения ошибок

В случае нашей программы морфологическая разметка является основой для поиска и разметки ошибок. Теги морфологической разметки являются базой для построения шаблонов ошибок.

Шаблон ошибки представляет собой заданный набор тегов, позво- ляющий автоматически извлекать определенный тип ошибки. Список ошибок, поиск которых удалось формализовать, основан на анализе извлеченных из корпуса именных и глагольных групп [Камши- лова 2015, Трефилова 2014] и включает в себя следующие типы: несо- гласование вспомогательного и смыслового глаголов в числе и лице, несогласование подлежащего со сказуемым в числе и лице, употре- бление полного инфинитива после модальных глаголов, употребление неопределенного артикля с существительным во множественном числе. Этот список пополняется.

Например, для случаев типа *can to [spend] with pet* (контекст глаголь- ной группы) шаблон имеет следующий вид:

if tokens [i][1]='MD' and tokens[i+1][1]='TO'

С помощью этого условия программа будет искать все случаи упо- требления частицы *to* после любого модального глагола.

В именных группах достаточно просто формализовать поиск ошибок на употребление неопределенного артикля с существительным во множе- ственном числе. Для случаев типа *become* ***a*** *real [friends], had* ***a*** *three [cats]* шаблон имеет следующий вид:

*if tokens [i][1]= ‘AT’ and (tokens [i+1][1]=’NNS’ or tokens[i+1] [1]=’NPS’)*

С помощью такого условия программа будет искать все случаи употре- бления неопределенного артикля с существительным во множественном числе на расстоянии одного токена. Если перед существительным стоит атрибут, то позиция существительного смещается на один два токена вправо, что легко задается в этом же условии.

‒

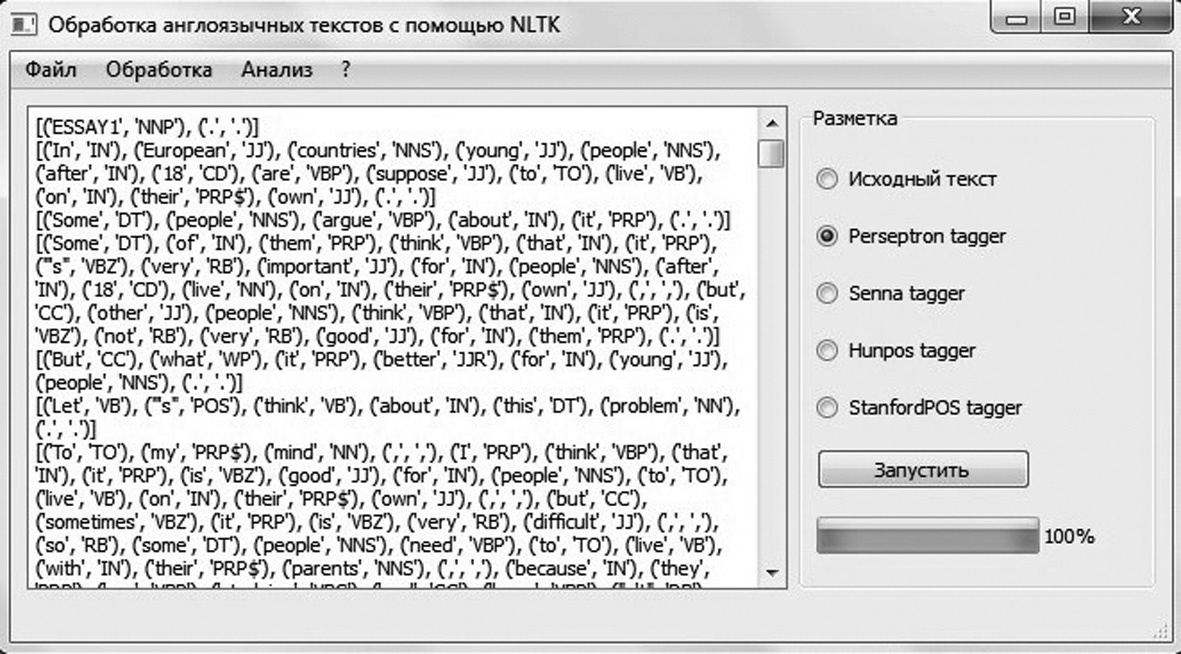
Предполагается, что ошибочным сочетаниям можно будет приписы- вать теги разметки ошибок автоматически или вручную в окне программы, однако этот модуль на данный момент находится в стадии разработки.

### Описание программы

Программа обработки англоязычных текстов корпуса SPbEFL LC написана на языке программирования Python с использованием пакета библиотек и программ для символьной и статистической обработки есте- ственного языка NLTK [NLTK 3.0 Documentation, Эл. Ресурс].

На данный момент программа состоит из трех модулей: нормализа- ция1, морфологический анализ, извлечение ошибок. Ведется работа над модулем разметки ошибок.

Модуль нормализации приводит тексты корпуса к виду, читаемому программой: 1 предложение — 1 строка.

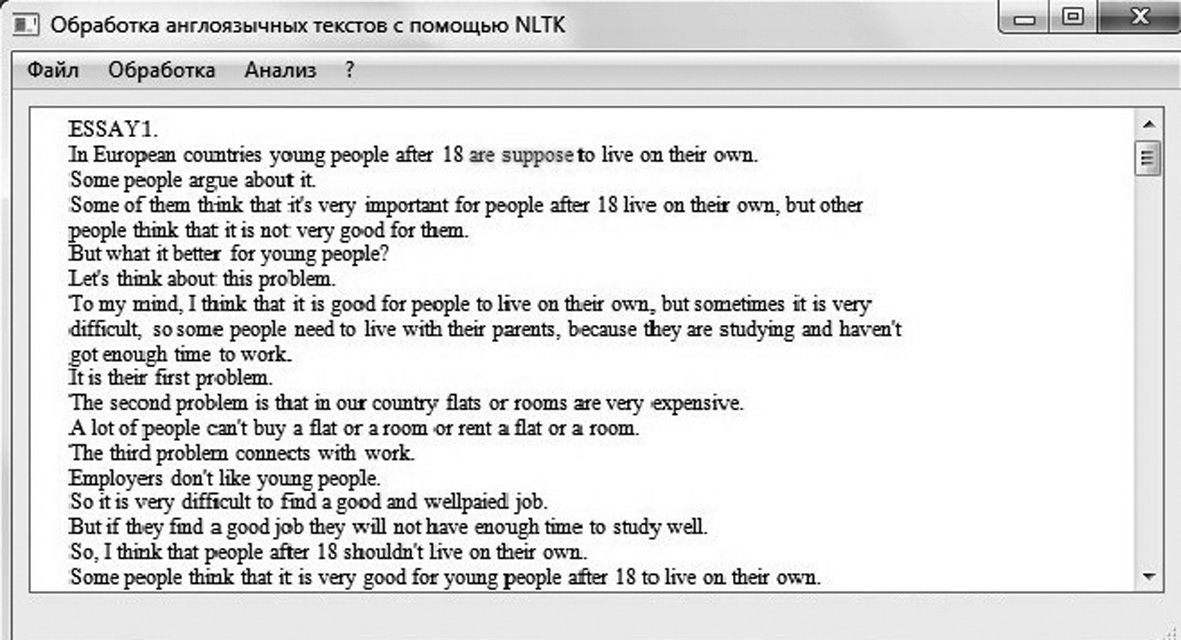
Модуль морфологического анализа приписывает каждому токену в тексте тег с частью речи. Следует отметить, что для морфологического анализа в данном модуле можно использовать 4 разных теггера из библио- теки NLTK: Perceptron tagger, Senna tagger, Hunpos tagger, StangordPos tagger (рис.1).

*Рис.1.* Автоматическая морфологическая разметка с помощью Perceptron tagger

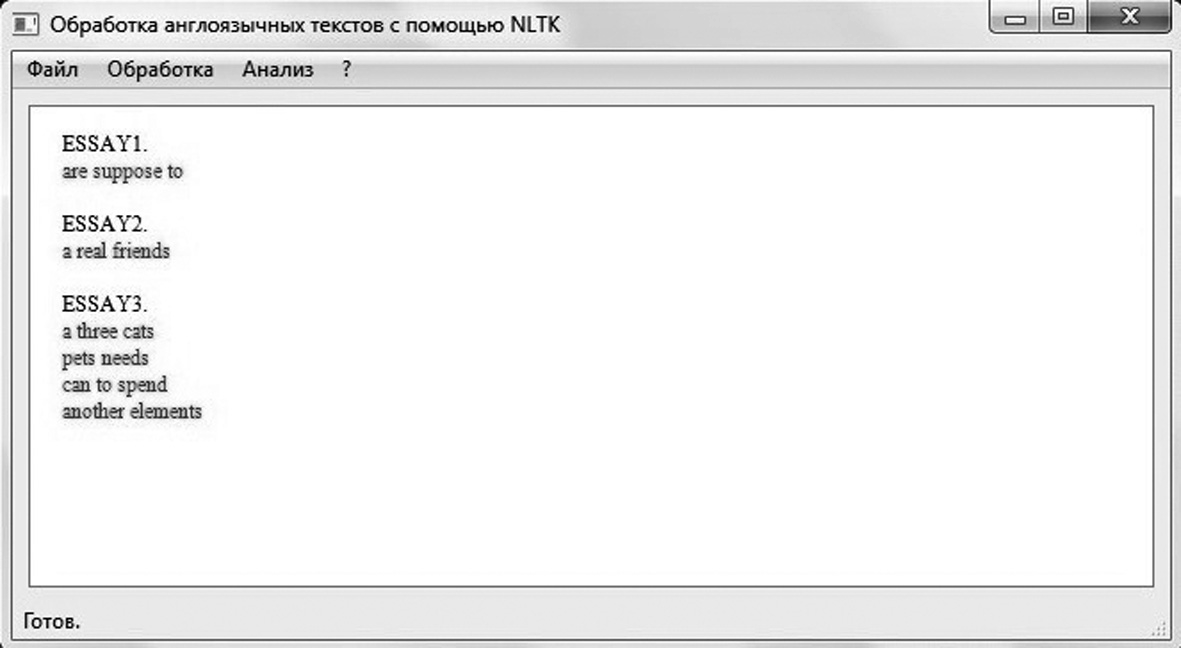
Модуль извлечения ошибок работает с уже размеченным текстом, который прошел ручную коррекцию (снятие омонимии, приписывание тегов словам, которые программа не определила в силу их неверного

1 Нормализация в программе не предполагает коррекцию орфографических и граммати- ческих ошибок, а также лемматизацию, поскольку это приведет к потере уникальности корпуса и его назначения.

написания). Вывод ошибок можно реализовать двумя способами: выделе- ние их в загруженном ранее тексте цветом [рис. 2] или извлечение только заданных ошибочных сочетаний в виде списка [рис. 3]. У каждого способа вывода ошибок есть свое назначение. Выделение ошибки в тексте позво- ляет увидеть ее контекст, список позволяет работать непосредственно с ошибочными сочетаниями.



*Рис.2.* Выделение ошибок в тексте (в окне выдачи цветом выделено *are suppose to*)



*Рис.3.* Вывод ошибок в виде списка

### Заключение

Автоматическая обработка текстов корпуса SPbEFL LC оказалась возможна благодаря использованию NLTK, а анализ извлеченных ранее именных и глагольных групп на наличие ошибок позволил формализовать поиск некоторых типов ошибок по всему корпусу.

Модуль разметки ошибок на данный момент находится в стадии раз- работки.

Автоматический поиск и извлечение ошибок из корпуса позволяет предполагать, что автоматическая разметка ошибок возможна хотя бы в рамках ограниченного списка. Но даже в рамках этого списка возникают трудности, связанные с формализацией извлечения ошибок, например на употребление неопределенного артикля, поскольку артикль и суще- ствительное зачастую разделены одним или несколькими атрибутами.

Таким образом, даже выдача ошибок по заранее заданному шаблону требует ручной проверки, после которой можно говорить об автоматиче- ской разметке ошибок.

**Литература**

* 1. *Камшилова О.Н., Трефилова Г.С.* (2015), Шаблоны для автоматического поиска ошибок в именных и глагольных группах (анализ корпуса ученических текстов). Труды международной конференции «Корпусная лингвистика — 2015», с. 245–252.
  2. *Трефилова Г.С.* (2014), Механизмы автоматического извлечения и разметки в учебном корпусе SPbEFL LC, 51 с.
  3. NLTK 3.0 Documentation, available at: <http://www.nltk.org/>(дата обращения: 17.09.2016).

**Literature**

1. *Kamshilova O.N., Treﬁlova G.S*. (2015), Shablony dlya avtomaticheskogo poiska oshibok v imennykh i glagol’nykh gruppakh (analiz korpusa uchenicheskikh tekstov). Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika — 2015», s. 245–252.
2. *Treﬁlova G.S.* (2014), Mekhanizmy avtomaticheskogo izvlecheniya i razmetki v uchebnom korpuse SPbEFL LC, 51 s.
3. NLTK 3.0 Documentation, available at: <http://www.nltk.org/>(data obrash- cheniya: 17.09.2016).

***Л.В. Татаринова***

**КОРПУСНОЕ ИССЛЕДОВАНИЕ КОНСТРУКЦИИ**

**«КАК БЫ»**

**Аннотация.** Целью исследования является анализ использования конструк- ции «как бы» в устной речи на основе Национального корпуса русского языка, что вызвано ее сверхчастотным употреблением в современной разговорной речи. На основе использования корпуса проводится исследование распространения конструкции «как бы» и употребления ее в качестве семантически опустошен- ной единицы или слова-паразита. НКРЯ послужил базой примеров, при анализе которых, представляется возможным установить, что сверхчастотное употре- бление конструкции привело к расширению, а также полному нивелированию ее первоначального значения — условно-предположительного. Статистические данные помогают проследить, как меняется семантика «как бы» в зависимости от контекста.

**Ключевые слова.** Семантически опустошенная единица, сверхчастотное употребление, значение, неуверенность.

***L.V. Tatarinova***

**CORPUS RESEARCH OF THE CONSTRUCTION “КАК БЫ”**

**Abstract.** The goal of investigation is the analysis of using construction как бы in oral speech on the basis of National Corpus of Russian Language. That was caused by the high frequency usage of the construction in the contemporary spoken language. The focus is on the examples with как бы as the so called semantically empty unit or junk word. The analysis of the examples supplies a clue to answering the question why this construction is widely-used in all spheres of speech. The data allow to follow changes in the semantic meaning of the unit including its primary meaning leveling (conditional and suppositional) according to the context.

**Keywords.** Semantically empty unit, high frequency usage, meaning, uncertainty.

Исследование посвящено проблеме выявления семантически опусто- шенных элементов в разговорной речи, представленных частицей (кон- струкцией) «как бы». Базовую семантику этой частицы фиксирует словарь: она призвана выражать в предложении значение условнопредположитель- ного сравнения [Словарь русского языка 1982]*.* Данная частица может рассматриваться и в качестве союза в современной грамматике русского языка, но граница между функциями этой конструкции в качестве двух частей речи практически не заметна, так как все сводится к выражению предположительности и условности, как было сказано выше.

**©** ТАТАРИНОВА Лариса Викторовна. — Tatarinova Larisa. Иркутский государственный университет, Россия. — Irkutsk State University, Russia. E-mail: [lorata@mail.ru](mailto:lorata@mail.ru)

*Отвечает как бы нехотя* — частица (выражает условное предполо- жение); *зашуршало в траве, как бы прополз кто то* — союз (выражает сравнение) [Словари и энциклопедии на Академике 2014].

Мы рассматриваем «как бы» в функции частицы, учитывая, что в узком смысле частицы — это разряд служебных слов для выражения различных оттенков отдельного слова в предложении или целого пред- ложения [Виноградов 2001].

Для исследования мы прибегаем к помощи корпусных технологий, используя НКРЯ, что позволяет установить характерные семантические признаки данной единицы, рассматривая ее окружение в корпусе. Своего рода банк примеров корпуса служит эмпирической поддержкой для наглядного представления о распространении конструкции с выраже- нием «как бы» и позволяет проследить ее использование в разных сферах жизни, а также для подтверждения тезиса о нивелировании ее первона- чального значения.

Прежде чем непосредственно перейти к анализу языкового материала с применением инструментов корпуса, представляется целесообразным сказать несколько слов об истории возникновения «как бы», семантически опустошенной единицы, выступающей как слово-паразит. Оказывается, что в позднесоветский период «как бы» воспринималось как элемент интеллигентской речи, и его постоянное употребление означало демон- стративный отказ от категоричности суждений. Позднее частое употребле- ние конструкции «как бы» стало проявлением жеманства, свойственного, как принято считать, женской речи. На сегодняшний день это слово прочно

«осело» в речи многих слоев населения [Михейкина 2010], что позволило современному поэту В. Романенко охарактеризовать это явление следую- щим образом: «“Закакбыкала” Россия — что ни фраза, то “как бы” — будто чья-то злая сила речь подняла на дыбы…» [Романенко 2013]*.*

Итак, для объективной картины развития семантической опустошен- ности конструкции «как бы» обратимся в корпусе к поиску точных форм [НКРЯ]. Конструкция «как бы» встречается в 13 748 документах, всего 71 928 вхождений. Для анализа было отобрано немногим более 500 при- меров сплошной выборки первых 50 страниц. В результате из всех примеров 205 содержали в себе конструкцию «как бы», которая в пред- ложении не несла никакой смысловой нагрузки и не выражала присущее ей значение, что составило примерно чуть меньше половины от всех анализируемых примеров.

Данные, полученные от сервиса «Распределение по годам», позволяют судить о распространении исследуемой конструкции относительно вре- менного критерия. Полученный график подтверждает гипотезу современ- ных лингвистов о том, что особое распространение конструкция получила во второй половине XIX в., а именно во времена творческой деятельности Ф.М. Достоевского, на что указывают также данные распределения по

метаатрибутам: этот автор возглавляет список, который ранжирует частот- ность употребления конструкции «как бы» в художественной литературе. Другие статистические данные также небезынтересны, в частно- сти, относительно гендерного показателя: авторы мужского пола более

склонны употреблять «как бы», чем авторы женского пола.

*Таблица.* Гендерный показатель употребления *как бы*

|  |  |  |
| --- | --- | --- |
| **Значение** | **Найдено документов** | **Найдено словоформ** |
| муж | 10052 | 64322 (85,76%) |
| жен | 1743 | 7279 (9,70% |

Безусловно, в данном случае нужно отметить, что для общей стати- стики значения не отбирались, но, учитывая результаты анализа примеров сплошной выборки, есть вероятность того, что на сегодняшний день, ска- жем, в сравнении с позднесоветским периодом, более частое употребление

«как бы» с опустошенным элементом относится к мужской половине. В связи с этим логичными представляются результаты статистики по частотности употребления конструкции в зависимости от тематики: пре- валирует сфера политики и общественной жизни, что совсем не противо- речит реальному положению дел — соотношение между полами в этих областях говорит в пользу мужского. На самом деле конструкция «как бы» в своем «паразитическом» употреблении [Кронгауз 2011:72] так широко распространилась, что ее использование совершенно не зависит от пола или статуса автора.

Чем обусловлена такая популярность этого, как метко заметил совре- менный поэт Сергей Соловьев, «вербального вируса»? («...У каждого времени есть свой вербальный вирус, нечто вроде лихорадки на губе. У нашего поколения — это словечко «как бы». Все реже человек говорит, все чаще — как бы хочет сказать…» [цит. по: Северская 2015]. Если сле- довать мнению, что это дань моде, то почему именно эта частица стала

«образцом»? И.А. Нагорный, исследуя коммуникативно-прагматические характеристики «как бы», считает, что данный речевой элемент как семантически опустошенная частица может функционировать в речи как полноценный языковой компонент, так как фиксирует точку зрения ква- лифицирующего событие субъекта [Нагорный 2015]. Данное утверждение нисколько не противоречит мнению о модальном смысле, который вносит частица в предложение. Вероятно, поэтому М. Кронгауз полагает, что

«как бы» характеризует человека в целом, его психологическое состояние.

«Фактически оно означает, что говорящий отказывается делать резкие и окончательные высказывания о мире, а каждый раз заявляет о своей неуверенности, об отсутствии у него права делать такие утверждения» [Кронгауз 2011:73]*.* Принимая во внимание тот факт, что мир (реальность)

категоризуется в языке, можно предположить, что языковая единица «как бы» манифестирует определенную степень неуверенности в сказанном и является «показателем времени», которое может быть охарактеризовано как нестабильное и неспокойное, поэтому «как бы» является своеобразной переработкой информации о мире.

Таким образом, наблюдение за функционированием «как бы» с так называемым семантически опустошенным элементом, подкрепленное корпусным исследованием, помогает выявить новые особенности в его семантике и объяснить его сверхчастотное употребление в речи.

**Литература**

1. *Виноградов В.В.* (2001), Русский язык (Грамматическое учение о слове). М.: Русский язык.
2. *Кронгауз М.* (2011), Русский язык на грани нервного срыва. 3D, с. 69–78.
3. *Михейкина С.Г*. (2010), Языковые парадоксы нашего времени. Русская речь. М., № 5, с. 62–64.
4. *Нагорный И.А. (2015),* Риски семантической трансформации языковых единиц: лексически опустошенная частица *как бы* и ее функции в речевой сфере: dspace.bsu.edu.ru›…14521…Riski\_semanticheskoy.pdf (дата обращения 20.09.2016).
5. *Национальный корпус русского языка*: [http://www.ruscorpora.ru/,](http://www.ruscorpora.ru/) (дата обра- щения 20.09.2016).
6. *Романенко В*. (2013), Жемчужины мысли: omiliya.org›content/kak-by-stikh. html (дата обращения 20.09.2016).
7. *Северская О.И*. (2015), Слова-паразиты: nkj.ru›archive/articles/3530/ (дата обращения 20.09.2016).
8. *Словарь русского языка* (1982), Т. 2. М.: Русский язык.
9. *Словари и энциклопедии на Академике* (2014): [http://dic.academic.ru/searchall.](http://dic.academic.ru/searchall) php?SWord=как+бы&stype=0&btype=1&p=0 (дата обращения 20.09.2016).

**Literature**

1. *Vinogradov V.V*. (2001), Russkiy yazyk (Grammaticheskoe uchenie o slove). M.: Russkiy yazyk.
2. *Krongauz M*. (2011), Russkiy yazyk na grani nervnogo sryva. 3D, s. 69–78.
3. *Mikheykina S.G*. (2010), Yazykovye paradoksy nashego vremeni. Russkaya rech’. M., № 5, s. 62–64.
4. *Nagorny I.A*. (2015), Riski semanticheskoy transformatsii yazykovykh yedinits: leksicheski opustoshennaya chastitsa kak by i eyo funktsii v rechevoy sfere: dspace. bsu.edu.ru›…14521…Riski\_semanticheskoy.pdf (data obrashcheniya 20.09.2016).
5. Natsional’nyi korpus russkogo yazyka: [http://www.ruscorpora.ru/,](http://www.ruscorpora.ru/) (data obrash- cheniya 20.09.2016).
6. *Romanenko V.* (2013), Zhemchuzhiny mysli: omiliya.org›content/kak-by-stikh. html (data obrashcheniya 20.09.2016).
7. *Severskaya O.I.* (2015), Slova-parazity: nkj.ru›archive/articles/3530/ (data obrashcheniya 20.09.2016).
8. Slovar’ russkogo yazyka (1982), T. 2. M.: Russkiy yazyk.
9. Slovari i entsiklopedii na Akademike (2014): [http://dic.academic.ru/searchall.](http://dic.academic.ru/searchall) php?SWord=kak+by&stype=0&btype=1&p=0 (data obrashcheniya 20.09.2016).

***М.В. Хохлова***

***\*К ВОПРОСУ О «ЗАБЫТЫХ» МЕРАХ СВЯЗАННОСТИ ДЛЯ ВЫЯВЛЕНИЯ СЛОВОСОЧЕТАНИЙ**1

**Аннотация.** Статистический аппарат активно используется в корпусной линг- вистике, позволяя исследователям количественно оценить получаемые данные. Наряду с наиболее часто используемыми мерами для вычисления силы связан- ности в рамках словосочетаний, или мерами ассоциации, существуют и иные коэффициенты. В статье представлен обзор некоторых из них.

**Ключевые слова.** Статистический аппарат, меры ассоциации, меры связан- ности, MI, мера Дайса, мера Жаккара, MS, z-score, сочетаемость.

***M.V. Khokhlova***

**ON RARELY USED STATISTICAL TOOLS FOR MEASURING COLLOCABILITY**

**Abstract.** Statistical tools play an active role in corpus linguistics supplying researchers with quantitative evaluation of the represented results. Along with the most frequently implemented association measures there are a number of other coefﬁcients that we dwell on in the paper.

**Keywords.** Statistical tools, measures of association, MI, Dice coefﬁcient, Jaccard coefﬁcient, minimum sensitivity, z-score, collocability.

В России первые серьезные попытки применить статистические методы на языковом материале были предприняты в начале XX в. исследо- вателями Н.А. Морозовым и А.А. Марковым. Направление, посвященное сопряжению статистических методов и традиционной лингвистики, полу- чило новый импульс в связи с появлением технических возможностей, позволивших привлечь вычислительные средства к решению ряда задач, в том числе получить новые данные и проверить гипотезы на большом материале.

В последнее время в связи с возросшей потребностью в автоматизиро- ванных системах большое внимание уделяется исследованиям, посвящен- ным коллокациям — словосочетаниям, которые обладают определенной воспроизводимостью в речи. Существуют различные статистические метрики для оценки сочетаемости слов. Иными словами, речь идет о статистической неслучайности словосочетаний, которая подвергается

1 Статья подготовлена в рамках работы по гранту Президента Российской Федерации для государственной поддержки молодых российских ученых № МК-5274.2016.6 «Исследова- ние статистических закономерностей сочетаемости лексических единиц».

© ХОХЛОВА Мария Владимировна. — Khokhlova Maria. Санкт-Петербургский госу- дарственный университет, Россия. — St. Petersburg State University, Russia. E-mail: [m.khokhlova@spbu.ru](mailto:m.khokhlova@spbu.ru)

количественной оценке. Ряд мер получил название мер ассоциации, или ассоциативных мер. Они позволяют вычислять силу связи между элемен- тами словосочетаний и основываются на частотах данных словосочетаний и входящих в них отдельных слов. Таким образом, может быть вычислена некоторая устойчивость, присущая лексическим единицам, позволяющая их расположить на шкале: от свободных сочетаний до фразеологизиро- ванных структур.

Первой работой, в которой стал обсуждаться статистический аппарат применительно к измерению устойчивости словосочетаний, стали труды Симпозиума по методам измерения статистической ассоциации приме- нительно к документам, представленным в машинном виде [Stevens et al. 1965]. В статье [Berry-Rogghe 1973] впервые применяется данный аппарат к извлечению словосочетаний. В работе [Manning, Schütze 1999] уже было дано описание нескольких мер, оценивающих связанность. В исследова- нии [Evert 2005] приводятся данные о статистическом аппарате, который используется для вычисления силы связанности компонентов словосоче- таний. На сегодняшний день наиболее полное описание статистического инструментария приводится в монографии [Pecina 2009]. Автор делит меры на три группы: 1) меры, учитывающие наблюдаемые и ожидаемые частоты слов, образующих словосочетание; 2) меры, оценивающие энтро- пию между словами; 3) меры, учитывающие контекст. В данной работе также иллюстрируется применение статистических мер на материале чешского и английского языков.

В качестве механизма для фильтрации словосочетаний используются

пороговые значения мер, списки стоп-слов и др. В наиболее простом случае разделение выделенных словосочетаний носит характер бинарного (коллокации и неколлокации). Наиболее часто в литературе, исследующей понятие коллокации и вычисление силы связанности, упоминаются меры MI, t-score и log-likelihood. В свою очередь, нам бы хотелось обратиться к метрикам, о которых говорится меньше, при этом некоторые из них являются производными от вышеперечисленных.

*MI* (коэффициент взаимной зависимости, взаимная информация, объем информации) сравнивает зависимые контекстно-связанные частоты с неза- висимыми, как если бы слова появлялись в тексте совершенно случайно:

*MI* (*n*, *c*)  log *f* (*n*, *c*)  *N* ,

2

*f* (*n*)  *f* (*c*)

где

*n* — ключевое слово; *c* — коллокат; *f* (*n*,*c*) — частота встречаемости ключе- вого слова *n* в паре с коллокатом *с*; *f* (*n*), *f*(*c*) — абсолютные (независимые) частоты ключевого слова *n* и слова *c* в корпусе (тексте); *N* — общее число словоформ в корпусе (тексте).

Мера *MI* вычисляет вероятность появления двух встречающихся вме- сте слов путем сравнивания произведения их частот в корпусе с наблю- даемыми частотами их совместной встречаемости. Разница между этими величинами выявит степень значимости их встречаемости.

Теоретически значения меры варьируются в пределах (– ∞; + ∞), хотя на практике оказываются ограниченными объемом корпуса. Если значение *MI* (*n*,*c*) больше 1, данное сочетание слов считается статистически значи- мым. В случае если *MI* (*n*,*c*) примерно равно 0, сочетание слов является менее статистически значимым, т. е. данные единицы появляются в паре крайне редко. *MI* (*n*,*c*) меньше 0 означает, что *n* и *c* находятся в отношении дополнительной дистрибуции. Вопрос о том, какие значения *MI* следует считать пороговыми, остается открытым. Так, в работе [Biber, Conrad, Reppen 1998: 266] говорится, что значения, намного превышающие 0, свидетельствуют, что слова встречаются не случайно.

Недостаток данной метрики заключается в том, что вес каждой отдель- ной коллокации тем больше, чем реже она встречается. Поэтому в случаях, когда частота сочетания мала, использование данной формулы может при- вести к неправильным результатам. Чтобы решить эту проблему, в ряде работ были предложены модификации данной меры.

*Дополненный (увеличенный) коэффициент взаимной информации* (augmented mutual information) учитывает частоту появления в корпусе одного из рассматриваемых слов при отсутствии другого и был предложен в работе [Zhang et al. 2008]:

*AugmentedMI* (*n*, *c*)  log2 *f* (*n*, *c*)  *N*

*f* (*n*, *~~c~~* )  *f* (*~~n~~*, *c*)

*Истинная взаимная информация* (true mutual information) призвана увеличивать вес высокочастотных словосочетаний, и для этого исполь- зуется произведение частоты появления рассматриваемого сочетания в корпусе и самой меры *MI*:

*TrueMI* (*n*, *c*)  *f* (*n*, *c*)  log *f* (*n*, *c*)  *N*

2

*f* (*n*)  *f* (*c*)

В работе [Oakes 1998: 171–172] эмпирически выводится формула, в которой величина *f*(*n*,*c*) возводится в куб (*MI*3-score):

*MI* 3  log2

*f* 3 (*n*, *c*)  *N* ,

*f* (*n*)  *f* (*c*)

Существует еще одна модификация данной меры, предложенная А. Килгарриффом, — мера log–log как расширение формулы *MI* [Evert 2004]. Для того чтобы уменьшить тенденцию, согласно которой низкоча-

стотные слова, встречаясь вместе, получают высокий ранг, MI умножается на log2 [*f*(*n*,*c*)]:

log  log  *MI*  log2  *f* (*n*, *c*)

*Мера Дайса* (Dice) связана с гармоническим средним и также пришла из информационного поиска, ее применение к языковым данным получило описание в работах [Smadja 1993; Dias et al. 1999].

*Dice =* 2 *f* (*n*, *c*)

*f* (*n*)  *f* (*c*)

Значение меры варьируется в диапазоне от 0 до 1.

Данная мера, как и MI, имеет тенденцию выдавать повышенные значения для редко встречающихся биграмм. В работе [Rychly 2008] был предложен модифицированный вариант меры Dice, который был призван решить эту проблему:

log *Dice*  14  log 2 *f* (*n*, *c*)

2

*f* (*n*)  *f* (*c*)

*Мера Жаккара* (Jaccard) является эквивалентной мере Dice и вычис- ляется как отношение совместной встречаемости двух слов к суммарной встречаемости отдельных слов в корпусе:

*Jaccard* *f* (*n*, *c*)

*f* (*n*, *c* )  *f* (*n*, *c*)  *f* (*n*, *c*)

Значение меры лежит в интервале [0; 1].

*Мера MS* (minimum sensitivity) была предложена в работе [Pedersen, Bruce 1996]. Ее значения колеблются в диапазоне от 0 до 1.

*MS*  min{ *f* (*n*, *c*) , *f* (*n*, *c*) }

*f* (*n*, *c*)  *f* (*n*, *c* ) *f* (*n*, *c*)  *f* (*n*, *c*)

Как отмечается в работе [Evert 2004], данная мера показала наи- лучшие результаты по сравнению с другими и была реализована в ряде систем.

Мера *z-score* (z-оценка) измеряет стандартное отклонение между наблюдаемой частотой встречаемости *n* рядом с *c* и ожидаемой частотой согласно нулевой гипотезе. Если при малых частотах слов в случае *MI* могут быть получены искаженные данные, то *z-score* не выдаст большой величины, поскольку разница между этими частотами двух слов в пере- счете на стандартное отклонение будет мала.

Впервые данная формула для вычисления сочетаемости применялась в работе [Berry-Rogghe 1973]:

*f* (*n*, *c*)  *f* (*n*)  *f* (*c*) *z*  *score* *N*

*f* (*n*)  *f* (*c*)

*N*

Большое значение *z-score* свидетельствует о том, что с определенной долей уверенности ассоциация между словами не случайна. Эта величина тем больше, чем больше *f* (*n*,*c*).

Результаты, полученные при автоматическом извлечении коллокаций, могут найти разнообразное применение. Статистические оценки для силы связанности лексических единиц могут быть использованы при снятии лексической неоднозначности, при поиске переводных эквивалентов в параллельных текстах и корпусах, при идентификации синонимов и антонимов. Также данные о сочетаемости лексических единиц востребо- ваны при разработке алгоритмов проверки текстов, автокоррекции, при построении онтологий, автоматической классификации и кластеризации лексики по группам и др.

**Литература**

1. *Biber D., Conrad S., Reppen R.* (1998), Corpus Linguistics: investigating lan- guage structure and use. Cambridge: Cambridge University Press.
2. *Berry-Rogghe G.* (1973), The Computation of Collocations and their Relevance in Lexical Studies. In The Computer and Literary Studies. Edinburgh: Edinburgh University Press.
3. *Church K., Hanks P.* (1990), Word association norms, mutual information, and lexicography. In Computational Linguistics. Vol. 16. Issue 1, pp. 22–29.
4. *Dias G., Guillore S., Lopes J.* (1999), Language independent automatic acquisi- tion of rigid multiword units from unrestricted text corpora. In Proceedings of of 6ème Conférence Annuelle sur le Traitement Automatique des Langues Naturelles (TALN), 12–17 July, Cargèse, France, pp. 333–339.
5. *Evert S.* (2004), The Statistics of Word Cooccurrences: Word Pairs and Colloca- tions. Dissertation, Institut für maschinelle Sprachverarbeitung, University of Stuttgart, available at: <http://purl.org/stefan.evert/PUB/Evert2004phd.pdf>(дата обращения 21.09.2016)
6. *Manning Ch., Schütze H.* (1999), Foundations of Statistical Natural Language Processing. Massachusetts: MIT Press.
7. *Oakes M.* (1998), Statistics for corpus linguistics. Edinburgh: Edinburgh Uni- versity Press.
8. *Pecina P.* (2009), Lexical Association Measures. Collocation Extraction. Prague: Institute of Formal and Applied Linguistics.
9. *Pedersen T., Bruce R.* (1996), What to infer from a description. Technical Report 96-CSE-04, Southern Methodist University, Dallas, TX.
10. *Rychly, P.* (2008), A lexicographer-friendly association score. In Proceedings of the Second Workshop on Recent Advances in Slavonic Natural Language Processing RASLAN 2008, pp. 6–9.
11. *Smadja F.* (1993), Retrieving collocations from text: Xtract. Computational Linguistics, 19 (1), pp. 143–177.
12. *Stevens M., Giuliano V., Heilprin, L.* (eds.) (1965), Proceedings of the Sympo- sium on Statistical Association Methods for Mechanized Documentation, Washington 1964, volume 269 of National Bureau of Standards Miscellaneous Publication.
13. *Zhang W., Yoshida T., Ho T., Tang X.* (2008), Augmented Mutual Information for Multi-Word Term Extraction. In International Journal of Innovative Computing, Information and Control, Vol. 8 (92), pp. 543–554.

**Literature**

1. *Biber D., Conrad S., Reppen R.* (1998), Corpus Linguistics: investigating lan- guage structure and use. Cambridge: Cambridge University Press.
2. *Berry-Rogghe G.* (1973), The Computation of Collocations and their Relevance in Lexical Studies. In The Computer and Literary Studies. Edinburgh: Edinburgh University Press.
3. *Church K., Hanks P.* (1990), Word association norms, mutual information, and lexicography. In Computational Linguistics. Vol. 16. Issue 1, pp. 22–29.
4. *Dias G., Guillore S., Lopes J.* (1999), Language independent automatic acquisi- tion of rigid multiword units from unrestricted text corpora. In Proceedings of of 6ème Conférence Annuelle sur le Traitement Automatique des Langues Naturelles (TALN), 12–17 July, Cargèse, France, pp. 333–339.
5. *Evert S.* (2004), The Statistics of Word Cooccurrences: Word Pairs and Colloca- tions. Dissertation, Institut für maschinelle Sprachverarbeitung, University of Stuttgart, available at: <http://purl.org/stefan.evert/PUB/Evert2004phd.pdf>(data obrashcheniya 21.09.2016)
6. *Manning Ch., Schütze H.* (1999), Foundations of Statistical Natural Language Processing. Massachusetts: MIT Press.
7. *Oakes M.* (1998), Statistics for corpus linguistics. Edinburgh: Edinburgh Uni- versity Press.
8. *Pecina P.* (2009), Lexical Association Measures. Collocation Extraction. Prague: Institute of Formal and Applied Linguistics.
9. *Pedersen T., Bruce R.* (1996), What to infer from a description. Technical Report 96-CSE-04, Southern Methodist University, Dallas, TX.
10. *Rychly, P.* (2008), A lexicographer-friendly association score. In Proceedings of the Second Workshop on Recent Advances in Slavonic Natural Language Processing RASLAN 2008, pp. 6–9.
11. *Smadja F.* (1993), Retrieving collocations from text: Xtract. Computational Linguistics, 19 (1), pp. 143–177.
12. *Stevens M., Giuliano V., Heilprin, L.* (eds.) (1965), Proceedings of the Sympo- sium on Statistical Association Methods for Mechanized Documentation, Washington 1964, volume 269 of National Bureau of Standards Miscellaneous Publication.
13. *Zhang W., Yoshida T., Ho T., Tang X.* (2008), Augmented Mutual Information for Multi-Word Term Extraction. In International Journal of Innovative Computing, Information and Control, Vol. 8 (92), pp. 543–554.

# АВТОМАТИЧЕСКИЙ И СТРУКТУРНЫЙ АНАЛИЗ ТЕКСТА

***В.С. Андреев***

**СООТНОШЕНИЕ ЭЛЕМЕНТОВ СТИХОТВОРНОГО ТЕКСТА**

**(корреляционный анализ)**

**Аннотация.** В работе рассматриваются соотношения частеречных характери- стик инициальной и финальной частей стихотворной строки с синтаксическими, ритмическими признаками. Материалом является база данных стихотворных текстов Г. Лонгфелло. Соотношения устанавливаются при помощи корреляци- онного анализа.

**Ключевые слова.** Стихотворный текст, корреляционный анализ, инициальная позиция, финальная позиция.

***V.S. Andreev***

**INTERRELATION OF ELEMENTS OF VERSE TEXT**

**(correlation analysis)**

**Abstract.** The study investigates the interrelation of part of speech characteristics of the words in the initial and ﬁnal positions of the verse line with syntactic and rhythmic features. The data source includes lyrics by H.W. Longfellow. Correlation analysis is used to establish the relationship of the parameters.

**Keywords.** Verse text, correlation analysis, initial position, ﬁnal position.

Одной из важных характеристик при анализе стихотворных текстов стало соотношение частей речи в тексте [Гаспаров 2012а; Naumann et al. 2012]. Эти данные помогают установить тип описания тем — ста- тическое либо динамическое [Гаспаров 2012б; Popescu et al. 2013], в целом ряде случаев выявить специфику авторского стиля и авторство [Tambouratzis G. et al. 2004]. Однако представляется важным сделать следующий шаг в изучении структуры текста — выявить взаимосвязь частеречных параметров с синтаксическими, ритмическими и рифмен- ными характеристиками стиха. Этот аспект и был положен в основу настоящего анализа.

© АНДРЕЕВ Вадим Сергеевич. — Andreev Vadim. Смоленский государственный универ- ситет, Россия. — Smolensk State University, Russia. E-mail: [vadim.andreev@gmail.com](mailto:vadim.andreev@gmail.com)

Следует отметить, что при анализе представленности частей речи в тексте могут быть использованы различные подходы. Согласно одному из них части речи подсчитываются по всему тексту без учета их положения в стихотворной строке. В соответствии с другим подходом учитывается локализация частей речи. В этой работе принимается второй подход.

Здесь при морфологическом анализе по выделению частей речи будут учтены только те слова, которые находятся в двух «экстремальных» позициях в стихе (стихотворной строке): в начале стиха и в его исходе. Важность именно этих позиций для стихотворной строки не вызывает сомнений [Гаспаров 1984].

В числе морфологических признаков к исследованию привлекаются следующие параметры: существительное (СУЩ), глагол (ГЛ), прилага- тельное (ПЛГ), наречие (НРЧ), местоимение (МЕСТ), служебное слово (СЛЖБ). Для разграничения позиций, в которых зафиксированы слова, используются обозначения «н» (начало строки) и «к» (исход строки).

Признаками, которые характеризуют стиль стихотворного текста, являются следующие:

1. Синтаксис и поэтический синтаксис: инверсия (полная), сложнопод- чиненное предложение, сложносочиненное предложение, маркированный восклицательным знаком или многоточием исход строки, синтаксический перенос (перенос паузы с исхода строки в ее середину), разрыв строки синтаксической паузой.
2. Ритмика: пропуски ударения на сильных местах (1, 2 иктах и последнем икте), увеличенная в слоговом составе анакруса и отягчение (ударность) анакрусы.
3. Рифма: отклонение от мужской рифмы, т. е. женская или дактили- ческая рифма.

В качестве материала исследования взяты лирические произведения классика американской литературы поэта-романтика Г. Лонгфелло.

В результате подсчетов была получена таблица, строками которой являются локализованные частеречные признаки, а столбцами — соотносимые с ними синтаксические, ритмические и рифменный при- знаки. Полученные количественные данные были использованы для выявления соотношений при помощи корреляционного анализа. Здесь использовалась мера Коула для качественных дихотомических признаков, а также предложенная им методика проверки статистической значимости результатов [Тулдава 1988; Cole 1949].

В таблице приводятся значения коэфициента корреляции между сопо- ставляемыми признаками начала и исхода строки с разноуровневыми при- знаками стиха. Статистически значимые соотношения выделены жирным шрифтом.

Как видно из таблицы, в целом ряде случаев между соотносимыми признаками наблюдаются статистически значимые зависимости.

Прежде всего, имеет смысл определить степень общей коррелирован- ности всех признаков начала и конца. Здесь вырисовывается достаточно интересная картина. Относительно синтаксических признаков неболь- шое преимущество имеется у признаков начала стиха. Относительно ритмических признаков соотношение начала и исхода одинаковое. И только по зависимостям с рифмой более значимой является степень связи исхода.

Последнее достаточно ожидаемо, так как именно в конечной позиции стиха реализуется рифма. С другой стороны, приблизительно одинаковая степень участия признаков начала и конца в ритмической структуре строки стоит отметить особо, так как можно было бы предположить, что именно исход будет преобладать и в этих случаях.

*Таблица***.** Соотношение частеречных характеристик начала и исхода строки и разноуровневых признаков стиха

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Инверсия | Сложно-подчиненное предложение | Сложно-сочиненное предложение | Маркированный исход строки | Перенос | Разрыв | Икт 1 | Икт 2 | Икт последний | Анакруса длина | Анакруса отягчение | Отклонение  от мужской рифмы |
| СУЩ-н | – 0,01 | **– 0,17** | **– 0,19** | 0,04 | **– 0,17** | 0,01 | **– 0,67** | **0,22** | 0,13 | **0,25** | – 0,02 | – 0,01 |
| Гл-н | 0,11 | **0,06** | **0,21** | 0,02 | **– 0,31** | 0,04 | – 0,02 | 0,04 | 0,00 | **– 0,25** | 0,01 | **– 0,25** |
| ПЛГ-н | – 0,24 | **– 0,47** | 0,00 | – 0,17 | 0,01 | **– 0,35** | **– 0,76** | – 0,19 | 0,03 | 0,05 | **– 0,55** | 0,05 |
| НРЧ-н | – 1,00 | 0,05 | – 0,20 | – 0,30 | **0,15** | **0,10** | 0,05 | – 0,23 | – 0,31 | **– 0,56** | **0,26** | 0,17 |
| МЕСТ-н | – 0,32 | **0,25** | 0,02 | **– 0,45** | – 0,14 | **0,12** | **– 0,89** | **– 0,69** | – 0,45 | – 0,13 | – 0,08 | 0,00 |
| СЛЖБ-н | 0,04 | – 0,08 | **– 0,76** | 0,03 | **0,28** | **– 0,78** | **0,88** | **– 0,86** | **– 0,63** | **– 0,88** | 0,02 | – 0,13 |
| СУЩ-к | – 0,23 | – 0,25 | – 0,04 | – 0,07 | **0,20** | **– 0,26** | – 0,01 | **0,50** | – 0,10 | **0,15** | 0,04 | **0,30** |
| Гл-к | **0,13** | **0,15** | 0,02 | – 0,18 | **– 0,17** | **0,13** | – 0,08 | **– 0,49** | – 0,03 | **– 0,33** | – 0,12 | **– 0,28** |
| ПЛГ-к | – 0,35 | 0,00 | – 0,05 | 0,01 | 0,01 | **0,06** | **– 0,59** | **– 0,41** | 0,05 | 0,05 | **– 0,34** | 0,02 |
| НРЧ-к | 0,02 | – 0,03 | 0,00 | **0,07** | **– 0,40** | – 0,10 | – 0,01 | **– 0,53** | – 0,50 | 0,02 | **0,08** | **– 0,60** |
| МЕСТ-к | 0,03 | 0,03 | **0,20** | – 0,02 | – 0,43 | **0,17** | – 0,43 | **– 0,73** | **0,08** | – 0,08 | – 0,08 | **– 0,62** |
| СЛЖБ-к | 0,02 | **0,17** | – 0,18 | **0,14** | – 0,26 | – 0,25 | **0,65** | **– 0,53** | 0,01 | **– 0,60** | 0,07 | – 0,34 |

По своим соотношениям с синтаксическими параметрами частереч- ные признаки начала строки в большей степени отрицательно ориенти- рованы, т. е. отражают несовместимость с синтаксическими признаками, ЧР-признаки исхода, напротив, в большем количестве случаев взаимно обусловлены с синтаксическими характеристиками. Относительно рит- мических параметров ЧР обеих локализаций имеют большее количество отрицательных зависимостей, чем положительных. И все же и здесь признаки начала строки имеют большее различие между количеством позитивных и негативных связей (10 отрицательных корреляций и 4 поло- жительные). В исходе это соотношение выглядит как 5:9.

Инверсия оказывается менее тесно связанной с ЧР, чем перенос. Обе эти конструкции достаточно сходны в том, что они образуют два семан- тических фокуса [Dorgeloh 1997]. Однако если инверсия реализует это в горизонтальном плане, то перенос — в вертикальном (объединяя две соседние строки). Получается, что для стихотворного текста более важна вертикальная семантическая структура. Сложноподчиненные предложе- ния более значимы для ЧР, чем сложносочиненные. Еще одним наблюде- нием является крайне высокая степень зависимости с ЧР синтаксического разрыва строки.

Среди ритмических признаков относительно больше корреляций с ЧР имеет пропуск ударения на втором икте. Как и следовало ожидать, между признаками Икт-1 и Икт-2 наблюдается своего рода оппозиция. Положи- тельным корреляциям одного из этих ритмических признаков соответствует либо отрицательная корреляция с той же частью речи, либо отсутствие значимой зависимости. Исключения составляют корреляции с одним при- знаком начала (МЕСТ-н) и одним признаком исхода (ПЛГ-к) строки.

В целом проведенный анализ позволяет сделать следующие выводы. Частеречные признаки двух крайних позиций в строке в значительной сте- пени коррелированы с разноуровневыми параметрами стихотворного тек- ста. При этом статистически значимых зависимостей от всех соотношений немногим меньше половины (45%). Сами корреляции варьируют от весьма тесных до относительно слабых, однако тесные корреляции и корреляции средней силы преобладают. Положительные, взаимно обусловливающие корреляции по числу и тесноте уступают отрицательным корреляциям, которые отражают несовместимость признаков. В процентном отношении первые составляют 48%, вторые — 62%.

**Литература**

1. *Гаспаров М.Л.* (1984), Эволюция русской рифмы // Проблемы теории стиха. Л., с. 3–36.
2. *Гаспаров М.Л.* (2012а), Точные методы анализа грамматики в стихе // М.Л. Гаспаров. Избранные труды. М., Т. 4, с. 23–35.
3. *Гаспаров М.Л.* (2012б), Пушкин и проблемы поэтической формы: язык и стих // Лингвистика стиха. Анализы и интерпретации. М.Л. Гаспаров. Избранные труды. М., Т. 4, с. 117–129.
4. *Тулдава Ю.А.* (1988), Об измерении связи качественных признаков в линг- вистике (1): Сопряженность альтернативных признаков // Квантитативная линг- вистика и автоматический анализ текстов: ученые записки Тартуского ун-та / отв. ред. Juhan Tuldava. Вып. 827, с. 146–162.
5. *Cole L.C.* (1949), The measurement of interspeciﬁc association // Ecology. Vol. 30, pp. 411–424.
6. *Dorgeloh H.* (1997), Inversion in modern English: form and function (Studies in discourse and grammar 6). Amsterdam and Philadelphia: John Benjamins Publishing Co. (Amsterdam and Philadelphia). Vol. 6.
7. *Naumann S. et al.* (2012), Aspects of nominal style / Naumann S., Popescu I.-I., Altmann G. // Glottometrics. Vol. 23, pp. 23–55.
8. *Popescu et al.* **(**2013), Descriptivity in Slovak lyrics / Popescu I.-I., Čech R., Best K.-H., Altmann G. **//** Glottotheory. Vol. 4 (1), pp. 92–104.
9. *Tambouratzis G. et al.* (2004) Discriminating the registers and styles in the Modern Greek language. Part 1: Diglossia in stylistic analysis / Tambouratzis G., Markantonatou S., Hairetakis N., Vassiliou M., Tambouratzis D. and Carayannis G. // Literary and Linguistic Computing. Vol. 19, No. 2, pp. 197–220.

**Literature**

1. *Gasparov M.L*. (1984), Evolyutsiya russkoy rifmy // Problemy teorii stikha. L., s. 3–36.
2. *Gasparov M.L*. (2012a), Tochnye metody analiza grammatiki v stikhe //

M.L. Gasparov. Izbrannye trudy. M., T. 4, s. 23–35.

1. *Gasparov M.L*. (2012b), Pushkin i problemy poeticheskoy formy: yazyk i stikh // Lingvistika stikha. Analizy i interpretatsii. M.L. Gasparov. Izbrannye trudy. M., T. 4, s. 117–129.
2. *Tuldava Yu.A*. (1988), Ob izmerenii svyazi kachestvennykh priznakov v lingvis- tike (1): Sopryazhennost’ al’ternativnykh priznakov // Kvantitativnaya lingvistika i avtomaticheskiy analiz tekstov: uchenye zapiski Tartuskogo un-ta / otv. red. Juhan Tuldava. Vyp. 827, s. 146–162.
3. *Cole L.C.* (1949), The measurement of interspeciﬁc association // Ecology. Vol. 30, pp. 411–424.
4. *Dorgeloh H.* (1997), Inversion in modern English: form and function (Studies in discourse and grammar 6). Amsterdam and Philadelphia: John Benjamins Publishing Co. (Amsterdam and Philadelphia). Vol. 6.
5. *Naumann S. et al.* (2012), Aspects of nominal style / Naumann S., Popescu I.-I., Altmann G. // Glottometrics. Vol. 23, pp. 23–55.
6. *Popescu et al.* **(**2013), Descriptivity in Slovak lyrics / Popescu I.-I., Čech R., Best K.-H., Altmann G. **//** Glottotheory. Vol. 4 (1), pp. 92–104.
7. *Tambouratzis G. et al.* (2004) Discriminating the registers and styles in the Modern Greek language. Part 1: Diglossia in stylistic analysis / Tambouratzis G., Markantonatou S., Hairetakis N., Vassiliou M., Tambouratzis D. and Carayannis G. // Literary and Linguistic Computing. Vol. 19, No. 2, pp. 197–220.

***Г.Н. Гумовская***

**ИССЛЕДОВАНИЕ СЕМАНТИЧЕСКОГО ПОТЕНЦИАЛА ТЕКСТА МЕТОДАМИ ТЕОРИИ ИНФОРМАЦИИ**

**Аннотация.** Статья посвящена изучению семантического потенциала текста с позиций теории информации. С помощью лингвистического моделирования исследованию подлежат системные связи плана содержания. Аналогом лингви- стического объекта являются математически представленные данные о количестве субъективной информации в тексте, являющейся мерой выразительности.

**Ключевые слова.** Теория информации, лингвистическая вероятность, нор- мальное распределение вероятностей, интервокабулярное расстояние, величина вероятности.

***G.N. Gumovskaya***

**RESEARCH OF THE SEMANTIC POTENTIAL OF THE TEXT BY METHODS OF THE THEORY OF INFORMATION**

**Abstract.** The article is devoted to the studying of the semantic potent from the positions of the Theory of information. Subjected to research are systemic ties of the plan of content by means of linguistic modelling. The analog of linguistic object is mathematically presented data on the quantity of the subjective information in the text which is a measure of its expressiveness.

**Key words.** Theory of information, linguistic probability, normal distribution of probabilities, probability size, intervocabulary distance.

Поскольку художественное произведение — текст — в восприятии, т. е. в своем актуальном бытии, оказывается информационным процес- сом, то появляется необходимость дополнить сведения о художествен- ном произведении, используя положения и методы теории информации. **Теория информации** является разделом прикладной математики, отно- сящимся к измерению количества информации и ее свойств, который устанавливает предельные соотношения для систем передачи данных. Как и любая математическая теория, теория информации оперирует матема- тическими моделями, а не реальными физическими объектами. В теории информации используется, главным образом, математический аппарат теории вероятностей и математической статистики.

Каждое словосочетание, нормативное или ненормативное, т. е. всякая комбинация по теории вероятности должна рано или поздно повториться. Степень вероятности соответствующего явления в языке определяет количество избыточной информации, облегчающей догадку

© ГУМОВСКАЯ Галина Николаевна. — Gumovskaya Galina. НИУ Высшая школа эконо- мики, Россия. — National Research University Higher School of Economics, Russia. E-mail: [goumovskayagaln@mail.ru,](mailto:goumovskayagaln@mail.ru) [ggumovskaya@hse.ru](mailto:ggumovskaya@hse.ru)

и тем самым понимание и восприятие речи. Вероятностные ограниче- ния, характеризующие язык, называются **лингвистической вероятно- стью** [Францерт 1982]. Этим понятием охватываются различные явле- ния: вероятность появления в речи отдельных элементов (слов, фонем), тех или иных сочетаний, грамматических моделей. Все это выявляется в статистической характеристике языка.

Вероятность словосочетаний подчиняется двум различным законо- мерностям, поэтому можно говорить о лексической и грамматической вероятностях словосочетаний. Лексическая вероятность определяется валентностью и сочетаемостью слов и зависит от их смысловых значе- ний. Эта зависимость не может считаться сугубо лингвистической, так как она обусловлена объективной действительностью, условиями жизни общества, очень изменчива как территориально, так и темпорально. **Лексическая вероятность** — чисто статистическая, не зависящая от грамматического строя языка.

Как известно, круг сочетаемости языковых единиц всегда более или менее ограничен. Эти ограничения наложены как логическими связями, обусловленными признаками понятия, заключенного в языковой единице в виде семантических компонентов, так и его функциональными показа- телями, т. е. употреблением. Например, следующее высказывание являет пример организации слов в соответствии с лексической валентностью, подтверждая лексическую вероятность.

*Catherine was a freelance translator of Italian commercial texts* [Roberts 2003: 76].

Однако для большей художественной выразительности автор может заставлять языковые единицы вступать в сочетания с другими единицами, с которыми до этого у них не было никаких контактов, т. е. нарушать лек- сическую вероятность. Нарушение структуры на одном участке является важным средством ее информационной активизации. А художественный текст — это не просто реализация структурных норм, но и их нарушение.

**Грамматическая вероятность** может быть двух видов:

* вероятность употребления определенной части речи в данной позиции;
* вероятность появления определенной грамматической формы в дан- ной позиции.

Первый вид имеет большое значение в языках с твердым порядком слов, каковым является английский, это чисто грамматическая законо- мерность.

*Her husband had got used to her running up to London for last-minute meetings with editors and publishers* [Roberts 2003: 76].

Второй вид — в языках с развитой флективной системой (русский язык).

*И тем самым, мечта моей мамы дать мне в руки какую-то хлебную профессию, не состоялась* [Минчин 2013: 229].

Обращение к понятиям теории информации для количественной характеристики художественного эффекта представляется логичным, потому что «информация математически определяется как мера раз- личия и описывает в абстрактном виде структурные отличия. Инфор- мация определяется через различие структур и тем самым через меры структур» [Францерт 1982: 165]. При функциональном сбое семан- тического ритма за счет нарушения сочетаемости языковых единиц определяющими характеристиками являются различия образной и необ- разной семантики и тип структурного отличия прямого значения от переносного.

Количественным параметром, соответствующим качественному параметру взаимодействия мало совместимых семантических потенциа- лов, является **величина вероятности** соединения этих слов в сочетание. Понятие «нормального распределения вероятностей», известное из тео- рии вероятности, связано со следующими тенденциями: во многих явле- ниях наибольшую вероятность имеют события из среднего интервала от 0 до 1, т. е. обычные события, а наименьшую вероятность — события из крайних интервалов, т. е. необычные. На примере текстовых фрагмен- тов рассмотрим, как реализуется принцип «нормальное распределение вероятности» в лингвистике. Актуализированные текстом семы можно назвать семантическим приращением, составляющим разницу между семантическими потенциалами ассимилятивного семантического ритма и ритма с функциональным сбоем [Гумовская 2015]. Это внутривокабу- лярное семантическое приращение, полученное в результате разницы между образной и необразной семантикой в сочетаниях типа *the ﬁre of my interest — the ﬁre of artillery; an eater of literature — an eater of bread, the blandishments of ﬂatter — the blandishments of cheese* поддаются количественному описанию. Художественно-эстетический эффект, достигаемый функциональным сбоем семантического ритма, характери- зуется тремя качественными параметрами: структурным, семантическим и художественно-эстетическим. Каждому параметру можно поставить в соответствие количественную характеристику, получив следственную цепочку, объясняющую истоки выразительности.

За математическую меру, характеризующую интервокабулярную связь

слов с несовместимыми семантическими идеями, целесообразно принять величину D, которая характеризует интервокабулярную связь компонентов словосочетания [1]. Величина D является мерой выразительности данного словосочетания и указывает на интервокабулярное расстояние между его компонентами, принимая значения любого числа из числового промежутка от 0 до 1. Лингвистический смысл нуля на оси интервокабулярных рассто- яний — нулевое расстояние между концептами слов, которое имеет место при сочетании слова с самим собой: *the ﬁre of the ﬁre, the blandishments*

*of blandishments, etc.* Лингвистический смысл единицы — максимально возможное расстояние между концептами сочетающихся слов, которое характерно для «патологических» сочетаний. Интервокабулярное рас- стояние находится по формуле:

*D*  1 2  *O* ,

*A*1  *B*1

где

*О* — количество общих сем; *А*1 — количество сем слова *А*; *В*1 — количество сем слова *В.*

Для наглядности возьмем два сочетания — нормативное и экс-

прессивное — с прилагательным *broad: a broad river* и *a broad word*. Семантическая структура лексемы *broad* может быть представлена в виде набора следующих сем: S1 — материальность, S2 — протяжен- ность в пространстве в направлении меньшего горизонтального размера (ширина), S3 — целостность сплошной поверхности, S4 — степень про- тяженности, S5 — оценка.

Семный состав лексемы *river* соответствует следующим компонен- там значения:

S1 — материальность,

S2 — протяженность в пространстве в направлении меньшего гори- зонтального размера (ширина),

S3 — протяженность в пространстве в направлении большего гори- зонтального размера (длина),

S4 — целостность сплошной поверхности,

S5 — протяженность в вертикальном направлении (глубина), S6 — движение,

S7 — жидкостный состав движущейся материи.

Интервокабулярное расстояние нейтрального сочетания *a broad*

*river*, соответствующее мере его выразительности, имеет следующее математическое значение:

*D =* 1 *–* 2 **·** *О / (A*1 *+ B*1*) =* 1 – 6 / 12 = 1 – 1/2 = 0,5

Семантической структуре слова word соответствует следующий набор сем:

S1 — материальность (звуковая или графическая репрезентация),

S2 — единица знакового кода, S3 — коммуникативная направленность, S4 — продукт деятельности мозга человека.

Констатируя наличие одной лишь общей семы у компонентов соче-

тания a *broad word*, вычислим интервокабулярное расстояние между

ними, соответствующее количеству субъективной информации, пере- даваемой этим сочетанием.

*D =* 1 – 2 **·** *О / (A*1 *+ B*1*) =* 1 – 2 **·** 1/9 = 1 – 2/9 = 0,77

Интервокабулярные расстояния, вычисленные по данной методике, располагаются в средних интервалах числового промежутка 0 – 1, при- чем при ассимилятивном семантическом ритме интервокабулярное расстояние *D*  0,5, а при семантическом ритме с функциональным сбоем *D*  0,77. Значение интервокабулярного расстояния экспрессив- ных аномалий всегда больше, чем в обычных, на некоторую величину. Таким образом, причиной функционального сбоя семантического ритма является достаточно большое — *D*  0,77 — интервокабулярное расстояние слов, сведенных в одну синтагму. Мерой выразительности экспрессивных аномалий, полученных в результате функционального сбоя семантического ритма, можно признать количество субъективной, художественно-эстетической информации. Лингвистический смысл их связи состоит в том, что «если составляющие текст высказывания максимально приближены к общеизвестным, тривиальным, то текст обладает минимальной информативностью» [Жирмунский 1997: 407] и, следовательно, функциональный сбой семантического ритма более информативен, чем ассимилятивный семантический ритм. Это под- тверждается примерами из художественной литературы. Нейтральному сочетанию *a broad river* противостоит экспрессивное *a broad word*, в котором лексема *broad* реализует лексико-семантический вариант

«грубый, непристойный» («крепкое словцо») с дифференцирующим

компонентом «отрицательное качество» и указывает на выходящий за рамки приличия грубый, вульгарный способ проявления эмоций.

*She was not a ﬂirt, not even a coquette — words dear to the heart of his generation, which loved to deﬁne things by a good,* ***broad****, inadequate* ***word*** *— but she was dangerous* [Galsworthy 1960:57].

*The moon made a pathway on the* ***broad river*** *for the light feet of Siva’s bride, and on the further bank a row of palm trees was delicately silhouetted against the sky* [Maugham 1978: 256].

Наше исследование показало, что применительно к лингвистике понятие «нормального распределения вероятностей» имеет следую- щие значения: наиболее вероятным является объединение в сочетание слов в соответствии с их лексической и грамматической валентностью, интервокабулярное расстояние между концептами которых принадлежит среднему интервалу промежутка 01, и наименее вероятно — крайним интервалам. Это подтверждается на практике более частотным форми- рованием синтагм с ассимилятивным семантическим ритмом, чем с дис- симилятивным ритмом с функциональным сбоем.

**Литература**

1. *Буданова Т.А.* (1985), Методы сравнительных лингвистических исследо- ваний. М.
2. *Гумовская Г.Н.* (2015), Гармоническая организация художественного про- изведения. М.: МПГУ.
3. *Зиндер Л.В.* (1958), О лингвистической вероятности: Вопросы языкознания,

№ 2, с. 121–125.

1. *Жирмунский В.М.* (1997), Теория литературы. Поэтика. Стилистика: Избран- ные труды. Л.: Наука.
2. *Францерт П.* (1982), Теория информации как модель отражения. М.- Бер- лин.

**Literature**

1. *Budanova T.A.* (1985), Metody sravnitel’nykh lingvisticheskikh issledovaniy. M.
2. *Gumovskaya G.N*. (2015), Garmonicheskaya organizatsiya khudozhestvennogo proizvedeniya. M.: MPGU.
3. *Zinder L.V*. (1958), O lingvisticheskoy veroyatnosti: Voprosy yazykoznaniya,

№ 2, s. 121–125.

1. *Zhirmunskiy V.M*. (1997), Teoriya literatury. Poetika. Stilistika: Izbrannye trudy. L.: Nauka.
2. *Frantsert P.* (1982), Teoriya informatsii kak model’ otrazheniya. M.- Berlin.

***М. В. Данильчук***

**МАРКЕТИНГОВАЯ ФОНОСЕМАНТИКА В СИСТЕМЕ НЕЙМИНГА**

**Аннотация.** В статье рассматриваются основные типы современного ней- минга, выявленные на материале корпуса брендов российского рынка. Предпола- гается, что в названиях, имеющих нулевое маркетинговое наполнение, первичным для потребителя является звуковой ряд, что должно быть доказано экспери- ментально. Процедура фоносемантического анализа может помочь определить основную информацию, которая соответствует звучанию имени, и установить, как оно воздействует на потребителя с точки зрения маркетинговой фоносемантики. **Ключевые слова.** Фоносемантика, маркетинговая фоносемантика, номина-

ция, нейминг.

***M. V. Danilchuk***

**MARKETING PHONOSEMANTICS IN THE SYSTEM OF NAMING**

**Abstract.** The paper considers the main modern types of naming based on the material of the corpus of brands on the Russian market. It is supposed that the sequence of sounds is prime for the consumer in the names with a zero marketing content but it should be proved by the experiment. The procedure of phonosemantic analysis can help to deﬁne the main information, which has the brand name according to its sound- ing, and its inﬂuence on the consumer for the purposes of marketing phonosemantics.

**Keywords.** Phonosemantics, marketing phonosemantics, naming.

Процесс номинации в лингвистике соответствует процессу нейминга в маркетинге. Механизмы обоих процессов сопоставимы друг с другом, хотя имеют разные цели. Новое имя нужно при появлении новых реалий, при расширении или сужении значений уже имеющихся слов и т. д. [Яга- фарова 2014]. Словообразование — основной способ обогащения словар- ного состава языка. Существуют уже устоявшиеся в языке словообразо- вательные модели, которые являются основными способами образования слов (словопроизводство, словосложение, конверсия). Другим способом расширения словаря является заимствование, причем словообразование здесь также играет важную роль, способствуя адаптации слова. Однако если номинация может давать имя как данной вещи, так и классу вещей одновременно, то нейминг нужен для того, чтобы создать уникальное имя для конкретной фирмы, товара или услуги.

Знакомство потребителя с товаром или услугой начинается с имени

бренда. Кроме того, имя, в отличие от логотипа компании и слогана,

© ДАНИЛЬЧУК Мария Витальевна. — Danilchuk Maria. Российский государственный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical Uni- versity of Russia, Russia. E-mail: [maryanadan@yandex.ru](mailto:maryanadan@yandex.ru)

не меняется с течением времени (подобные случаи встречаются редко). С именем потребитель сталкивается во всех каналах коммуникации: на упаковке или вывеске, в СМИ, в аудио- и видеорекламе, кроме того, он произносит название при покупке и т. д. [Скок 2011]. Именно поэтому разработку названий для товаров и услуг выделили в отдельный вид дея- тельности — нейминг.

Нейминг использует лексику конкретного языка, но, во-первых, запас лексики при всем его объеме не безграничен, во-вторых, не вся лексика соответствует задачам номинации бренда, в-третьих, все чаще проводятся судебные разбирательства по защите прав интеллектуальной собствен- ности с проведением нейминговой экспертизы. Нейминговая экспер- тиза — это термин, образовавшийся на стыке лингвистики, ономастики и юриспруденции, обозначающий процедуру, необходимую для защиты прав интеллектуальной собственности на торговую марку [Ремчукова 2015]. Решением задач нейминга может стать создание неологизмов для новых российских брендов или для адаптации названия иностранного бренда. Если предположить, что при восприятии неологизмов и адапти- рованных иностранных слов первичным является именно звуковой ряд, то фоносемантический анализ названия является одной из важнейших составляющих нейминга.

Для анализа типологии имен брендов методом сплошной выборки

был составлен пилотный корпус брендов, выявивший следующие типы номинаций («бренд» и «торговая марка» в данной статье понимаются как синонимы):

1. Названия, содержащие имена собственные, фамилии производителей

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| Lipton | Чай | Английский бренд. Заимствовано  для российского рынка без изменений |
| Mcdonalds | Сеть ресторанов быстро- го питания | Американский бренд. Заимствовано для российского рынка без изменений |
| Коркунов | Шоколад | Российский бренд |
| Тинькофф | Банк | Российский бренд |

1. Названия-аббревиатуры

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| BMW | Автомобиль | Немецкий бренд. Заимствовано  для российского рынка без изменений |
| KFC | Сеть ресторанов быстро- го питания | Американский бренд. Заимствовано для российского рынка без изменений |
| КамАЗ | Автомобиль | Российский бренд |
| МТС | Мобильная связь | Российский бренд |

1. Названия, созданные с использованием фоносемантических при- емов (аллитерация, ассонанс, звукоподражание, рифма и т. д.)

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| Sсhweppes | Напиток | Швейцарский бренд. Заимствовано для российского рынка без изменений. Название является фамилией основателя, но акцент в имени бренда сделан именно на использование звукоподражания |
| Coca-cola | Напиток | Американский бренд. Заимствовано для российского рынка без изменений |
| Yo-Yo | Игрушка | Американский бренд. Заимствовано для российского рынка без изменений |
| Вимм-Билль-Данн | Продукты питания | Российский бренд, но с 2011 г. компания поглощена американской компанией PepsiCo |
| Кис-кис | Ирис (конфеты) | Российский бренд |

1. Названия, созданные с использованием произвольных, реальных слов из существующего запаса лексики

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| Dove | Шоколад | Американский бренд. Заимствовано для российского рынка без изменений |
| Apple | Техника | Американский бренд. Заимствовано для российского рынка без изменений |
| Добрый | Напиток | Российский бренд |
| Swell | Напиток | Российский бренд |

1. Названия-неологизмы

|  |  |  |
| --- | --- | --- |
| Имя | Товарная категория | **Комментарий** |
| Nissan | Автомобиль | Японский бренд. Заимствовано  для российского рынка без изменений.  Слово “Nissan” было собрано  из первых букв двух слов названия компании “Nixon Sanyo” |
| Fanta | Напиток | Немецкий бренд. Заимствовано для российского рынка без изменений. Сокращение от “Fantacstic” |
| Microsoft | Программное обеспе- чение | Американский бренд. Заимствовано для российского рынка без изменений.  Сложение слов “Microcomputer Software” |
| Canon | Фототехника | Японский бренд. Заимствовано для российского рынка без изменений. Скорректированное имя буддийского бога милосердия Кванона |
| Yandex | Поисковая система | Российский бренд |
| Шармэль | Кондитерские изделия | Российский бренд |

1. Названия с использованием окончаний off, of, офф, оф

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| Davidoff | Сигареты, кофе | Швейцарский бренд. Заимствовано для российского рынка без изменений |
| Таксовичкоф | Такси | Российский бренд |
| Чайникофф | Сеть кофеен | Российский бренд |
| Тинькофф | Банк | Российский бренд |
| Nemiroff | Алкогольная продукция | Российский бренд |

1. Названия с использованием интернет-символов (буква i, названия доменов, хештеги)

|  |  |  |
| --- | --- | --- |
| **Имя** | **Товарная категория** | **Комментарий** |
| iдеЯ! | Интернет-магазин бытовой техники | Российский бренд |
| iТехника  и iАксессуары | Ремонт техники | Российский бренд |
| #Кофейня № 1 | Ресторан | Российский бренд |
| Мусора. больше.net | Вывоз мусора, экологи- ческие проекты | Российский бренд. Название домена является частью фразы |
| Кенгу.ru | Агентство интернет- маркетинга | Российский бренд. Название домена является частью слова |

С точки зрения маркетинговой фоносемантики интерес представляет не только третий тип номинации. Так, например, в неологизмах звуковой ряд также несет информацию о слове. Степень воздействия звукового ряда в разных типах номинации и эффективность передачи ассоциаций, соответствующих рекламируемому товару или услуге, можно определить с помощью экспериментальной оценки номинаций респондентами. Если результаты дальнейших исследований покажут, что звуковой ряд является основным или даже единственным источником информации для опреде- ленного рода номинаций, процедура фоносемантического анализа будет иметь решающее значение при создании нового имени, которое изна- чально имеет нулевое маркетинговое наполнение [Вегенер 2012]. Слово, изначально не имеющее маркетингового наполнения, уникально и имеет больше возможностей войти в речь, приобрести определенный вес в языке, при этом сделать бренд легко узнаваемым [Данильчук 2016].

Таким образом, необходимо выявить экспериментально, что является основным источником информации для потребителя в каждом типе номи- нации: семантика, фоносемантика, графика и т. д. Фоносемантический анализ важен тогда, когда звуковой ряд является основным источником информации. Процедура фоносемантического анализа имени бренда нуж- дается в детальном описании для последующего использования в рамках

маркетинговой фоносемантики как для анализа существующих номина- ций, так и для того, чтобы заложить нужные разработчику ассоциации в новое имя на этапе его создания.

**Литература**

* 1. *Вегенер Ю.С.* (2012), Нейминг в системе формирования и продвижения бренда // Омский научный вестник, № 1 (105), с. 260–262.
  2. *Данильчук М.В*. (2016), Фоносемантический анализ при решении вопросов номинации в маркетинге: постановка задачи // Вестник современной науки, № 2, с. 12–14.
  3. *Ремчукова Е.Н., Соколова Т.П., Махиянова Л.Р.* (2015), Проблемы ней- минга: новые русские урбанонимы в лингвокреативном аспекте // Гуманитарные, социально-экономические и общественные науки, № 9, с. 328–330.
  4. *Скок Д.О.* (2011), Проблемы и роль нейминга в восприятии рекламных сообщений различными целевыми группами // Вестник МГИМО Университета,

№ 5, с. 232–234.

* 1. *Ягафарова Г.Н.* (2014), О факторах, влияющих на процесс номинации в языке // Фундаментальные исследования, № 12 (ч. 11), с. 2505–2508.

**Literature**

1. *Vegener Yu.S*. (2012), Neyming v sisteme formirovaniya i prodvizheniya brenda

// Omskiy nauchnyi vestnik, № 1 (105), s. 260–262.

1. *Danil’chuk M.V*. (2016), Fonosemanticheskiy analiz pri reshenii voprosov nomi- natsii v marketinge: postanovka zadachi // Vestnik sovremennoy nauki, № 2, s. 12–14.
2. *Remchukova E.N., Sokolova T.P., Makhiyanova L.R*. (2015), Problemy neyminga: novye russkie urbanonimy v lingvokreativnom aspekte // Gumanitarnye, sotsial’no- ekonomicheskie i obshchestvennye nauki, № 9, s. 328–330.
3. *Skok D.O*. (2011), Problemy i rol’ neyminga v vospriyatii reklamnykh soob- shcheniy razlichnymi tselevymi gruppami // Vestnik MGIMO Universiteta, № 5, s. 232–234.
4. *Yagafarova G.N*. (2014), O faktorakh, vliyayushchikh na protsess nominatsii v yazyke // Fundamental’nye issledovaniya, № 12 (ch. 11), s. 2505–2508.

***Ф.А. Десятириков, А.В. Белоусов, А.Г. Белоусова, Е.Н. Десятирикова***

**КОМПЬЮТЕРНЫЙ АНАЛИЗ РЕЧЕВОГО ПСИХОИНФОРМАЦИОННОГО ВОЗДЕЙСТВИЯ**

**Аннотация.** Разработана комплексная методика, определяющая наличие несанкционированного речевого воздействия, формирующего заданные пове- денческие реакции слушателя. Реализован программный продукт, позволяющий провести анализ речи на предмет присутствия эмоционального воздействия на слушателя, определить уровень несанкционированного речевого воздействия и ожидаемые решения.

**Ключевые слова.** Психоинформационное воздействие, несанкционированное воздействие, компьютерный анализ речи, объектно-ориентированное програм- мирование.

***F.A. Desyatirikov, A.V. Belousov, A.G. Belousova, E.N. Desyatirikova***

**COMPUTER ANALYSIS OF UNAUTHORIZED SPEECH EXPOSURE**

**Abstract.** A comprehensive methodology that determines the presence of an unauthorized speech inﬂuence, forming the listener speciﬁed behavioral reactions. It implemented a software product that allows to hold a speech analysis for the presence of the emotional impact on the listener, to determine the level of unauthorized speech inﬂuence.

**Keywords.** Psycho-Informational impact, unauthorized exposure, computer speech analysis, object-oriented programming.

В современном мире человек находится под воздействием огромного потока информации, поэтому актуальным является анализ речевых сигна- лов на предмет присутствия несанкционированного (скрытого по целям) психологического воздействия, имеющего целью, например, навязывание слушателю определенной точки зрения.

Общая задача работы — компьютерный анализ речи на присутствие несанкционированного воздействия — разделяется на подзадачи (1 ÷ 6):

1. выявление влияния звуков речи на общую содержательность слова;
2. определение корреляции звуков речи и цветов светового спектра;

© ДЕСЯТИРИКОВ Феликс Андреевич. — Desyatirikov Felix. МБОУ «Гимназия им. Н.Г. Басова при Воронежском государственном университете». — Basov Gymnasia (Voronezh State University), Russia. E-mail: [science2000@ya.ru](mailto:science2000@ya.ru)

© БЕЛОУСОВ Алексей Вадимович. — Belousov Alexey. Воронежский опорный универ- ситет (ВГТУ и ВГАСУ). — Voronezh State Technical University, Russia.

© БЕЛОУСОВА Алла Генриховна. — Belousova Alla. МБОУ «Гимназия им. Н.Г. Басова при Воронежском государственном университете». — Basov Gymnasia (Voronezh State University), Russia.

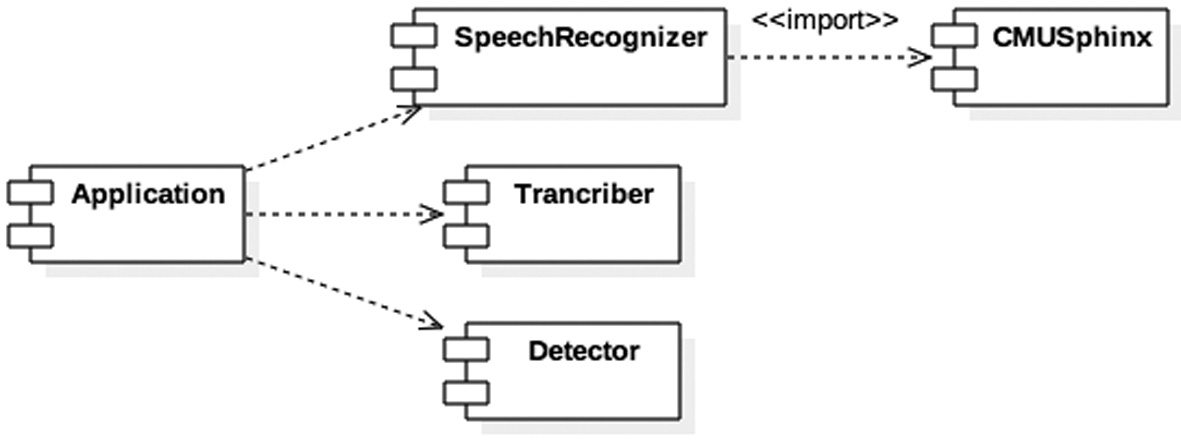
© ДЕСЯТИРИКОВА Елена Николаевна. — Desyatirikova Elena. Воронежский государ- ственный университет. — Voronezh State University, Russia.

1. изучение воздействия цвета на эмоциональное состояние человека;
2. реализация преобразования речи в текст, а также ее транскриби- рование;
3. разработка алгоритма для расчета частотности звуков в речи и выделения основной палитры цветов речи по звукам, отклоняющимся от среднестатистической частоты употребления;
4. прогнозирование эмоциональных реакций человека на речевое воздействие.

Зависимость между звуком и цветом изучается давно и рассматрива- ется в контексте фонемно-цветовой синестезии [Журавлев 1991, Пугачева и др. 2008]. Цвет вызывает у человека психофизиологическую реакцию, оказывающую влияние на его разум и тело, поэтому, установив звуко- цветовые зависимости, можно прогнозировать поведение слушателя. Ранее [Белоусова и др. 2016] доказано, что звуки речи заключают в себе определенную информативную содержательность, формирующую эмо- циональное состояние слушателя.

**Описание технической реализации работы**

Для программной реализации подзадач (1÷ 6) использовались:

* CMUSphinx — программное обеспечение для распознавания речи, построенное на основе СММ и n-граммных моделей;
* Java — объектно-ориентированный язык программирования (выбор языка обусловлен удобной интеграцией с CMUSphinx, а также широким набором стандартных коллекций для хранения данных);
* IntelliJ IDEA — интегрированная среда разработки, включающая в себя текстовый редактор, компилятор, отладчик, средства интегра- ции с системами контроля версий и другие полезные при разработке ПО инструменты;
* Apache Maven — фреймворк для автоматизации сборки проектов на основе описания их структуры в файлах на языке Project Object Model, являющемся подмножеством XML.

*Рис. 1.* Диаграмма компонентов приложения (нотация UML)

Нами разработан программный продукт (***ПП***) Maven-приложение (рис. 1), состоящее из нескольких взаимосвязанных частей:

* Application — инициализирующая часть программы, посредством

которой осуществляется доступ к остальным компонентам;

* Detector — часть приложения, обрабатывающая транскрибирован- ную строку для выделения отклоняющихся от нормы звуков и их цветов;
* SpeechRecognizer — часть программы, отвечающая за преобразо- вание речи в строку текста (обращается к библиотекам CMUSphinx для подключения необходимого функционала обработки речи);
* Transcriber — данный компонент занимается транскрибированием текстовой строки, полученной от компонента распознавания речи.

На вход приложения поступает речь, которая распознается классом SpeechRecognizer. Полученный текст используется классом Transcriber, на выходе которого получается последовательность звуков. В классе Detector вычисляются частотности звуков и сравниваются со среднеста- тистическими частотностями, затем выделяется цветовая гамма обраба- тываемой речи.

### SpeechRecognizer (распознавание речи)

Класс SpeechRecognizer использует библиотеки CMUSphinx для распознавания речи. При инициализации методом Application объекта класса SpeechRecognizer начинается построение необходимых библиотек, расположенных в папке libraries (содержит файлы акустической модели языка для правильного распознавания фонем, список примерно из 500 тыс. слов русского языка и их транскрипции, лингвистические модели языка). После сборки всех библиотек и инициализации объекта класса SpeechRecognizer выполняется метод recognize(). А именно запускается процесс распознавания речи, при этом в консоль печатается уведомление о начале записи. Пока CMUSphinx будет улавливать речевые сигналы, запись будет продолжаться. В ином случае сработает команда на преоб- разование записанной речи в текст, и за ней — команда на освобождение памяти от записанной речи и результата преобразования. В консоли появится сообщение об успешной записи речи. Далее преобразованный текст выводится в консоль и возвращается классу Application для после-

дующего использования классом Transcriber.

### Transcriber (транскрибирование речи)

Данная часть приложения транскрибирует текстовую строку, передан- ную в конструктор класса. Например, при входной строке: «проверка рабо- тоспособности данного компонента», результат получится следующим:

«p r o vj e r k a r a b o t o s p o s o b n o s tj i d a n n o g o k o m p o nj e n t a».

В работе класса Transcriber учтены данные [Бондарко 2006].

После исполнения вышеперечисленных функций класс Transcriber выводит результат работы в консоль и возвращает его в класс Application для последующего использования классом Detector.

### Detector (определение цветовой гаммы речи по частотностям звуков, отклоняющихся от нормы)

В конструкторе класса Detector определены зависимости между зву- ками русского языка и их среднестатистической частотой употребления в речи. Используемые в приложении частотности звуков основаны на дан- ных, полученных в работах [Журавлев 1991, Ляшевская и др. 2009].

Основной метод detect() данного класса приведен в листинге 1.

*Листинг 1 — Метод detect() класса Detector*

void detect() { countSounds();

System.out.println(“Общее количество звуков в строке = « + allSound- sCount + «\n»);

System.out.println(«Количество повторений звуков: \n» + soundsCounts

+ «\n»);

calculateSoundsFrequencies();

System.out.println(«Вычисленные частотности звуков (кол-во повторе- ний/общее количество): \n» + soundsFrequencies + «\n»);

System.out.println(«Среднестатистические частотности звуков: \n» + averageSoundsFrequencies + «\n»);

compareSoundsFrequencies();

System.out.println(«Разницы частотностей (вычисленная /статистиче- ская):»);

System.out.println(«Выше нормы « + mostDifferentSounds); System.out.println(«Ниже нормы « + lessDifferentSounds + «\n»);

deﬁneSoundsColors();

System.out.println(«Выделяющиеся в речи цвета: \n» + speechColors);

}

В процедуре countSounds происходит разбиение строки, поступившей на вход конструктора класса Detector, вычисление общего числа звуков в строке и подсчет количества повторений каждого звука. После выпол- нения данного метода в консоль выводятся полученные данные.

Затем в методе calculateSoundsFrequencies() для звуков высчитывается частотность путем деления количества повторений каждого элемента на общее количество звуков в строке. Полученные результаты расчетов выводятся в консоль.

Следующий этап — сравнение вычисленных частотностей со средне- статистическими в процедуре compareSoundsFrequencies(). Для этого зна-

чения по каждому звуку делятся друг на друга, и, если отношение оказыва- ется больше 2, то частота употребления данного звука превышена в 2 раза, в таком случае звук и его вычисленная разница частотностей добавляется в объект mostDifferentSounds класса HashMap. Если отношение меньше 0,5, то частотность конкретного звука слишком мала, и звук с разницей его частотностей заносится в объект lessDifferentSounds класса HashMap. После этого объекты сортируются методом sortByValues.

По отсортированным в порядке убывания разницам частотностей зву- ков в методе deﬁneSoundsColors() определяются соответствующие звукам цвета, которые выносятся в выходную строку Maven-приложения.

### Экспериментальное подтверждение работы приложения

Проведены эксперименты с текстами, потенциально содержащими скрытое воздействие на слушателя. В частности, проанализирован реклам- ный слоган «*Наша компания рекомендует вам приобрести этот автомо- биль для принятия участия в розыгрыше профессиональной автомойки*». В стандартном наборе призывающих к покупке слов программа нашла отклоняющиеся в 4.4 и 2.7 раза, соответственно звуки [j] (Й) и [rj] (мягкий Р), окрашивающие речь в яркие цвета красного, оранжевого и коричне- вого. Эти цвета и попали в итоговую строку (рис. 2).

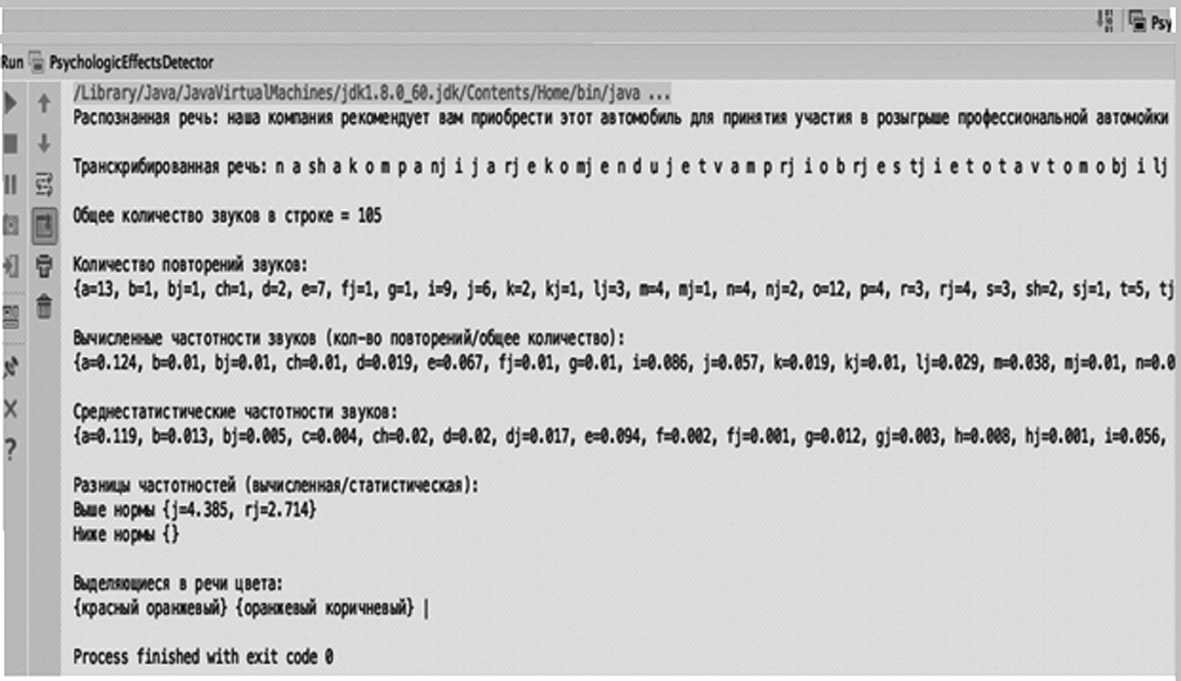
Для интерпретации результатов использованы данные [Купер 2001]. Оранжевого цвета в речи больше всего, он вызывает у слушателя теплые чувства и благотворный эмоциональный фон. Красный привлекает вни- мание покупателя, а также побуждает его к активности. Коричневый цвет способствует чувству доверия и надежности. В итоге получается, что звуки в, казалось бы, обычных для продавца словах побуждают клиента к покупке.

### Заключение

При речевом воздействии можно оказывать несанкционированное воз- действие на слушателя. Результатом его является формирование нужного эмоционального состояния, поведенческих реакций и решений слушателя. Компьютерная реализация анализа речи на присутствие скрытого по целям воздействия на слушателя включает в себя: распознавание речи; транскрибирование речи; алгоритм выделения цветовой гаммы речи по частотностям звуков, отклоняющихся от нормы употребления; подпро- грамму вывода о наличии несанкционированного воздействия.

Разработанный ***ПП*** позволяет провести анализ любого цвето-звукового воздействия на человека (реклама на естественном языке, речевое и цве- товое сопровождение компьютерных игр, речь участника переговорного процесса и др.) и выявить несанкционированное психо-информационное воздействие.

236



*Рис. 2*. Работа приложения с речью

**Литература**

1. *Белоусова А.Г., Десятириков Ф.А., Каширских С.Г., Назаренко М.О.* (2016), Системный анализ речевого несанкционированного воздействия / Системный анализ в проектировании и управлении: сб. науч. тр. XX МНПК. СПб.: Изд-во Политехн. ун-та, ч. 1, с. 470–474.
2. *Бондарко Л.В.* (2006), Фонетика современного русского языка. СПб.: Изд-во СПбГУ, 386 с.
3. *Пугачева Т.Н., Дымшиц М.Н., Кулакова С.В*. (2008), Звуко-цветовая синесте- зия и методы ее исследования: <http://www.vaal.ru/show.php?id=114>(дата обращения 10.06.2016).
4. *Журавлев А.П*. (1991), Звук и смысл. М.: Просвещение, 160 с.
5. *Купер М., Мэтьюз А*. (2001), Язык цвета. М.: ЭКСМО-Пресс, 144 с.
6. *Ляшевская О.Н., Шаров С.А.* (2009), Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 1090 с.

**Literature**

1. *Belousova A.G., Desyatirikov F.A., Kashirskikh S.G., Nazarenko M.O.* (2016), Sistemnyi analiz rechevogo nesanktsionirovannogo vozdeystviya / Sistemnyi analiz v proektirovanii i upravlenii: sb. nauch. tr. XX MNPK. SPb.: Izd-vo Politekhn. un-ta, ch. 1, s. 470–474.
2. *Bondarko L.V*. (2006), Fonetika sovremennogo russkogo yazyka. SPb.: Izd-vo SPbGU, 386 s.
3. *Pugacheva T.N., Dymshits M.N., Kulakova S.V*. (2008), Zvuko-tsvetovaya sinesteziya i metody eyo issledovaniya: <http://www.vaal.ru/show.php?id=114>(data obrashcheniya 10.06.2016).
4. *Zhuravlev A.P*. (1991), Zvuk i smysl. M.: Prosveshchenie, 160 s.
5. *Kuper M., Met’yuz A*. (2001), Yazyk tsveta. M.: EKSMO-Press, 144 s.
6. *Lyashevskaya O.N., Sharov S.A*. (2009), Chastotnyi slovar’ sovremennogo russkogo yazyka (na materialakh Natsional’nogo korpusa russkogo yazyka). M.: Azbukovnik, 1090 s.

***Е.Ю. Дубинина***

**РАЗРАБОТКА МОДЕЛИ**

**АВТОМАТИЧЕСКОГО ВЫДЕЛЕНИЯ ИНФОРМАТИВНЫХ ФРАГМЕНТОВ ТЕКСТА**

**Аннотация.** В статье рассматриваются основные вопросы, возникающие в процессе автоматического реферирования научного текста, а также базовые принципы, лежащие в основе операций по выделению его информативных фраг- ментов. Описывается эксперимент, связанный с сопоставительным анализом авторских рефератов и соответствующих им научных статей. По результатам анализа устанавливаются базовые принципы автоматического реферирования, которые служат основой моделирования этого процесса.

**Ключевые слова.** Автоматическое реферирование, авторский реферат, имен- ные группы, информативные фрагменты.

***E.U. Dubinina***

**DEVELOPMENT OF THE MODEL FOR AUTOMATIC EXTRACTION OF INFORMATIVE TEXT FRAGMENTS**

**Abstract.** The paper considers the basic problems arising during automatic abstract- ing of the scientiﬁc text, as well as the basic principles for extraction of informative text fragments. Therefore the experiment connected with the comparitive analysis of the author’s abstracts and corresponding scientiﬁc articles is described. The analysis results determine the basic principles of automatic abstracting to form a basis of this process modelling are established.

**Keywords.** Automatic summarization, author’s abstract, noun phrases, informa- tive fragments.

Постоянное возрастание объемов текстовой информации вызывает необходимость создания и совершенствования систем автоматической переработки текста. Среди систем автоматической переработки текста осо- бый интерес вызывают системы автоматического реферирования. Целью таких систем является создание вторичного текста (реферата), который объективно передает содержание исходного текста. В настоящее время рефераты, созданные с помощью компьютерной системы, формируются, как правило, из предложений исходного текста, «взвешенных по важно- сти» [Тарасов 2010]. Затем в реферат отбирается определенное количество наиболее информативных предложений.

При разработке систем автоматического реферирования наиболее важной является задача выделения информативных фрагментов текста.

© ДУБИНИНА Екатерина Юрьевна. — Dubinina Ekaterina. Санкт-Петербургский государ- ственный университет аэрокосмического приборостроения, Россия. — St. Petersburg State University of Aerospace Instrumentation, Russia. E-mail: [eka609@yandex.ru](mailto:eka609@yandex.ru)

Для того чтобы решить эту задачу, необходимо выяснить, каким образом распределяется в тексте наиболее важная информация. Если предполо- жить, что основная информация распределяется в тексте не равномерно, а находится в каких-либо его определенных фрагментах, то именно эти фрагменты можно использовать при создании автоматического реферата. Получить данные о закономерностях распределения основной инфор- мации в тексте можно, исследовав его структуру, в частности структуру научной статьи.

Для выявления наиболее информативных фрагментов текста был проведен сопоставительный анализ текстов авторских рефератов и соот- ветствующих им научных статей. В качестве материала исследования были использованы научные статьи, взятые из сборников научной конференции и научных журналов по судостроению. В результате предварительного анализа было отобрано 100 статей. Все статьи принадлежат англоязычным авторам и обладают заранее заданной жесткой структурой. Обязательными компонентами в них являются: заглавие, авторский реферат и сама статья, состоящая из введения, основной части и заключения.

Первоначальная задача при разработке модели выделения информа- тивных фрагментов текста состояла в создании словаря антипризнаков. Целью такого словаря является удаление неинформативной лексики [Беляева 2001]. В словарь антипризнаков входит служебная лексика, а также общая и частично общенаучная лексика (см. пример в табл. 1).

*Таблица 1.* Частотный словарь общенаучной лексики (фрагмент)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ранг** | **Частота** | **Словоформа** | **Ранг** | **Частота** | **Словоформа** |
| 1 | 1346 | used | 18 | 256 | increased |
| 2 | 1188 | using | 19 | 254 | performed |
| 3 | 449 | during | 20 | 251 | available |
| 4 | 447 | new | 21 | 244 | points |
| 5 | 441 | considered | 22 | 183 | located |
| 6 | 434 | presented | 23 | 182 | previous |
| 7 | 433 | found | 24 | 182 | related |
| 8 | 425 | use | 25 | 182 | smaller |
| 9 | 423 | well | 26 | 180 | known |
| 10 | 395 | proposed | 27 | 176 | according |

Помимо словаря антипризнаков, был создан словарь типизированных конструкций. В этот словарь входили словосочетания из авторских рефе- ратов и статей, которые являются показателями важности предложения. Такие конструкции, расположенные, как правило, в начальной позиции предложения, передают информацию о целях и задачах проведенного исследования, его актуальности, а также его результатах (например: *this study aims; the current study has emphasized; the objective of this research*). Подобные типизированные конструкции были выделены и занесены в специальный словарь.

Следующая задача состояла в выделении информативных лексических единиц (ЛЕ), содержащихся в заглавиях и авторских рефератах. Для их выделения была разработана специальная процедура, состоящая из трех этапов. На первом этапе к каждой статье составлялся список всех ЛЕ, входящих в заглавие и реферат. На втором этапе эти списки сравнива- лись со словарем антипризнаков. В результате этого сравнения удалялись неинформативные ЛЕ. На третьем этапе из оставшихся ЛЕ выделялись простые именные группы (ИГ), которые считались ключевыми и являлись объектом дальнейшего анализа.

ИГ, выделенные из заглавия и авторского реферата, исследовались относительно их употребления в тексте. Подсчитывалась частота встречае- мости каждой ключевой ЛЕ и степень ее распространенности. Под степенью распространенности понимается количество вхождений каждой ключевой ЛЕ, выделенной из заглавия и авторского реферата, в структурные разделы статьи, т. е. в такие ее компоненты, как введение, основная часть и выводы. На основании полученных данных для каждой статьи были составлены таблицы ключевых ЛЕ (см. табл. 2). Также была составлена сводная таблица, отражающая распределение ИГ разной длины в массиве текстов.

*Таблица 2*. Ключевые лексические единицы, выделенные при сопоставлении реферата, заглавия и статьи (фрагмент)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Лексические единицы | Частота в разделах | | | | | Суммарная частота | Распространенность |
| Заглавие | Реферат | Введение | Основная часть | Выводы |
| fatigue life |  | 1 | 3 | 20 | 3 | 27 | 4 |
| ﬁnite element analysis |  | 1 | 1 | 3 | 1 | 6 | 4 |
| ﬂuid-structure interaction |  | 1 | 1 | 1 |  | 3 | 3 |
| full-scale measurements |  | 1 | 4 | 1 |  | 6 | 3 |

Рассмотрение таблиц показало достаточно большое разнообразие входящих в них ЛЕ. В анализируемом массиве содержится 1718 разных ИГ. Из них наибольшую частоту имеют двухкомпонентные ИГ, сводная частота которых равна 708. Следует отметить, что, если однокомпонент- ные ИГ представляют собой, как правило, ЛЕ с общим значением (напри- мер: *data, function, process*), то двух- и трехкомпонентные ИГ являются в основном терминологическими словосочетаниями (*actuator dynamics, turbulent ﬂow, controllable transom ﬂap*). В работе принято, что одно-, двух- и трехкомпонентные ИГ являются малокомпонентными, а ИГ, содержащие 4 и более элементов, — многокомпонентными.

Кроме того, в результате анализа таблиц выяснилось, что в заглавиях и авторских рефератах используется большое количество многокомпо- нентных ИГ. В самих статьях эти многокомпонентные группы практиче- ски не используются: как правило, они разбиваются на двух- или трех- компонентные коллокации. Исходя из этого, был сделан вывод о том, что именно двух- и трехкомпонентные ИГ передают важную информацию о содержании текста.

На следующем этапе необходимо исследовать закономерности распре- деления ключевых ЛЕ в отдельных компонентах статьи. Для того чтобы выяснить, распределяются ли ключевые ЛЕ равномерно по всему тексту или же они расположены в каких-либо его определенных фрагментах, была использована теория «сильных» позиций текста. Согласно этой теории наиболее важная информация расположена в фиксированных фраг- ментах текста, например, в заглавии, в первом и последнем предложении, во вводной и заключительной части [Москальчук 2003]. Соответственно, было сделано предположение, что такой компонент статьи, как основная часть, содержит менее значимую информацию. Исходя из этого предпо- ложения, ИГ, выделенные из заглавия и авторского реферата каждой ста- тьи, изучались относительно их использования в основной части текста. Изучение количества разных ИГ, выделенных в заглавии, авторском рефе- рате и основной части, показало, что большинство авторов используют ключевые ЛЕ из заглавия и реферата во введении и выводах, лишь незна- чительная часть ЛЕ, используемых в авторских рефератах, содержится в основной части статьи.

На основании этого был сделан вывод о том, что в основной части

статьи содержится менее значимая информация и, как следствие, этот ком- понент статьи можно не учитывать при выделении наиболее информатив- ных фрагментов текста. Исходя из полученных данных были предложены принципы реферирования, которые затем использовались при разработке модели выделения информативных фрагментов текста. К таким принци- пам относятся следующие: полное исключение из процесса компрессии такого компонента статьи, как основная часть; оценка информативности

предложений по частоте и степени распространенности ключевых ЛЕ, по количественному составу ИГ, а также по наличию в них типизированных конструкций.

Именно эти принципы были использованы при разработке модели выделения информативных фрагментов текста. Базовые компоненты разработанной модели включают базу знаний, которая содержит словарь антипризнаков, список типизированных конструкций, а также формулы для подсчета весов предложений. Предлагаемая модель реализует воз- можность участия пользователя на этапе обучения системы, выбора модели обработки текстов и организована как набор модулей. Задачей этих модулей является определение «весовых» характеристик каждого предложения на основе наличия в предложении ключевых ЛЕ, типизиро- ванных конструкций, малокомпонентных ИГ в совокупности с другими дополнительными весами.

Итак, разработанная модель позволяет выявить наиболее информа- тивные фрагменты исходного текста, тем самым сократив время его обра- ботки, что, в свою очередь, дает возможность пользователю определить дальнейшие режимы работы с документами массива.

**Литература**

* 1. *Беляева Л.Н.* (2001), Лингвистические автоматы в современных информа- ционных технологиях. СПб.
  2. *Москальчук Г.Г.* (2003), Структура текста как синергетический процесс. М.
  3. *Тарасов С.Д.* (2010), Метод тематического ранжирования в задачах авто- матического сводного реферирования, Искусственный интеллект и принятие решений, № 2, с. 36–41.

**Literature**

1. *Belyaeva L.N.* (2001), Lingvisticheskie avtomaty v sovremennykh informatsion- nykh tekhnologiyakh. SPb.
2. *Moskal’chuk G.G*. (2003), Struktura teksta kak sinergeticheskiy protsess. M.
3. *Tarasov S.D*. (2010), Metod tematicheskogo ranzhirovaniya v zadachakh avto- maticheskogo svodnogo referirovaniya, Iskusstvennyi intellekt i prinyatie resheniy,

№ 2, s. 36–41.

***А.А. Мартиросян***

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ КАК МЕТОД ИССЛЕДОВАНИЯ ТЕКСТА**

**Аннотация.** В докладе представлены результаты исследования художе- ственного текста в плане соотношения локальных тем, в совокупности пере- дающих его идею; вводится понятие кумулятивного показателя распределения концептуальной информации. При анализе художественных текстов используется метод математического моделирования, с помощью которого свойства языковых явлений концентрированно отражаются в конкретных математических моделях, показанных в таблицах. Использование математики в лингвистике ориентировано на повышение надежности и точности методов лингвистических исследований.

**Ключевые слова.** Тема, математическое моделирование, кумулятивный показатель, концептуальная информация.

***A.A. Martirosyan***

**MATHEMATICAL MODELING AS A METHOD OF TEXT STUDY**

**Abstract.** The report presents the results of study of the ﬁctional text in terms of correlation of the local themes rendering as a whole its idea; the report introduces the notion of cumulative index of distribution of conceptual information. When analyzing the ﬁctional text the method of mathematical modeling is used enabling to represent lan- guage features in particular mathematical models shown in tables. Using mathematics in linguistics aims to enhance reliability and accuracy of the methods of linguistic studies. **Key words.** Theme, mathematical modeling, cumulative index, conceptual infor-

mation.

Современная лингвистика стремится не просто к точным наблюде- ниям, а к созданию формальных моделей. Постепенно происходит мето- дологический переход от субъективной интроспекции к объективным корпусным методам исследования. Корпусная лингвистика занимается разработкой принципов построения и использования корпусов текстов, так как текст является одним из основных источников языкового материала, необходимого для лингвистических исследований. Лингвистические кор- пусы, размеченные по разным лингвистическим параметрам, стали созда- ваться с целью работы с большими объемами текстов для извлечения из них нужной информации. Для больших объемов исследования материала потребовалось применение новых методов анализа, включающих метод математического моделирования, квантитативный метод.

© МАРТИРОСЯН Анна Арменовна. — Martirosyan Anna. НИУ Высшая школа эконо- мики, Россия. — National Research University Higher School of Economics, Russia. E-mail: [annamartirossian@mail.ru](mailto:annamartirossian@mail.ru)

Метод математического моделирования основан на использовании формально-абстрактных структур и может быть приложим к различным объектам действительности и познания, каковым, по нашему мнению, является и художественный текст. Само математическое описание языка восходит к представлению Ф. де Соссюра о языке как о механизме, про- являющемся в речевой деятельности, результатом которой являются последовательности речевых единиц, подчиняющиеся определенным закономерностям и допускающие математическое описание [Большая советская энциклопедия 1990].

Смена «описательного» вектора на «моделируемый» продиктована установкой на большую достоверность, прогностический характер и квантитативность. Воспроизводящая модель строится по схеме: «линг- вистический объект — гипотеза — аналог лингвистического объекта» [Гумовская 2015: 21]. Мы полагаем, что модель, воспроизводящая языко- вой объект должна не только показать объект схематически, но и помочь сгенерировать новые знания об объекте. Под моделью нами понимается

«такая мысленно представляемая или материально реализованная система, которая, отображая или воспроизводя объект исследования, способна замещать его так, что ее изучение дает нам новую информацию об этом объекте» [Штофф 1966: 19].

Существуют разные способы передачи авторского замысла в художе- ственном произведении. Индивидуально-авторское отношение описы- ваемых в произведении событий передает концептуальная информация, представляющая собой творческое переосмысление этих событий. Рассмотрим, как автор распределяет основные и второстепенные темы в художественном произведении на примере трех романов современных английских писателей: «Одержимый» Майкла Фрейна [Frayn 1999],

«Крушение империй: великая осада» Уильяма Напьера [Napier 2012] и «Вид на Мерси» Рут Хамильтон [Hamilton 2010].

Основная идея романа «Одержимый» заключается в том, что сильная страсть к искусству, как, впрочем, и страсть к чему бы то ни было, может полностью подчинить себе человека. В самом названии произведения отражается характеристика героя, одержимого идеей понять одну из картин Брейгеля и стать ее обладателем.

«Крушение империй: великая осада» — исторический роман, осно- ванный на реальных событиях XVI в. — осаде турками острова Мальты, который принадлежал рыцарям ордена госпитальеров. Главная мысль произведения заключается в том, что сила духа людей может преодолеть непреодолимые обстоятельства: в данном случае несколько сотен рыцарей смогли отразить нападение многотысячной армии.

«Вид на Мерси» повествует о судьбе женщины, которая, вырастив троих детей, поняла, что не может больше жить с мужем, который на про-

тяжении всей их совместной жизни обманывал ее. Основная мысль романа состоит в том, что в любом возрасте судьбу можно взять в свои руки и изменить, главное — сделать решительный шаг, перестать подавлять свои эмоции и не бояться перемен.

В исследовании особенностей художественного произведения, большое значение имеет для нас анализ тем. В романе «Одержимый» нами выявлены 42 темы, в совокупности передающие идею произ- ведения, в романе «Крушение империй…» — 99 тем**,** а в романе «Вид на Мерси» — 107. Ниже приведены таблицы с подробным анализом тем по каждому произведению, где указываются процент встречаемости каждой темы в тексте произведения и коррелятивный кумулятивный показатель распределения концептуальной информации (КИ).

*Таблица 1A*. Анализ тем в романе «Одержимый»

|  |  |  |  |
| --- | --- | --- | --- |
| **Название темы**  **(общее кол-во 42)** | **Частотность появления темы (кол-во диктем)** | **%**  **встречаемости темы** | **Кумулятивный показатель распределения КИ (%)** |
| Hero’s emotions | 190 | 13,24 | 13,24 |
| Laura | 125 | 8,71 | 21,95 |
| Research | 119 | 8,29 | 30,24 |
| Neighbour | 96 | 6,69 | 36,93 |
| Kate’s disapproval | 82 | 5,71 | 42,65 |
| Hero’s anxiety | 82 | 5,71 | 48,36 |
| Остальные темы | … | … | … |
| Kate’s suggestion | 2 | 0,14 |  |
| Итого | 1435 | 100 | 100,00 |

*Таблица 1B.* Анализ тем в романе «Крушение империй…»

|  |  |  |  |
| --- | --- | --- | --- |
| **Название темы**  **(общее кол-во 99)** | **Частотность появления темы (кол-во диктем)** | **%**  **встречаемости темы** | **Кумулятивный показатель распределения КИ (%)** |
| Nicholas | 379 | 12,37 | 12,37 |
| The fairhead | 150 | 4,90 | 17,27 |
| The knights | 148 | 4,83 | 22,10 |
| La Valette | 132 | 4,31 | 26,41 |
| Fight | 127 | 4,15 | 30,56 |
| Blackbeard | 125 | 4,08 | 34,64 |
| Остальные темы | … | … | … |
| Traitor | 1 | 0,03 |  |
| Итого | 3063 | 100,00 | 100,00 |

*Таблица 1C.* Анализ тем в романе «Вид на Мерси»

|  |  |  |  |
| --- | --- | --- | --- |
| **Название темы**  **(общее кол-во 107)** | **Частотность появления темы (кол-во диктем)** | **%**  **встречаемости темы** | **Кумулятивный показатель распределения КИ (%)** |
| Lucy | 265 | 12,26 | 12,26 |
| Alan | 171 | 7,91 | 20,17 |
| Сhildhood friend | 161 | 7,45 | 27,61 |
| Neighbour | 124 | 5,74 | 33,35 |
| Lizzie | 100 | 4,63 | 37,97 |
| Moira | 88 | 4,07 | 42,04 |
| Остальные темы | … | … | … |
| Ggratitude | 1 | 0,05 |  |
| Итого | 2162 | 100,00 | 100,00 |

В первом столбце каждой таблицы представлены все темы, встре- чающиеся в произведении, в порядке их убывания. Второй столбец указывает частотность повторяемости каждой темы. Так, в романе

«Одержимый» наиболее часто встречается тема “hero’s emotions” — 190 раз, в «Крушении империй…» — тема “Nicholas” 379 раз, а в романе

«Вид на Мерси» — тема “Lucy” 265 раз. В третьем столбце приводится процент встречаемости каждой темы. В четвертом столбце приводятся данные о том, какой процент от использованных тем передает основное содержание произведения, т. е. концептуальную информацию (рядом с каждой темой указывается сумма процентов встречаемости всех тем, расположенных выше данной темы, т. е. употребленных чаще нее). Сумма этих процентов встречаемости обозначена как кумулятивный показатель распределения концептуальной информации.

Так, в романе «Одержимый» (см. табл. 1А) вторая по частоте исполь- зования тема “Laura” имеет процент встречаемости 8,71%, а ее кумулятив- ный показатель равен 21,75% (13,24% + 8,71%), т. е. тема “Laura” и пре- дыдущая тема “hero’s emotions” вместе передают 21,75% концептуальной информации всего произведения. Это есть процент повторяемости в тексте первых двух наиболее часто встречающихся тем. Аналогично в следую- щей строке указана сумма трех наиболее часто встречающихся тем.

Четвертая колонка показывает коррелятивное убывание значимости каждой последующей темы в передаче главной темы произведения. Отрицательная значимость (ее практическое отсутствие) наблюдается в последнем элементе таблицы (т. е. почти 100% смысла передано другими темами), а максимально значимый для этого произведения показатель представлен в первом элементе 13,24% («Одержимый»), 12,37% («Кру- шение империй…») и 12,26% («Вид на Мерси»).

Данные анализа текстов, представленные в таблицах (здесь они пред- ставлены в сокращенном виде), позволяют сделать следующие выводы: чем больше коррелятивный кумулятивный показатель распределения КИ, тем меньшую значимость представляет данная тема для текста; основное содержание произведения 60–80% передается с помощью 15–30% от всех тем этого произведения.

Релевантность и частотность любого лингвистического явления проще проследить на большом объеме текстов, в связи с чем дальнейшие корпусные исследования художественных текстов представляются пер- спективным направлением.

**Литература**

1. Большая советская энциклопедия (1990), URL: https://dic.academic.ru/dic. nsf/bse/107208/Математическая (дата обращения 15.03.2016).
2. *Гумовская Г.Н.* (2015), Математическое моделирование как метод исследо- вания текстовой ритмики // Язык, культура, речевое общение: материалы между- народной научной конференции, посвященной 90-летию профессора М. Я. Блоха. Ч. 1. М.: Факультет иностранных языков МПГУ, с. 19–25.
3. *Штофф В.А.* (1966), Моделирование и философия. М.: Наука, 304 с.
4. *Frayn M.* (1999), Headlong. London: Faber and Faber, 394 p.
5. *Hamilton R.* (2010), Mersey View. Pan Books, London, 405 p.
6. *Napier W.* (2012), Clash of Empires: The Great Siege. Orion Books, London, 398 p.

**Literature**

1. Bol’shaya sovetskaya entsiklopediya (1990), URL: https://dic.academic.ru/dic. nsf/bse/107208/Matematicheskaya (data obrashcheniya 15.03.2016).
2. *Gumovskaya G.N.* (2015), Matematicheskoe modelirovanie kak metod issle- dovaniya tekstovoy ritmiki // Yazyk, kul’tura, rechevoe obshchenie: materialy mezhd- unarodnoy nauchnoy konferentsii, posvyashchennoy 90-letiyu professora M.Ya. Blokha. Ch. 1. M.: Fakul’tet inostrannykh yazykov MPGU, s. 19–25.
3. *Shtoff V.A.* (1966), Modelirovanie i ﬁlosoﬁya. M.: Nauka, 304 s.
4. *Frayn M.* (1999), Headlong. London: Faber and Faber, 394 p.
5. *Hamilton R.* (2010), Mersey View. Pan Books, London, 405 p.
6. *Napier W.* (2012), Clash of Empires: The Great Siege. Orion Books, London, 398 p.

***В.Р. Нымм, К.Р. Пиотровская, Ю.П. Еремеева, И.С. Макогон***

**СТАТИСТИЧЕСКИЕ МЕТОДЫ В ИССЛЕДОВАНИИ ПРОБЛЕМЫ КОМПЬЮТЕРНОГО ОБУЧЕНИЯ ЯЗЫКУ**

**Аннотация.** В контексте исследования проблемы компьютерного обучения иностранным языкам в статье рассматривается использование статистических методов для: а) оценки и сравнения различных алгоритмов управления процессом обучения; б) отбора содержания баз обучающего контента.

**Ключевые слова.** CALL (компьютерное обучение иностранным языкам), алгоритм управления процессом обучения, среднее значение, дисперсия, стан- дартное отклонение, конкордансер, статистический подход.

***V. Nymm, X. Piotrowska, Y. Yeremeeva, I. Makogon***

**STATISTICAL METHODS**

**IN CALL-PROBLEM RESEARCH**

**Abstract**. The paper reviews the using of statistical methods for evaluation and comparison of different learning management algorithms and text-mining selection of learning content in the context of CALL-problem research.

**Keywords.** CALL (computer aided language learning), learning management algo- rithm, mean value, dispersion, standard deviation, concordanser, statistical approach.

Начиная с замечательной работы Л.А. Растригина и М.Х. Эренштейн [Растригин 1988] в качестве целей компьютерного обучения, реализуе- мого в рамках упражнений типа «стимул — реакция», рассматриваются следующие две.

Первая определяет желаемое состояние обучаемого на момент заверше- ния процесса обучения и выражается в форме неравенства *Q* ( *p* )  *q* , где *Q* (*p*) — функционал качества обучения, *p* ( *t* )  ( *p*1 ( *t* ) , *p*2 ( *t* ) , ..., *pn* ( *t* ) ) — вектор параметров состояния обучаемого, *t* — время, а *q* — пороговое значение функционала.

Вторая цель является экстремальной и состоит в минимизации вре- мени достижения желаемого состояния обучаемого.

© НЫММ Волдемар Рихардович. — Nymm Voldemaar. Российский государственный педа- гогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [jetnomm@gmail.com](mailto:jetnomm@gmail.com)

© ПИОТРОВСКАЯ Ксения Раймондовна. — Piotrowska Xenia. Российский государствен- ный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [krp62@mail.ru,](mailto:krp62@mail.ru)

© ЕРЕМЕЕВА Юлия Павловна. — Yeremeeva Yulia. Российский государственный педаго- гический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [pavlof](mailto:pavloffna@bk.ru)[fna@bk.ru](mailto:fna@bk.ru)

© МАКОГОН Ирина Станиславовна. — Makogon Irina. Российский государственный педа- гогический университет им. А.И. Герцена (Россия). — Herzen State Pedagogical University of Russia, Russia. E-mail: [jetnomm@gmail.com](mailto:jetnomm@gmail.com)

Учитывая достаточно опосредованную связь между целевым нера- венством и временем, за которое это условие должно быть достигнуто, в теории управления поиск решения задач, подобных данной, перево- дится на язык стратегий. Другими словами, решение задачи переводится в область поиска эвристического построения, обеспечивающего доста-

точно быструю сходимость функционала *Q* ( *p* ) к пороговому значе-

нию *q*. Возникает естественный вопрос о том, как сравнивать алгоритмы, реализующие различные решения задачи. Один из возможных подходов к решению вопроса состоит в том, чтобы:

а) заменить достаточно субъективное (в рамках компьютерного обу- чения) понятие «минимизации времени достижения желаемого состояния обучаемого» на «минимизацию суммарного числа предъявлений стимулов упражнений до достижения желаемого состояния обучаемого»;

б) с помощью средств сбора данных о ходе обучения, встроенных в программу, которая реализует алгоритм, получить информацию для вычисления статистических характеристик, достаточных для его оценки и сравнения с другими алгоритмами.

В качестве примера реализации такого подхода рассмотрим алгоритм, построенный на базе одноэлементной модели обучения парным ассоциа- циям, которую иногда называют «все или ничего» [Аткинсон и др. 1969: 139]. В основе этой модели лежит утверждение о том, что ассоциация, связывающая стимул и реакцию, не может формироваться частично или постепенно. Предполагается, что, однажды возникнув, ассоциация сохра- няется в течение длительного времени и обучаемый будет давать верный ответ при каждом новом предъявлении ему стимула. Для одноэлементной модели состояние обучаемого в любой момент времени описывается

вектором *p* ( *t* )  ( *p*1 ( *t* ) , *p*2 ( *t* ) , ..., *pn* ( *t* ) ) , каждая координата которого

имеет значение 0, если по i-му элементу ассоциация не сформирована, и значение 1 в противном случае.

Полное описание алгоритма приведено в статье [Нымм, Пиотровская, Быстрая 2015], здесь же мы отметим лишь его существенные характери- стики и результаты тестирования разработанного алгоритма.

Для задачи обучения лексике иностранного языка в качестве критерия завершения обучения по каждому отдельному упражнению рассматрива- ется событие, состоящее в том, что обучаемый в ходе одного из сеансов дает верный ответ с первого же в текущем сеансе предъявления стимула этого упражнения. Если считать, что в день реализуется не более одного сеанса, то это означает, что верный ответ с первого раза был дан не менее, чем через 24 часа после предыдущего предъявления этого стимула. Обу- чение в целом реализуется в ходе последовательных сеансов и продолжа- ется до тех пор, пока все упражнения, представленные в базе обучающей информации, не окажутся помеченными как элементы, по которым обу-

чение завершено, т. е. когда значения всех координат вектора p, вектора параметров состояния обучаемого, станут равными 1.

По завершении процесса обучения, которое отслеживается програм- мой, реализующей алгоритм, она (т. е. программа) допускает обучаемого к контрольному тестированию.

В ходе реализации алгоритма встроенные в программу средства сбора информации о ходе обучения формируют наборы данных, включающие, в частности: а) количество сеансов, которое потребовалось обучаемому до завершения обучения по каждому упражнению; б) число предъявлений стимула упражнения в ходе этих сеансов. Вместе с результатами контроль- ного тестирования эти данные необходимы для оценки эффективности алгоритма и последующих аналитических исследований.

Эффективность алгоритма оценивалась по результатам эксперимента, в котором принимала участие группа из 50 студентов филологического факультета. Материал изучения включал 100 языковых клише (устойчивых словосочетаний, вводных конструкций и т. п.) по общенаучной тематике. Процесс обучения осуществлялся в ходе последовательных сеансов. Интервал времени между сеансами у разных студентов колебался от 24 до 72 часов.

На основе обработки данных были получены следующие результаты. Среднее число предъявлений стимула (по группе в целом), которое потре- бовалось для завершения обучения по одному упражнению, оказалось рав- ным 4,64. Разброс этого показателя для отдельных обучаемых достаточно велик: от 2,72 до 7,81, стандартное отклонение — 1,57.

Среднее число сеансов, в ходе которых эти предъявления реализо- вывались, оказалось достаточно малым — всего 1,91. Это означает, что

«выучивание» правильного ответа в среднем осуществлялось примерно за два сеанса так, что на третьем сеансе обучаемый мог воспроизвести правильный ответ с первого раза. Разброс значений этого показателя также достаточно мал: от 1,51 до 2,37, стандартное отклонение — 0,32.

Непосредственные результаты обучения оценивались по данным контрольного тестирования. Несмотря на то что процедура тестирования была достаточно утомительной (она проводилась по всему материалу обучения, а не по выборке из него, как это делается обычно), результаты обучения оказались более чем оптимистичными. Доля правильных отве- тов для разных обучаемых оказалась в диапазоне от 0,61 до 0,9, среднее значение — 0,779, стандартное отклонение — 0,0832.

Не меньшую роль способны играть статистические методы при формировании информационного обеспечения систем компьютерного обучения. Хотя объективные методы отбора учебных терминологических единиц (например, статистико-вероятностный подход, получивший свое развитие в работах Р.Г. Пиотровского и Г.П. Луна) известны с шести-

десятых годов XX в. [Luhn 1957; Пиотровский и др. 1977], до сих пор отбор терминологических единиц производится подчас субъективно, исключительно с опорой на личный опыт и интуицию преподавателя или исследователя. Особенно остро эта проблема стоит при изучении профессионально ориентированного языка и связана с ускоряющимися темпами информационного обмена и появлением новых специализиро- ванных подъязыков, с лексико-грамматическим составом, не описанным в традиционных учебниках. Сегодня статистико-вероятностный подход переживает определенный ренессанс и в России [Лаздинь 2009], и за рубежом, на что указывают, например, работы группы исследователей из Японии и Новой Зеландии. Эта группа создала ряд открытых и надежных сервисов под названием AntConc, которые активно используются в кор- пусных исследованиях [Nation, Anthony 2016)].

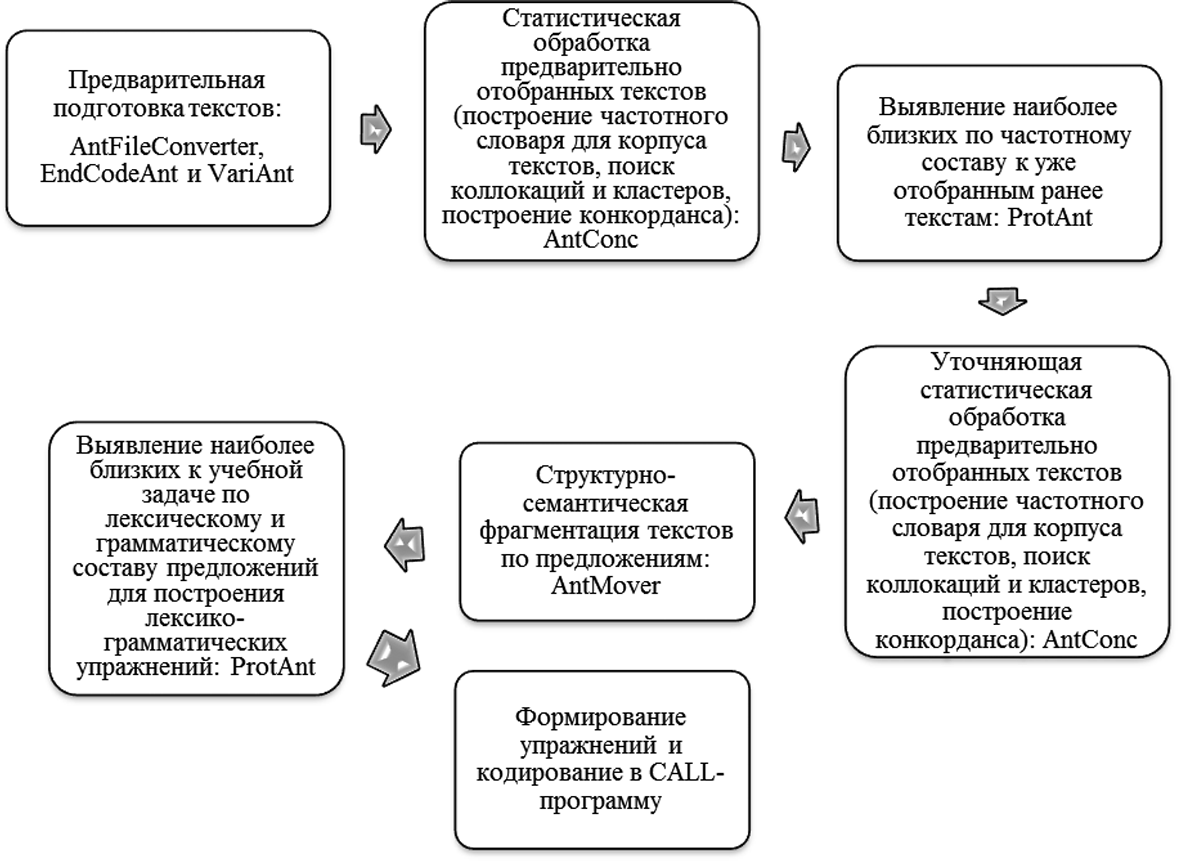
В настоящее время на базе этих сервисов нами разрабатывается тех-

нологическая цепочка операций по квантитативному отбору содержания обучения и созданию информационного обеспечения программы (см. рис. 1). Целевое назначение программ, входящих в семейство программ AntConc отражено в табл.

*Таблица*. Целевое назначение программ группы AntConc

|  |  |  |
| --- | --- | --- |
| **№ п/п** | **Название программы** | **Назначение** |
| 1. | AntConc | Конкордансер (частотный словарь, коллокации, конкор- данс) |
| 2. | AntPConc | Построитель параллельных текстов для последующего анализа текста и построения конкордансов |
| 3. | AntMover | Анализатор семантической структуры текста |
| 4. | AntWordProﬁler | Профилирование словаря и проверки сложности текста |
| 5. | ProtAnt | Инструмент для поиска файла прототипа |
| 6. | TagAnt | Инструмент для теггирования по частям речи |
| 7. | AntFileConverter, EndCodeAnt, VariAnt | Сервисные программы (перекодировка символов, спел- леров и подготовка формата текстов к последующей обработке) |

Предварительно необходимо перевести заранее отобранные тексты в нужный формат, используя программы AntFileConverter, EndCodeAnt и VariAnt. Далее с помощью программы AntConc построить частотный словарь, выделить ключевые слова и выявить устойчивые словосочетания с ключевыми словами. Объем исследуемых текстов далее наращивается путем использования программы ProtAnt по введенным ключевым сло- вам, которые выявлены на предварительном этапе или по эталонному тексту. Эта программа осуществляет отбор текстов, наиболее близких по лексическому составу к поставленной учебной задаче. После этого необ-

ходимо вернуться снова к работе по лингво-статическому исследованию в программе AntConc, но уже расширенного состава учебных текстов, а также к профилированию полученного словаря и разметке текстов. Для удобства формирования учебных упражнений с помощью программы AntMover про- изводится фрагментация и структурно-семантическая разметка текстов по предложениям с сохранением каждого предложения в отдельный файл. Далее программой ProtAnt снова обрабатываются уже файлы отдельных предложе- ний и выявляются те из них, которые по лексическому или грамматическому составу подходят для генерации упражнений в CALL-программу (см. рис.).

*Рис*. Технологическая цепочка формирования обучающего контента

Организованный комплекс выступает поддержкой преподавания таких дисциплин, как «Квантитативная лингвистика и новые информационные технологии», «Компьютерная лингводидактика», «Новые информацион- ные технологии в филологии», «Сетевой инструментарий лингвиста».

**Литература**

* 1. *Аткинсон Р., Бауэр Г., Кротерс Э.* (1969), Введение в математическую теорию обучения. М.: Мир.
  2. *Лаздинь Т.А.* (2009), Основы статистической оптимизации преподавания иностранных языков // Вестник Санкт-Петербургского университета. Сер. 9. Вып. 3. СПб.: Изд. СПбГУ.
  3. *Нымм В.Р., Пиотровская К.Р., Быстрая В.А*. (2015), CALL-технологии: управление процессом обучения языку // Высокотехнологичная образовательная среда: сборник статей Международной научно-практической конференции. СПб.: ООО «Книжный дом», с. 172–179.
  4. *Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А.* (1977), Математическая лингвистика: учеб. пособие для пед. институтов. М.: Высшая школа, 383 с.
  5. *Растригин Л.А.* (1988), Адаптивное обучение с моделью обучаемого / Л.А. Растригин, М.Х. Эренштейн. Рига: Изд. Зинатне, 160 с.
  6. *Luhn H.P.* (1957), A Statistical Approach to Mechanized Encoding and Searching of Literary Information // IBM Journal of Research and Development, Vol. 1, Issue: 4.
  7. *Nation P., Anthony L.* (2016) Measuring vocabulary size. In Handbook of Research in Second Language Teaching and Learning, Vol. III, E. Hinkel (Ed.). New York: Routledge, 2016.

**Literature**

1. *Atkinson R., Bauer G., Kroters E.* (1969), Vvedenie v matematicheskuyu teoriyu obucheniya. M.: Mir.
2. *Lazdin’ T.A.* (2009), Osnovy statisticheskoy optimizatsii prepodavaniya inos- trannykh yazykov // Vestnik Sankt-Peterburgskogo universiteta. Ser. 9. Vyp. 3. SPb.: Izd. SPbGU.
3. *Nymm V.R., Piotrovskaya K.R., Bystraya V.A.* (2015), CALL-tekhnologii: uprav- lenie protsessom obucheniya yazyku // Vysokotekhnologichnaya obrazovatel’naya sreda: sbornik statey Mezhdunarodnoy nauchno-prakticheskoy konferentsii. SPb.: OOO «Knizhnyi dom», s. 172–179.
4. *Piotrovskiy R.G., Bektaev K.B., Piotrovskaya A.A*. (1977), Matematicheskaya lingvistika: ucheb. posobie dlya ped. institutov. M.: Vysshaya shkola, 383 s.
5. *Rastrigin L.A.* (1988), Adaptivnoe obuchenie s model’yu obuchaemogo /

L.A. Rastrigin, M.Kh. Erenshteyn. Riga: Izd. Zinatne, 160 s.

1. *Luhn H.P.* (1957), A Statistical Approach to Mechanized Encoding and Searching of Literary Information // IBM Journal of Research and Development, Vol. 1, Issue: 4.
2. *Nation P., Anthony L.* (2016) Measuring vocabulary size. In Handbook of Research in Second Language Teaching and Learning, Vol. III, E. Hinkel (Ed.). New York: Routledge, 2016.

# ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ В ФИЛОЛОГИИ

***А. С. Балуева***

**НАУЧНЫЙ ТЕКСТ**

**В НОВЫХ УСЛОВИЯХ ФУНКЦИОНИРОВАНИЯ**

**Аннотация.** В статье рассматривается проблема изменения научного тек- ста под влиянием действия функционально-адаптивных процессов в речи и социокультурных процессов. На примере анализа современных научных текстов выделяются определяющие тенденции функционально-адаптивных процессов в речи, связанные с субъектной идентификацией автора текста, изменением пред- ставлений об адресанте и адресате, социальной востребованностью общества в трансляции научных знаний.

**Ключевые слова**: научный текст, функционально-адаптивные процессы в речи, социокультурные факторы.

***A. S. Balueva***

**SCIENTIFIC TEXT IN THE NEW CONDITIONS OF FUNCTIONING**

**Abstract.** The paper covers the issue of scientiﬁc text change under the inﬂuence of the functional adaptive processes in speech and sociocultural processes. Analysing modern scientiﬁc texts, the author reveals the determining tendencies of the functional adaptive processes in speech related to the subject identiﬁcation of the text author, change of ideas about the addresser and addressee, and social demand of society for scientiﬁc knowledge transfer.

**Keywords.** scientiﬁc text, functional adaptive processes in speech, sociocultural factors.

Проблема изменения научного текста тесно связана с системными результатами развития общества и действия социокультурных и гео- политических процессов, отражающих зависимость научного текста от внешнего мира. В первую очередь речь идет о развитии всех видов техни- ческих средств, позволяющих взаимодействовать с широким кругом лиц и распространять научную информацию. При осуществлении трансляции научного знания «акт признания состоит в его использовании в той или иной форме. Степень его использования может быть разной: обращение

БАЛУЕВА Александра Сергеевна. — Balueva Alexandra. Российский государственный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [baspublish@mail.ru](mailto:baspublish@mail.ru)

к нему, запрос; ознакомление с ним <...> производство нового знания на базе использованного...» [Макаров 2003]).

Социальная роль научного текста как части системы познания заключается в отражении действительности и служит для исследования актуальных социокультурных проблем, формирования ценностных и нрав- ственных ориентиров, отражения государственных интересов, необходи- мых в социальной деятельности общества. Язык как часть социальной деятельности и как средство социального взаимодействия отражается в разнообразных дискурсах, интерпретирующих социокультурную реаль- ность. В дискурсе речевое поведение, выбор языковых средств и структур, направлено на реализацию социальной функции текста (информирование, объяснение явлений, действий, их обобщение, воздействие, регулирова- ние, управление, просвещение, воспитание и др.) и установление через научный текст диалога как в межсубъектном взаимодействии с пользова- телем, так и с обществом, поэтому важно иметь общие социокультурные коды, помогающие в восприятии, понимании и интерпретации не только текста, но и реальности.

В научном дискурсе автор текста реализует прежде всего свои комму- никативные задачи, а также задачи, которые обусловлены действием рече- вого стандарта в научной коммуникации. В последнее время наблюдается весьма устойчивая тенденция изменения правил представления результа- тов научного исследования. Речь идет о «включении» в коммуникативный акт читателя, который формирует свои коммуникативно-прагматические задачи. Автор научного текста, а точнее, научной статьи, должен сфор- мировать набор языковых (вербальных и невербальных) средств, направ- ленных на адаптацию текста для адекватного его понимания и оценки значимости научных результатов. В процессе адаптации научного текста для читателя автор использует эффективные приемы, осуществляет функ- циональный выбор речевых структур и языковых средств, тактических и стратегических действий. Анализ изменений речевого стандарта научной коммуникации на фоне осмысления функционально-адаптивных процес- сов, происходящих в современной речи, а также изменений механизмов восприятия, понимания и интерпретации является актуальной задачей для исследователей.

Для определения функционально-адаптивных процессов в речи было осуществлено научное исследование 600 научных статей с использованием коммуникативно-дискурсивного анализа научного текста1.

1 Исследование проводилось в 2016 г. на материале статей научных рецензируемых журна- лов и материалах научных конференций на кафедре образовательных технологий в фило- логии РГПУ им. А.И. Герцена под руководством доктора филологических наук, профес- сора Н.Л. Шубиной.

Установлено, что в настоящее время в научных статьях наблюдается изменение не только структуры, но и функциональное назначение науч- ного текста. В научных статьях наряду с традиционными описаниями результатов эксперимента, подходов к исследуемой проблеме, отражается авторская индивидуальность, проявляющаяся в заранее продуманной модели текста, выборе языковых средств (ср.: статья — дискуссия «Пред- сказуемость — технология интонационного процесса в музыкальной культуре Востока/Запада»; статья — расследование «Загадочная доль- цайна»; статья — побуждение «What A Likable World, или Почему био- логия познания привлекательна для лингвиста»; статья — комментарий

«Метаграфемика в «художественном паспорте» А. Белого». Целью такого научного текста является не только информирование или выявление причинно-следственных связей, но и определение через актуализацию, оценку предмета исследования роли того или иного явления в повсед- невной жизни, желание сделать его общеизвестным и доступным для понимания и оценки.

Автор передает восприятие мира со своей точки зрения, демонстрируя собственную включенность, привлекая для этого возможные средства авторизации (например, *автор предлагает считать, автор отстаивает позицию, считаю правильным называть* и др.) Обстоятельства, мотивы, побуждения автора часто оформляются как цели и могут, к примеру, выра- зить несогласие с мнением окружающих:

«Нередко мы слышим в эфире, обыденных разговорах и менторских наставлениях: “Весь цивилизованный мир...”, “Надо жить как во всем цивилизованном мире...” и пр. Фразы эти словно подчеркивают, что есть, мол, где-то цивилизации, в Северной Америке, например, или Западной Европе, а здесь в России — Российской Федерации нам со всеми особен- ностями своих социально-экономических отношений и системой управ- ления до этой самой цивилизации идти, мол, еще и идти**.**» (О.Я. Гелих, НМ № 11.С. 22)

Авторское намерение реализуется в ***коммуникативных стратегиях*** автора, которые он использует в целях **информирования**, **воздействия** на читателя, чтобы заинтересовать, вызвать реакцию читателя.

Стратегия ***информирования*** с целью дать новое знание/новое пред- ставление о предмете исследования тесно связана с стратегиями убежде- ния, оценки, формирования отношения и др. Например, стратегия ***апелля- ции*** к читателю с целью описания обстоятельств может также содержать, ***презентацию***, ***признание***, ***оценку***, ***констатацию факта***:

«Однако, получив **от известного литературоведа С. И. Тиминой** в подарок **роман А. Белого “Москва”** с напутствием **по-новому осмыс- лить** пунктуационно-графическое его оформление, **признаюсь**, что анализ творческих попыток А. Белого использовать графический рисунок

текста, а также его пунктуационную аранжировку “мерцающих смыслов” **существенно изменило сложившееся представление** о так называемой авторской пунктуации.» (Н.Л. Шубина, УНж № 20. С. 44)

«**Вопрос** о месте познавательного приема в системе научного знания **всегда был актуальным**.» (Н.К. Матросова, НМ № 11. С. 41) (***актуали- зация ситуации***)

«Однако именно тут и кроется та опасность, **избегнуть** которую **необходимо**, чтобы анализ опыта не превратился в его чудовищную про- тивоположность.» (А.Е. Радеев, НМ № 11. С. 32) (***обоснование действий***)

«Общепризнано, что английский язык превратился в самую доходную статью экспорта <...> **поскольку все еще господствует миф** о том, что лучшими преподавателями иностранного языка могут быть только так называемые “носители языка, то есть native speakers”.» (В.В. Кабакчи. Глобанглизация: легко ли быть языковым лидером?// Актуальные про- блемы современной лингвистики. Язык во всех измерениях: сб. статей к юбилею профессора И.Б. Руберт.Вып.3. СПб.: СПбГУЭФ, 2011. С. 193– 200.) (***формирование мнения***)

«В 2010 г. сотрудники музея протоисториии доистории (Берлин) пре- доставили мне возможность работать с хранящимся в этом музее архивом

<...> Пользуясь случаем, **выношу признательность** нашим берлинским коллегам за возможность работы в их архиве и фондах.» (В.И. Кулаков, УНж 20. С. 17) (***выражение признательности****)*

Для ***обобщения*** результатов или концепции используется **стратегия иллюстрирования:**

«**Суммируя вышесказанное**, все возможные пути решения онто- логической проблемы сознания **можно представить в виде наглядной схемы**...» (И.. Черепанов, НМ № 10. С. 83)

Коммуникативной целью современных научных текстов является **воздействие** на читателя. Стремление автора использовать средства, при- влекающие внимание, вызывающие интерес, побуждающие к действию, иронично иллюстрирующих основную мысль, чтобы передать экспрес- сию, реализуется в результате использования в речи

1. ***метафор, окказионализмов, вопросов, крылатых выражений, цитат:***

**«“Воспитание глаза”** и методические проблемы формирования профессионального восприятия у студентов-искусствоведов» (О.В. Суб- ботина, НМ № 10. С. 142)

1. ***уточнений,*** заключающихся в скобки (поясняющие содержание, даты, названия произведений, ФИО деятелей эпохи и др.):

«Особые случаи перевода глаголов движения с русского на итальян- ский язык (**стилистика и переводческая эквивалентность**).» (Р. Пит- талуга, УНж № 20. С. 147)

«Наиболее крупной новой работой, посвященной библеизмам, явля- ется вышедшая в 2011 г. книга К.Х. Мартлинга **“Смерть в кастрюле” —**

Doden i gryton (сюжет об отравленной пище).» (Л.А. Нарусевич, УНж № 20. С. 218)

**«“Культурный ренессанс начала века”** (Н. А. Бердяев) неразрывно связан с творчеством А. Белого.» (Н.Л. Шубина, УНж № 20. С. 44)

С целью **привлечения внимания** к содержанию научного текста автор использует следующие приемы:

* 1. ***указание на важность информации, с использованием таких структур, как*** *особо следует отметить важную роль, особую актуаль- ность приобретают, важно отметить* и др.

«**Особую актуальность** вопрос реализации программы развития позитивного родительства, по нашему мнению, приобретает также при рассмотрении вопроса социализации молодежи в условиях детского дома.» (Е.Н. Рыбакова, НМ № 11. С. 17)

* 1. ***комментарий с целью воздействия*** : *интересно, получает наи- высший балл, это выразительный пример, ключевая роль отводится, наиболее влиятельной* и др.

«Усиление контраста рельефа и фона - **ведущая, но не единственная** направленность раннеромантического фактурообразования**.**» (Е.Л. Алек- сандрова, УНж № 20. С. 111)

* 1. ***конструкции, демонстрирующие отношение*:** *к сожалению, как правило, однако, очевидно, тем не менее, несомненно, без сомнения* и др.

«...что **является совершенно уникальным** и не встречается на дру- гих инструментах того времени. » (А.А. Кизиляев, УНж № 20. С. 256)

С целью **управления** процессом восприятия, **управления внима- нием** адресата используется метатекстовый комментарий и следующие конструкции и средства:

1. для фиксации этапов своей деятельности ***для общей деятельности*** с читателем: *далее, рассмотрим структуру каждого компонента, при- ступим к рассмотрению первого случая, сделаем ряд важных выводов, заметим, перейдем к следующему вопросу, перейдем к обсуждению, как будет показано далее и др.:*

«**Посмотрим сейчас последние два примера**, где присутствует выражение *вышел на улицу*, которые **для удобства** будут сопоставлены вместе.» (Р. Питталуга,УНж № 20. С. 153)

«**Как показало дальнейшее развитие** этой идеи, синтез здесь не получился. » (Е. Ю. Сиверцев, НМ № 10. С. 87)

1. для ***активизации внимания*** используют вводные слова, регули- рующие ход рассуждения: *итак, следовательно, предположим, допу- стим* и др.

«**Итак**, в первую очередь продуктивной можно считать ту тео- рию, которая последовательно раскрывает новое проблемное поле.» (О.Я. Гелих, НМ № 11. С. 33)

1. ***конструкции, содержащие рекомендации***: *стоит учитывать, необходимо различать, следует применять, лучше употребить* и др.

«**Стоит разделить** деятельность научно-литературного общества и общества по изучению народной литературы и краеведению.» (Р. В. Доро- хина, НМ № 11. С. 61)

На **интеллектуальное воздействие** направлены ***метакогнитивные регуляции***, осуществляемые посредством стратегий (обоснование, объ- яснение, убеждение), понимание которых требует информационных, критических, культурных умений. **Управление** текстом направлено на решение цели, координацию действий, контроль, с использованием авто- ром как языковых, так и параграфических средств.

«Вышедший в 1939 г. словарь крылатых слов и выражений П. Хольма... до сих пор считается наиболее фундаментальным исследо- ванием по этому вопросу (**его 15-е издание в редакции С. Экбу вышло в 1989 г**.).» (Л. А. Нарусевич. Унж № 20. С. 217) (***Уточнение***)

«Не отрицая наличия в каждом тексте «авторства», **смею предполо- жить**, что русская пунктуация формировалась как средство организации текста.» (Н. Л. Шубина, УНж № 20. С. 43) (***Предположение***)

«В этом отрывке **мы специально решили** включить и контекст, предшествующий моменту, когда Самойленко возвращается в гостиную, **чтобы показать**, как они изначально оказались в кабинете Никодима Александровича. » (Р. Питталуга,УНж № 20. С. 151) ; «Сам осужденный

имел лишь номер, носил маску (**чтобы не запоминался**), был обязан молчать. » (А.Н. Яшин, НМ № 10. С. 27) (***Объяснение***)

«**Это свидетельствует еще и о практике обучения** китайских сту- дентов за рубежом в этот период.» (Юань Фаньфань, НМ № 10. С. 44-45) (***Доказательство***)

Для правильного понимания смыслового содержания **с** целью **объ- яснения явлений** используются ***риторические вопросы***, которые помо- гают сфокусировать внимание читателя на концептуальную информацию, аргументацию доказательства или направить ход рассуждения в нужное русло. Постановка вопроса, помогает читателю выдвинуть гипотезу, имплицитный и эксплицитный смысл текста, определиться с оценкой, вызвать реакцию:

«Вопрос **“**Cui bono” (Кому это выгодно**?**) на первый взгляд является риторическим. преимущества языкового лидерства, по крайней мере на первый взгляд, очевидны.» (В.В. Кабакчи 2011. С. 193–200)

Для выяснения читателем неизвестной **информации, осуществления поиска** используются ***вопросно-ответные конструкции***:

«Зададимся вопросом, **насколько правомерным может быть подоб- ное соотнесение музыкальных образов веберовской сонаты с прооб- разами Фауста и Мефистофеля?**»(Е.В. Бабухина, УНж 20. С. 190)

или с целью **обратить внимание на исследуемый предмет**:

«Однако основное внимание исследователей обращено, как правило, к функциям и прагматике этих языковых образований (фразеологических единиц) в современных языках без учета их исторического развития. Воз- никает вопрос**: подвергаются ли развитию эти языковые комплексы или остаются неизменными и именно поэтому воспринимаются как своего рода архаизмы?**» (Г.А. Бабаева, УНж № 20. С. 37)

Использование средств аргументации, убеждения с привлечением цитат выявляет четкую авторскую позицию, согласие/несогласие с дру- гим исследователем или является способом ***организации диалогового общения,*** приглашением к дискуссии. При этом в условиях современных технологических способов быстрого воспроизведения и передачи инфор- мации автор осуществляет защиту интеллектуальной частной собствен- ности, четко отделяя собственный текст от чужого, используя средства метаграфемики:

«В данном утверждении **нельзя не согласиться с теорией Д.С. Докучаева** о понимании “региональной идеи”. » (С.С. Касаткина,

НМ № 10. С. 57) или

«Данный памятник археологии, **зафиксированый мною** как “грун- товой могильник-3 Окунево”, является южной частью могильника Гре- битен.» (В.И. Кулаков, УНж № 20. С. 16)

Средства аргументации, убеждения, служащие ***для облегчения вос- приятия и фокусирования внимания:***

«Все духовные предметы существуют в определенной связи друг с другом, **а значит**, **можно сделать вывод**, что символ и есть бытие.» (Н. М. Пустовойт, НМ № 11. С. 83)

Для **осмысления, понимания текста** на содержательном уровне авторы используют различные варианты ***оценки*** событий **или** вводят информацию с целью корректной **интерпретации** текста.

Оценка содержит прямое и косвенное воздействие при взаимодей- ствии, стимулирует осуществление поиска новых научных идей, связывает с прошлым опытом, сопровождает процесс деятельности:

«Все **идеи были призрачные**, потому и ассоциации были их же проис- хождением. Они, идеи, попросту отражались **как в «Королевстве кривых зеркал»** и вызвали невероятные ассоциации…» (И.Н. Супаров, НМ № 11. С. 66) (*Отрицательная оценка ситуации :критика, недовольство, жалоба*)

«Но в отличие от традиционных изображений девушек, на фото можно заметить **взаимодействие модели с камерой**, она как бы **стре- мится показать весь свой шарм и достоинства**, чтобы привлечь вни- мание зрителей.» (В.Б. Пантелеева, УНж № 20. С. 171) (*положительная оценка ситуации похвала)*

Невербальные знаки выражения ***оценки*:** *пунктуация*: восклицатель- ные и вопросительные знаки, многоточие, тире.

«“Пирамиды Мавроди и залоговые аукционы” вышли из тени про- шлого на передний план. О них надо говорить спокойно, ибо они стали обычным, рядовым и социальным явлением. Но почему они были запу- щены в народ**?!**» (И.Н. Супаров, НМ № 11. С. 68)

*Графические*: выделение шрифтом, разбивка, черта и др.

«Развитие технологии, включающее в себя не только орудия труда и ремесел, но и, **что еще более важно**, *управленческих навыков*, — гово- рится здесь, — функционально привязано к процессам формирования цивилизации» (**курсив наш**. — *О. Г*.) (О. Я. Гелих. НМ № 11. С. 26)

В условиях интегративных процессов в науке многие явления **описы- ваются** с различных аспектов. По ***уточнению*** значения терминов часто используются ***отсылки на словарь*** или чужой текст, такая отсылка может содержать ***комментарий*** и оформляется в тексте в скобках или отдельно, в примечании.

«Упоминания о дольцайне встречаются **во многих литературных произведениях**. Их можно найти **в “Песеннике Баены”** (“Cancionero de Baena”), **составленном Хуаном Альфонсо де Баена** (Juan Alfonso de Baenaок. 1375–1434)...» (А.А. Кизиляев, УНж № 20. С. 252–253)

«Поэтому вариант, предложенный двумя другими переводчиками (соответственно Mancosu и Bietti), *entro* **кажется отчасти менее точным** особенно потому, что создается впечатление, что моменты до и после раз- говора не связаны между собой...» (Р. Питталуга, УНж 20. С. 152) ***(Для обоснования***)

Кроме этого, наблюдаются приемы распределения информации. Рас- пределяя текст в соответствии с намерением, автор помогает читателю сориентироваться в структуре и композиции текста. Имеются в виду приемы организации внимания: ***логические усилители слов****,* уточняю- щие смысл высказывания; ***текстовые метаграфемы***, осуществляющие выделительную функцию (шрифты, рамки, прописные буквы, пробелы, отступы и др.; ***структурные выделения*** (втяжки, эпиграфы, списки, примеры, таблицы, сноски, примечания); элементом композиционного построения текста являются ***подписи к рисункам*** (в красную строку, выключка в край, над рисунком, сбоку рисунка и др.);

Таким образом, современный научный текст направлен на реализацию коммуникативных задач не только автора, но и читателя, реализует лич- ностные намерения, идеи, взгляды автора, основывается на диалоговом принципе для социального взаимодействия. Выявление особенностей изменения научного текста играет важнейшую роль для понимания меха- низма кодирования/декодирования и специфики репрезентации научной информации.

**Принятые сокращения**

НМ — научный журнал «Научное мнение»; УНж — «Университетский научный журнал».

**Литература**

* 1. *Макаров В. Л*. Экономика знаний: уроки для России. [URL:http://vivovoco.](http://vivovoco/) astronet.ru/VV/JOURNAL/VRAN/SESSION/VRAN5.HTM (дата обращения 14.09.16)

**Literature**

1. *Makarov V. L.* Ekonomika znaniy: uroki dlya Rossii. URL: [http://vivovoco.](http://vivovoco/) astronet.ru/VV/JOURNAL/VRAN/SESSION/VRAN5.HTM (data obrashcheniya 14.09.16)

***Е.И. Баранчеева***

**ОПЫТ РАЗРАБОТКИ И ВНЕДРЕНИЯ ЭЛЕКТРОННЫХ КУРСОВ ПО ИНОСТРАННОМУ ЯЗЫКУ**

**В АППАРАТНО-ПРОГРАММНОЙ ЯЗЫКОВОЙ СРЕДЕ ELANG**

**Аннотация.** Работа посвящена вопросам использования информационно- коммуникационных технологий в преподавании иностранных языков, обсуждается опыт разработки аппаратно-программной языковой среды, которая предназначена для создания электронных учебных курсов по иностранным языкам. В качестве примера анализируется методическая деятельность рабочей группы по созда- нию электронных курсов русского языка как иностранного и английского языка. В заключение авторы останавливаются на возможностях использования разрабо- танной языковой среды преподавателями и студентами.

**Ключевые слова.** Информационно-коммуникационные технологии в препо- давании иностранных языков, электронная языковая среда, преподавание русского языка как иностранного.

***E.I. Barancheeva***

**DESIGN AND APPLICATION OF FOREIGN LANGUAGE COURSES IN ELECTRONIC LANGUAGE ENVIRONMENT ELANG**

**Abstract.** The current paper is devoted to the issues of application of informational and communicative technologies in foreign language teaching. These issues are in the focus of the paper through the discussion of the designing of the electronic language environment eLang (elang.nstu.ru) in NSTU (Russia). The author describes the condi- tions of elaborating of this educational environment, designed for the language courses. As a conclusion the paper touches the possibilities of using of the newly elaborated language environment both for students and language instructors.

**Key words.** Informational and communicative technologies in foreign language teaching, Electronic language environment, teaching Russian as a foreign language.

Развитие методики преподавания иностранных языков в наши дни свя- зано с активным изучением возможностей использования информационно- коммуникационных технологий (ИКТ) в процессе обучения и с созданием новых ресурсов для обучения иностранному языку, в том числе для обу- чения профессиональному иностранному языку. В современной методике

© БАРАНЧЕЕВА Екатерина Игоревна. — Barancheeva Ekaterina. Новосибирский государ- ственный технический университет, Россия. — Novosibirsk State Technical University, Russia. Е-mail: [barancheeva\_ek@yahoo.com](mailto:barancheeva_ek@yahoo.com)

важное значение принадлежит теоретическим работам, посвященным разработке качественных учебных материалов на основе актуальных лингвометодических подходов, а также новым аппаратно-программным продуктам, которые могут стать частью учебного процесса и повысить уровень преподавания иностранных языков в высшей школе. Данная работа связана с проблематикой современной компьютерной лингводидак- тики и предлагает анализ реализации проекта по разработке электронной языковой среды для изучения иностранных языков.

В 2012 г. Новосибирский государственный технический университет (НГТУ) начал реализовывать проект под названием «Разработка универ- ситетской аппаратно-программной языковой среды». Работа над данной языковой средой была частью более широкого университетского проекта

«Модернизация образовательного процесса» программы стратегического развития НГТУ на 2012–2016 гг. и курировалось Институтом дистан- ционного образования НГТУ. Цель проекта — на основе современных требований электронной лингводидактики разработать онлайн-оболочку для создания электронных учебных материалов по иностранным языкам, предназначенных для самостоятельного совершенствования уровня вла- дения иностранными языками студентами, магистрантами, аспирантами, преподавателями и сотрудниками НГТУ. Официальный университетский выпуск новой аппаратно-программной среды состоялся 14 мая 2014 г. Новая языковая среда получила название eLang (elang.nstu.ru). В рамках данного проекта была создана рабочая группа из преподавателей ино- странных языков и технических специалистов, которые подготовили концепцию новой аппаратно-программной языковой среды, сформировали техническое задание, разработали банк учебных материалов по различным иностранным языкам.

На различных этапах работы осуществлялось тестирование языковой среды, корректировался дизайн, функциональные возможности среды, типы учебных заданий. Работа по тестированию аппаратно-программной среды была связана с апробированием как вариантов работы препода- вателя в данной системе (в режиме создания и редактирования курсов, модулей, отдельных заданий), так и возможностей самостоятельной работы студента. Тестирование режима «Преподаватель» включал реше- ние вопросов по авторизации пользователя-преподавателя (или несколь- ких преподавателей, работающих над курсом), по процессу создания курса и его структуры, по процессу создания упражнений различных типов и созданию дополнительных настроек упражнений. Тестирование системы в режиме пользователя-студента решало вопросы построения индивидуальной траектории обучения, возможностей быстрого доступа к различным разделам модуля и курса, возможностей самопроверки,

контроля и анализа допущенных ошибок. По результатам тестирования электронной оболочки были сформированы предложения для технических разработчиков системы по совершенствованию ее функционирования в режиме «Преподаватель»/«Студент».

Создание и редактирование заданий проводилось преподавателями самостоятельно в онлайн-режиме редактирования курса. Оболочка также имела опцию режима просмотра (т.е. в процессе наполнения курса пре- подаватель мог видеть создаваемый курс так, как его увидят студенты).

В работе над созданием оболочки приняли участие преподаватели раз- личных иностранных языков НГТУ, работающие на разных факультетах. В результате к концу 2013/2014 уч. г. были созданы электронные курсы по русскому языку как иностранному (уровень A2), английскому языку (English B1 и Press Ofﬁce (English for PR), немецкому языку (Deutsch für Anfänger — A1), деловому японскому языку и китайскому языку для начи- нающих. Начата разработка курса польского языка (уровень В1).

Разработанный нами учебный курс по русскому языку как иностран- ному (базовый уровень — A2) ориентирован на обучение иностранных студентов, владеющих русским языком в объеме элементарного уровня. Достижение базового сертификационного уровня владения русским язы- ком позволит удовлетворить основные коммуникативные потребности иностранных граждан средствами русского языка в социально-бытовой и социально-культурной сферах общения. Владение языком на данном уровне позволит быстрее адаптироваться в условиях новой социально- культурной сферы, устанавливать контакты с носителями русского языка в условиях языковой среды и вне ее, подготовиться к изучению русского языка на пороговом (первом сертификационном) уровне. Учеб- ный курс по русскому языку как иностранному направлен на взаимос- вязанное обучение аспектам языка и видам речевой деятельности для формирования у иностранных учащихся коммуникативной компетент- ности. Построение курса базируется на принципах коммуникативно- ориентированной методики, прежде всего тематико-ситуативном прин- ципе организации учебного материала.

Методическая основа создаваемого курса и содержание учебных материалов по русскому языку как иностранному определяются стан- дартными требованиями, предъявляемыми к базовому уровню общего владения русским языком как иностранным. При разработке курса мы ориентировались, прежде всего, на государственные образовательные программы, стандарты и лексические минимумы, принятые в Российской Федерации ([Программа по русскому языку 1984], [Программа-справочник 1995], [Государственный стандарт 2001], [Образовательная программа

2001], [Андрюшина 2011; Антонова 2012]).

Курс для самостоятельного изучения русского языка как иностранного базового уровня включает входной тест (60 заданий), 9 модулей и итого- вый тест (75 заданий). Основные темы модулей следующие:

* Модуль 1. О себе
* Модуль 2. Моя семья
* Модуль 3. Моя учеба
* Модуль 4. Мой дом
* Модуль 5. В магазине
* Модуль 6. У врача
* Модуль 7. Мой рабочий день
* Модуль 8. Свободное время
* Модуль 9. Мой город. Транспорт

Девять модулей соответствуют темам и ситуациям общения, которые прописаны в Государственном стандарте по русскому языку как ино- странному в разделе «Содержание коммуникативно-речевой компетенции» [Государственный стандарт 2001], именно эти темы значимы для обучения иностранных учащихся продуцированию диалогической и монологи- ческой речи. Каждый модуль содержит различные виды заданий, в том числе работу с текстами разной степени сложности. Общее количество заданий по модулям различалось, но не превышало 30 заданий. Разрабо- танная оболочка не предполагает контактное расположение нескольких заданий, связанных с одним текстом, — притекстовых и послетекстовых заданий. Данный курс РКИ содержит тексты в графической форме и в форме аудиозаписей, дополненных транскриптом. Большая часть тек- стов имеет страноведческую направленность, позволяет узнать о жизни в современном российском городе, о российской семье, учебе в россий- ском вузе и т. д.

Сейчас перед авторами стоят задачи интеграции созданного курса в учебный процесс, а также адаптации курса для целей самостоятель- ного изучения языка внешними пользователями. Авторы курса надеются также обогатить текстовую базу, записав оригинальные аудио- и видео- материалы. К сожалению, все аудио- и видеоматериалы, находящиеся в открытом доступе, не соответствуют базовому уровню владения русским языком: аутентичных материалов данного уровня или недо- статочно, или нет.

Отметим, что коллектив ИДО НГТУ и преподаватели иностранных языков создали электронную языковую среду eLang, которая позволяет пользователю изучать иностранный язык самостоятельно, в интерак- тивном режиме, в рамках традиционного университетского курса или удаленно. Такая среда расширяет и возможности преподавателей, которые теперь могут:

* + создавать необходимое количество учебных курсов по иностранным языкам без помощи программистов;
  + создавать в структуре каждого курса интерактивные задания, спра- вочные материалы и тесты;
  + оформлять учебные задания в учебные модули;
  + использовать 12 видов интерактивных заданий;
  + загружать в систему заданий текстовые и мультимедийные мате- риалы (изображения, аудио, видео);
  + редактировать разработанные курсы.

Студенты получают возможность выбрать курс иностранного языка, пройти стартовое и итоговое тестирование, определить собственный поря- док освоения материалов курса и, далее, обращаться к материалам курса неограниченное количество раз. Дополнительными преимуществами являются быстрый доступ к справочным материалам, пошаговая оценка выполнения заданий, обращение к мультимедийным ресурсам и интер- фейсу на изучаемом языке.

**Литература**

1. *Андрюшина Н.П. и др.* (2011), Лексический минимум по русскому языку как иностранному. Базовый уровень. Общее владение. СПб.: Златоуст, 116 с.
2. *Антонова В.Е., Нахабина М.М., Толстых А.А.* (2012), Типовые тесты по русскому языку как иностранному. Элементарный уровень. Общее владение. Варианты, СПб.: Златоуст, 112 с.
3. Государственный стандарт по русскому языку как иностранному. Базовый уровень [под. ред. М.М. Нахабиной]. 2-е изд., испр. и доп. М. — СПб.: Златоуст, 2001, 32 с.
4. *Иванова А.С., Соболева Н.И., Пушкова Н.П.* (2014), Программа-справочник по русскому языку как иностранному. С комментарием на английском языке. М.: Изд-во РУДН, 293 с.
5. Образовательная программа по русскому языку как иностранному. Элемен- тарный уровень. Базовый уровень. Первый сертификационный уровень [под. ред. З.И. Есиной]. М.: Изд-во РУДН, 2001, 134 с.
6. Программа по русскому языку для студентов-иностранцев, обучающихся на подготовительных факультетах вузов СССР. М.: Русский язык, 1984.
7. Программа-справочник «Русский язык как на ладони». М.: Изд-во РУДН, 1995.

**Literature**

1. *Andryushina N.P.* i dr. (2011), Leksicheskiy minimum po russkomu yazyku kak inostrannomu. Bazovyi uroven’. Obshcheye vladenie. SPb.: Zlatoust, 116 s.
2. *Antonova V.E., Nakhabina M.M., Tolstykh A.A*. (2012), Tipovye testy po russkomu yazyku kak inostrannomu. Elementarnyi uroven’. Obshcheye vladenie. Varianty, SPb.: Zlatoust, 112 s.
3. Gosudarstvennyi standart po russkomu yazyku kak inostrannomu. Bazovyi uroven’ [pod. red. M.M. Nakhabinoy]. 2-e izd., ispr. i dop. M. — SPb.: Zlatoust, 2001, 32 s.
4. *Ivanova A.S., Soboleva N.I., Pushkova N.P*. (2014), Programma-spravochnik po russkomu yazyku kak inostrannomu. S kommentariem na angliyskom yazyke. M.: Izd-vo RUDN, 293 s.
5. Obrazovatel’naya programma po russkomu yazyku kak inostrannomu. Elementa- rnyi uroven’. Bazovyi uroven’. Pervyi sertiﬁkatsionnyi uroven’ [pod. red. Z.I. Yesinoy]. M.: Izd-vo RUDN, 2001, 134 s.
6. Programma po russkomu yazyku dlya studentov-inostrantsev, obuchayush- chikhsya na podgotovitel’nykh fakul’tetakh vuzov SSSR. M.: Russkiy yazyk, 1984.
7. Programma-spravochnik «Russkiy yazyk kak na ladoni». M.: Izd-vo RUDN, 1995.

***М.Я. Блох***

**КОММУНИКАТИВНАЯ ОТВЕТСТВЕННОСТЬ**

**В УСЛОВИЯХ ИНФОРМАЦИОННОЙ РЕВОЛЮЦИИ**

**Аннотация.** В работе рассматривается проблема ответственности комму- никантов друг перед другом за эффективное проведение акта речевого общения. Говорящий должен четко и удобопонятно передать сообщение, слушающий дол- жен постараться верно и полно понять сообщение, слышащий не должен мешать коммуникантам общаться за исключением случаев каких-либо форсмажорных обстоятельств. На этом фоне уточняется положение о пресуппозиции, преобра- зуемое в положение о понятийной триаде, выступающей в виде одного общего понятия — «коммуникативная суппозиция», и двух частных понятий — соот- ветственно «пресуппозиция» (говорящий) и «постсуппозиция» (слушающий).

**Ключевые слова.** Коммуникативная ответственность, коммуникант, гово- рящий, слушающий, слышащий, коммуникативная суппозиция, пресуппозиция, постсуппозиция.

***M.Y. Blokh***

**COMMUNICATIVE RESPONSIBILITY IN CONDITIONS OF INFORMATIONAL REVOLUTION**

**Abstract.** Considered is the problem of the communicants” responsibility before one another for the adequate result of the communicative action. The speaker should render the message in a legible and intelligible form, the listener should try intensely to understand the message, the hearer should not interfere with the communication disturbing it unless some extraordinary occurrence forces him to. Against this background the principle of presupposition is reformed, being developed into the principle of notional triad: one notion is general — “communicative supposition”, while the other two are special, namely the ﬁrst − “presupposition” (the speaker”s part of mental activity), the second − “postsupposion” (the listener”s part of mental activity). **Keywords.** Communicative responsibility, communicant, speaker, lictener, hearer,

communicative supposition, presupposition, postsupposition.

Информационная революция представляет собой радикальное и лави- нообразное преобразование характера обмена информацией в современ- ном сообществе человечества. Эта революция началась с изобретения телеграфа, затем радио и телефона и вместе с этими областями общения продолжилась бурным развитием полиграфии с линотипом и газетно- журнальным бумом, а потом изобретением телевизора, а потом вдруг кибернетикой и бушующим развитием компьютера и далее внезапно интернетом и всей коммуникативно-информационной деятельностью,

© БЛОХ Марк Яковлевич. — Blokh Mark. Московский педагогический государственный университет, Россия. — Moscow Pedagogical Snate University, Russia. E-mail: [blmy2@mail.ru](mailto:blmy2@mail.ru)

связанной с ним. Возникает вопрос: а куда должно деться «естественное» устное речевое общение? Неужели оно обречено на отмирание, и мы, сидя за дружеским столом, будем общаться посредством ноутбуков и планшетов через международную информационную сеть?

Сделаем «крутое» предположение, что этого не произойдет, что в обо- зримом будущем человек останется естественно говорящим. В самом деле, если хоть немного вдуматься, то в смущенном мозгу высветится простая истина, что за всеми как угодно замысловатыми передачами текста стоит неотчуждаемая от личности человеческая устная речь. И проблема комму- никативной ответственности, о которой я собираюсь говорить, относится в первую очередь именно к этой речи — органичной и бесценной принад- лежности личности.

Говоря о коммуникативной ответственности — ответственности в общении — спрашиваем: ответственность кого? Ответственность перед кем? Ответственность за что?

Поскольку мы рассматриваем проблему ответственности в **общении**, то ясно, что под «кого» и «кем» имеются в виду **коммуниканты** — лица, ведущие разговор или переписку [Арискина 2011]. А ответственность **за что**? Вот в этом существо дела. Это — наша **искомая величина.**

Чтобы ее найти, следует рассмотреть понятие эффективности обще- ния, а конкретнее эффективности непосредственного акта общения, реа- лизуемого диалогическим единством или диалогемой: вызов — отзыв, вопрос — ответ… [Блох, Поляков 1992].

Речь в человеческом общении — это **главная** часть поведения. Рас- крывая закономерности **речевого** поведения, мы должны строго различать поведение **говорящего** и поведение **слушающего,** прибавив к ним также и поведение **слышащего** (лица, воспринимающего речь, но непосред- ственно не участвующего в акте общения) [Блох 2013]. Если акт общения и его назначение находятся в цивилизованных рамках взаимной добро- желательности (как теперь говорят, «кооперативности»), то обычные рекомендации (или «правила») для коммуникантов можно сформулировать следующим образом.

Правила **для говорящего:** 1) правило внятности речи; 2) правило понятности речи; 3) правило вежливости (поведенческой уместности) речи; 4) правило стилеуместности речи (это правило, по существу, является собственно лингвистическим, затрагивая информационные типы речи); 5) правило кооперативного вызова (на нужную реакцию);

1. правило переакцентуации лиц (в центре внимания — слушающий);
2. правило оптимальной протяженности (краткости) речи; 8) правило разумности (логичности) речи; 9) правило интереса (интересная речь для слушающего); 10) Правило кооперативного паралингвального поведения.

Правила **для слушающего:** 1) правило внимания; 2) правило не пре- рывания; 3) правило выражения интереса; 4) правило поощрения к продол- жению речи; 5) правило активной (оптимальной) реакции на обращение Правила **для слышащего:** 1) не вмешиваться в общение, т. е.

не мешать! 2) допустить вмешательство в релевантных форс-мажорных обстоятельствах (уничижительное высказывание о слышащем, сговор против него и т. д.).

Все эти **правила прагматики** или **«правила речевого этикета»** — продолжение традиционных **«правил хорошего тона»** в европейском воплощении.

Мы отметили существенные поведенческие правила **кооперативного** общения. С другой стороны, в плане эффективности акта общения с точки зрения достижения **коммуникативной цели некооперативным** говоря- щим указанные «правила хорошего тона» закономерно пересматриваются. Положим, цель говорящего — унизить слушающего, оскорбить его, запугать его, привести в замешательство, погрузить в состояние простра- ции и т. д. И все наши добрые рекомендации отбрасываются в сторону. Вступает в бурное действие высказывательный аспект **воздействия, импрессия**, поворачивающаяся стороной **отрицательной выразитель-**

**ности**.

Аспект выразительности, соединенный с аспектами внятности и членораздельности, должен строго учитываться при реализации **любой** коммуникативной цели, кроме того случая, когда цель говорящего — сформировать нарочито невнятное высказывание (сравните, например, высказывание высокомерного чиновника, обращенное к униженному просителю). Нормальная же кооперативная речь, как я показал в одном из научных докладов, должна удовлетворять качеству ЧУВ, быть «чув- ной»: **членораздельной (внятной), уважительной, выразительной** (с точки зрения адекватности коммуникативной цели говорящего). Можно

«чувную» речь переформулировать в «речь в три В»: ВВВ — «Внятная, Вежливая, Выразительная».

У меня был случай показать, что из всех напастей, обрушивающихся в наше время на живой язык — таких как засорение его дублирующими заимствованиями, переполнение разного рода сленгизмами-жаргонизмами, огрубление элементами просторечия и т. д. — самой тяжелой напастью является усиливающаяся тенденция к **нечленораздельной, невнятной речи**. Набирает **катастрофически** все большую и большую силу **скорого- ворка**, пренебрегающая тем простым фактом, что в нормальном общении мы говорим **не для себя, а для слушающего**. Такой скороговорке пре- даемся мы и сами по себе, включаясь в бешеный темп информационных потоков; такой скороговорке обучает нас и речь радио- и телекорреспон- дентов, отвечающая утрированно заданной цели передать **все больше**

**и больше информации в единицу времени**. На деле такая скороговорка ведет к обратному эффекту: важная информация в указанном **ущербном канале** деформируется и теряется.

А она **не должна** деформироваться и теряться. Не должна — при осо- знании говорящим **коммуникативной ответственности** перед слушаю- щим. Я, говорящий, **отвечаю** за то, чтобы слушающий хорошо **услышал** мое высказывание и хорошо **понял** его, понял мой **содержательный посыл,** заключенный в вызове, и передал бы мне запрашиваемую инфор- мацию в нужном мне **адекватном отзыве**.

Итак, говорящий является **коммуникативно-ответственным перед слушающим**. Если говорящий говорит невнятно, говорит, при прочих равных условиях, непонятно для слушающего, то он является **коммуни- кативно безответственным** перед слушающим, и как таковой он достоин лишь осуждения, которое, естественно, призвано вести его к комму- никативному исправлению!

С говорящим теперь все ясно. С него **нелицеприятно спрашивается**. Но как же мы должны отнестись к **слушающему**? Имеется ли у слушаю- щего некая коммуникативная ответственность, или жизнь его полностью вольготна, и он, будучи пассивным коммуникантом, ни за что не отвечает? Само собой понятно, что вопрос этот является риторическим. Конечно, у слушающего имеется ответственность, да еще какая! Она состоит в его моральном долге перед говорящим слушать и слышать его речь, неукосни- тельно стараясь **понять все ее посылы**, все ее явно и неявно выраженные смыслы [Блох 2007].

И когда мы обращаемся к коммуникативной ответственности слу- шающего, мы оказываемся перед большой проблемой соотносительного запаса, или фонда, знаний коммуникантов — такого **общего информа- ционного запаса**, который необходим для их взаимопонимания в ходе общения.

Рассмотрение вопроса о фонде общих знаний коммуникантов, как известно, осуществляется в лингвистической дисциплине, назвавшей себя лингвопрагматикой — наукой об использовании языка в непосред- ственных актах общения. Фонд общих знаний связывается с понятием **пресуппозиции — «пред-предположения»**. Пресуппозиция так и опреде- ляется по своей букве как фонд общих знаний говорящего и слушающего, необходимый для адекватного восприятия и понимания **речи говорящего со стороны слушающего** [Арутюнова 1973].

Плодотворность этого понятия самоочевидна; его использование в раз- ных школах языкознания помогло глубоко раскрыть явление подтекста и подойти с новых позиций к проблеме **смысла сообщения**.

Понятие пресуппозиции, выдвинутое во второй половине прошлого века, закономерно развиваясь по мере его применения в лингвистических

наблюдениях, все более отчетливо обращалось своей релевантностью не к слушающему как таковому, а к говорящему. В самом деле, ведь не слушаю- щий, а говорящий начинает общение; не слушающий, а говорящий делает изначальное предположение в рамках формирования вышеуказанного

«фонда общих знаний». А раз это так, то осмыслить и определить понятие пресуппозиции следует теперь не просто в качестве соответствующего

«фонда общих знаний», а как **предположение говорящего** о наличии **у слушающего** суммы сведений, необходимых для адекватного понимания высказывания (или смысла высказывания).

Следовательно, пресуппозиция есть движение мысли **говорящего**. Но параллельно с движением мысли говорящего должно, по необходи- мости, осуществляться и движение мысли **слушающего**, направленное на интерпретацию воспринимаемого высказывания. Если у говорящего имеется предположение относительно отправного информационного фонда слушающего, то у слушающего, со своей стороны, должно быть сформировано предположение об отправной («оперативной») **инфор- мационной базе** говорящего, необходимое для адекватного понимания его высказывания. Если предположение говорящего называется «пре- суппозицией», то предположение слушающего разумно назвать **«пост- суппозицией».** Таким образом, мы будем различать, во-первых, общее понятие «**коммуникативной суппозиции**», во-вторых, частное понятие

«коммуникативной **пресуппозиции**» (говорящий) и, в-третьих, частное понятие «коммуникативной **постсуппозиции**» (слушающий).

Коммуникативная суппозиция реализуется в конкретных условиях двух контекстов: горизонтального (широкая ситуация общения) и вер- тикального (общий культурный фон общения). Для достижения успеха в акте общения суппозиции коммуникантов должны возможно более строго соответствовать друг другу. Чтобы соответствовать друг другу, а в идеале быть эквивалентными или тождественными, они должны адекват- ным образом строиться на аналогичных оперативных информационных базах, т. е. на релевантном фонде общих знаний, органически связанных с обоими контекстами высказывания — горизонтальным и вертикальным. Указанное требование настолько важно, что может быть возведено в ранг **коммуникативно-когнитивного закона**. Нарушение этого закона (т. е. закона эквивалентности коммуникативной суппозиции), выражающееся в возникновении **ложной пресуппозиции** или **ложной постсуппозиции**, ведет к **информационному сбою** — ошибочной интерпретации смысла высказывания **слушающим. Эта ошибочная интерпретация сплошь и рядом происходит из-за коммуникативной безответственности слушающего**.

Проблема **коммуникативной суппозиции** и непосредственно свя-

занная с ней проблема **коммуникативной ответственности**, лежащих

в основе достижения взаимопонимания в акте общения, относятся к числу важнейших проблем человеческого общежития. В свете сформулирован- ного закона эквивалентности коммуникативной суппозиции, необходимой для взаимопонимания участников акта общения, **зачин** общения, т. е. установление коммуникативного контакта, должен быть опосредован **включением аппарата супппозици**. Соответственно, **завершение** обще- ния, размыкание коммуникативного контакта, должно сопровождаться **выключением аппарата суппозиции**. Ошибочная интерпретация высказывания, вызванная ложными суппозициями, приводит к различным недоразумениям и бедам, в том числе обидам и ссорам говорящих, а в случаях общения политических сторон — к межкультурным и межгосу- дарственным конфликтам [Тер-Минасова 2004]. Отсюда особая актуаль- ность рассмотрения этих проблем для современного этапа существования человечества, отличающегося бурным развитием межкультурного взаимо- действия людей на всех уровнях общественного устройства.

**Литература**

* 1. *Арискина О. Л.* (2011), Языковая и коммуникативная личность: различные подходы к исследованию: Вестник Челябинского государственного университета,

№ 25.

* 1. *Арутюнова Н.Д.* (1973), Понятие пресуппозиции в лингвистике: Известия АН СССР, Сер. литературы и языка, т. 32, № 1.
  2. *Блох М.Я., Поляков С. М.* (1992), Строй диалогической речи, М.
  3. *Блох М.Я.* (2007), Диалог, монолог и фактор слушающего: Научные труды Московского педагогического государственного университета, Филологические науки, сб. статей. М.
  4. *Блох М.Я.* (2013), Язык, культура и проблема регуляции речевого общения // Язык. Культура. Речевое общение, № 2*.*
  5. *Тер-Минасова С.Г.* (2004), Язык и межкультурная коммуникация. М.

**Literature**

1. *Ariskina O. L*. (2011), Yazykovaya i kommunikativnaya lichnost’: razlichnye podkhody k issledovaniyu: Vestnik Chelyabinskogo gosudarstvennogo universiteta,

№ 25.

1. *Arutyunova N.D*. (1973), Ponyatie presuppozitsii v lingvistike: Izvestiya AN SSSR, Ser. literatury i yazyka, t. 32, № 1.
2. *Blokh M.Ya., Polyakov S. M*. (1992), Stroy dialogicheskoy rechi, M.
3. *Blokh M.Ya.* (2007), Dialog, monolog i faktor slushayushchego: Nauchnye trudy Moskovskogo pedagogicheskogo gosudarstvennogo universiteta, Filologicheskie nauki, sb. statey. M.
4. *Blokh M.Ya*. (2013), Yazyk, kul’tura i problema regulyatsii rechevogo obsh- cheniya // Yazyk. Kul’tura. Rechevoe obshchenie, № 2.
5. *Ter-Minasova S.G*. (2004), Yazyk i mezhkul’turnaya kommunikatsiya. M.

***Л.К. Гейхман, И.В. Ставцева***

**ПЕДАГОГИЧЕСКИЙ ПОТЕНЦИАЛ МАССОВЫХ ОТКРЫТЫХ ОНЛАЙН-КУРСОВ**

**Аннотация.** Статья посвящена анализу педагогической сущности феномена массовых открытых онлайн-курсов (МООК): в фокусе внимания авторов их особенности в сравнении с традиционным обучением в вузе и потенциал для высшего профессионального образования. Приводится сравнение зарубежных и российских платформ. Детально анализируется специализация по лингводи- дактике и методике преподавания английского языка, разработанная ППС госу- дарственного университета штата Аризоны, «Сертификат TESOL, Часть 1. Как преподавать английский»: дается ее характеристика с педагогической точки зрения и рассматриваются перспективы использования данного сертификата и подобных ему в системе российского высшего образования.

**Ключевые слова.** Массовые открытые онлайн-курсы (МООК), высшее про- фессиональное образование, методика преподавания английского языка, Coursera.

***L.K. Geykhman, I.V. Stavtseva***

**MOOCS EDUCATIONAL VALUE**

**Abstract.** The paper deals with the Massive Open Online Courses (MOOC) edu- cational phenomenon: the authors emphasize their peculiarities in comparison with traditional education and the opportunities they create for professional higher education. The paper suggests Russian and international learning platforms comparison. Besides, there is a detailed analysis of English teaching specialization, made by Arizona State University *TESOL. Part I. How to teach English*. The authors point out its educational features and show its potential for Russian educational system.

**Keywords.** Massive open online courses (MOOCs), professional higher education, TESOL, Coursera.

С распространением влияния новых информационных технологий и интернета массовые открытые онлайн-курсы (МООК, от *англ*. Massive Open Online Courses — MOOC) набирают популярность, развиваются, перерастая из информационных источников в полноценный способ повы- шения квалификации и получения профессионального образования. Так, на платформе [https://www.coursera.or](http://www.coursera.org/)g/ существует система специализаций (specializations), каждая из которых построена на определенном наборе МООК (например, см. [7] — специализация по методике преподавания английского языка, состоящая из трех курсов и дипломного проекта).

© ГЕЙХМАН Любовь Кимовна. — Geykhman Lyubov. Пермский национальный исследо- вательский политехнический университет, Россия. — Perm National Research Polytechnic University, Russia. E-mail: [lyubageykhman7@gmail.com](mailto:lyubageykhman7@gmail.com)

© СТАВЦЕВА Ирина Вячеславовна. — Stavtseva Irina. Южно-Уральский государственный университет, Россия. — South Ural State University, Russia. E-mail: [stavtsevaiv@susu.ru](mailto:stavtsevaiv@susu.ru)

В свою очередь, платформа [https://www](http://www.edx.org/).edx.or[g/](http://www.edx.org/) предлагает курсы для старшеклассников, программы профессионального образования и возмож- ность «перезачета» университетских предметов. На https://www.udacity. com/ активно рекламируются программы получения «нано»-степени (nanodegree program). В России МООК в основном активно продвигают крупные университеты и образовательные интернет-площадки. Кроме того, в 2015 г. появилась Национальная платформа открытого образования https://openedu.ru/, предлагающая онлайн-курсы, которые разрабатываются в соответствии с требованиями федеральных государственных образова- тельных стандартов и нацелены на возможность встроить данные про- дукты в учебный план любого российского вуза. Что же делает данные предложения успешными? Каковы педагогические особенности МООК? Что их отличает от традиционного способа получения высшего образова- ния в стенах университета или повышения квалификации? Каким образом можно использовать МООК в системе высшего российского образования вообще и в его лингводидактической отрасли в частности?

Характеристики успешности МООК в целом заложены в самом

названии: массовость и открытый доступ из любой точки планеты через интернет, когда любой желающий может записаться на курс и пройти его в комфортное время, находясь у себя дома. Так, миссия МООК видится в создании учебных курсов совместными усилиями лучших представите- лей международного педагогического сообщества и возможности сделать их доступными каждому желающему в любой точке земного шара. Важно отметить, что данная миссия активно реализуется посредством множества МООК-платформ в разных странах, предоставляющих доступ к курсам для сотрудников университета. В последнее время все больше университе- тов в России и за рубежом не только выкладывают курсы на МООК плат- формах, но и вырабатывают систему «перезачета» дисциплин студентам (см., например, статью об опыте ТГУ [Бабанская 2015]).

В качестве составляющих МООК А.А. Андреев приводит [Андреев 2013]: 1) изучение содержания; 2) консультации; 3) участие в вебинарах;

1. постоянный контроль. Данные составляющие могут быть по-разному реализованы, однако носят универсальный характер. Из организационно- педагогических особенностей МООК справедливо выделяют: 1) бес- платное и добровольное обучение; 2) большое количество обучающихся на курсе; 3) использование интернет-инструментов; 4) возможность проводить занятия различных видов; 5) активизация самостоятельной деятельности обучающихся [Андреев 2013; Бабанская 2015; Романова, Апухтин, Белоножкин 2011; Тихомирова 2016].

Дадим комментарии по организационно-педагогическим пунктам. Во-первых, обучение на МООК, разумеется, условно бесплатное:

способы монетизации курсов существуют и самый распространенный из

них — плата за сертификат с некоторыми ограничениями, которые всту- пают в силу в случае отсутствия данной платы (например, автоматическая оценка промежуточных контрольных тестов на курсе “Lesson Planning with the ELL in Mind” возможна только при оплате курса [Manos 2016]). Кроме того, последние исследования рынка свидетельствуют о росте про- даж сертификатов онлайн-курсов за счет их бюджетности по сравнению с традиционными курсами повышения квалификации [Носкова 2016].

Во-вторых, «занятия различных видов» — понятие не безграничное: например, классические лекции в формате МООК действительно приоб- ретают совершенно новый смысл, а вот взаимодействие с реальным препо- давателем ограничено, или, если более точно, опосредовано технологиями, которые вносят свои коррективы [Гейхман, Ставцева 2015].

В-третьих, нельзя однозначно положительно или отрицательно оценить пункт «активизация самостоятельной деятельности», поскольку, если по статистике только около 10–12% [Бабанская 2015] слушателей заканчи- вает курс с сертификатом, а большинство его не проходят полностью, это означает, что «активизация» происходит у единиц. С другой стороны, здесь возникает множество факторов, как субъективных (мотивация, приоритеты и планирование времени слушателя), так и объективных (содержание курса, его организация).

В целом если смотреть глобально, то успешность МООК с педаго- гической точки зрения основывается на «социальном конструктивизме» [Романова, Апухтин, Белоножкин 2011] как особой форме взаимодействия, возникающей в той или иной степени между всеми участниками кон- кретного курса. Возникает вопрос, каким образом это можно реализовать в курсе лингводидактики и методики преподавания иностранного языка? Наш интерес к МООК по данной дисциплине обусловлен ее практико- ориентированной спецификой: ведь полноценный диплом по методике предполагает проведение уроков. Кроме того, важно, каким образом можно организовать по интернету взаимодействие с опытными педаго- гами и методистами, а также общение слушателей между собой. Поиски разработанных МООК по методике привели нас к специализации TESOL от Arizona State University [Dixon, Shewell, Haraway, Cinco 2016], состоя- щей из 5 МООК, которые можно проходить как целиком, так и каждый

отдельно.

Специализация TESOL. Part I. How to teach English является первой из двух, рассчитана на тех, у кого нет опыта преподавания английского языка и охватывает следующие темы: основы методики преподавания английского языка (Foundational Principles), теории овладения языком (Theories of Second Language Acquisition), планирование урока и контроль (Lesson Design and Assessment). Рассмотрим отдельные составляющие данной специализации.

*Изучение основного содержания* на курсе организовано посредством небольших видеолекций (около 5 мин), сгруппированных по неделям. В видео информация подается метафорично, рассматриваются примеры ведения занятий, которые тем не менее представляют собой актерски разыгранные видеофрагменты, в которых камера направлена на учителя, стоящего перед воображаемым классом. *Дополнительные материалы* представлены в текстовом формате .pdf и даются дозировано. Содержание курса разбито на максимально малые блоки, что позволяет не перегрузить слушателя.

*Консультации* организованы через форум, в котором активно уча- ствуют сами слушатели, что снимает нагрузку с тьюторов. Организация консультаций через форум делает курс гибким, не привязанным к опреде- ленному времени. И хотя курс не доступен в режиме ‘on demand’ и нужно записываться на определенную сессию, наличие форума технически позволяет сделать курс доступным в любое время.

*Вебинары* на курсе не представлены, что, с одной стороны, делает курс более гибким, а с другой — лишает слушателей возможности общения с авторами, максимально приближенного к реальному.

*Контроль* организован через небольшие тесты по видеолекциям (quizes), тесты — контрольные точки, проверяющие знания по несколь- ким блокам (check-points), письменные работы, проверяемые другими слушателями. Последний курс в специализации — это дипломный проект (capstone project), предполагающий видеозапись проведенных занятий и анализ видеофрагментов занятий других слушателей.

Оценивая курс в целом и рассматривая перспективу рекомендации данного курса студентам, изучающим методику преподавания иностран- ных языков, а также слушателям программ повышения квалификации, отметим несколько моментов.

Во-первых, безусловным достоинством курса является точность и кор- ректность используемой терминологии, освоение которой полезно как для российских студентов, так и для преподавателей. Например, в курсе рассматриваются гипотезы об усвоении языка: input hypothesis, affective ﬁlter и др.

Во-вторых, авторы предлагают такую методику планирования урока, которая предельно конкретна, но при этом открыта для педагогического творчества. В каждом уроке рассматриваются следующие части: warm up and objective discussion, instruction and modeling, guided practice and inde- pendent practice и assessment. В-третьих, за счет постоянной взаимной про- верки творческих работ расширяется кругозор слушателя, обогащаемый видением предмета представителями разных стран и культур.

С точки зрения перезачета дисциплины «Методика преподавания иностранных языков» при наличии сертификата данного курса возникают

вопросы, требующие внимательного рассмотрения. Главный из них — содержательный, поскольку, с одной стороны, вполне естественно то, что типичная рабочая программа дисциплины, ориентированная на россий- ский стандарт и российскую методическую школу, лишь частично соот- носится с представленным на международной платформе американским курсом. С другой стороны, рассматриваемый курс не всеобъемлющ, а, скорее, наоборот, очень конкретен и соответственно ограничен, т. е. возможно, что он мог бы стать заменой некоторой части дисциплины или рекомендован в качестве дополнительного источника информации.

**Литература**

* 1. *Андреев А.А.* (2013), Интернет-технологии и модели обучения в среде Интернет. М.: МИПК.
  2. *Бабанская О.М.* (2015), Механизмы включения МООК в образовательные программы высшего образования: опыт Томского государственного универси- тета. Развитие единой образовательной информационной среды: материалы XIV Международной научно-практической конференции, с. 92–94.
  3. *Романова Г.М., Апухтин А.В., Белоножкин Ю.Н.* (2011), Разработка когнитивно-продуктивной модели и методики реализации педагогики социального конструкционизма в области подготовки кадров для сферы туризма и гостепри- имства. Промежуточный отчет по НИР. Сочи.
  4. *Гейхман Л.К., Ставцева И.В.* (2015), Новые технологии в профессиональ- ном образовании: учебное пособие. Челябинск: Издательский центр ЮУрГУ.
  5. *Носкова А.* (2016), Онлайн образование или оффлайн: кто кого? URL: http:// edumarket.ru/library/coaching/31917/ (16.09.2016).
  6. *Тихомирова Е.* (2016), Живое обучение: Что такое e-learning и как заставить его работать. М.: Альпина Паблишер.
  7. *Daphne Koller* (2012), What we’re learning from online education. URL: http:// [www.ted.com/talks/daphne\_koller\_what\_we\_re\_learning\_from\_online\_education.html](http://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.html) (16.09.2016).
  8. *Dixon Sh., Shewell J., Haraway A.M., Cinco J.* Arizona State University MOOC “TESOL. Part I. How to teach English”. URL: [https://www.coursera.org/specializations/](http://www.coursera.org/specializations/) tesol (16.09.2016).
  9. *Manos E., McLaughlin C.* Arizona State University MOOC Lesson Planning with the ELL in Mind. URL: [https://www.coursera.org/learn/ell-lesson/home/welcome](http://www.coursera.org/learn/ell-lesson/home/welcome) (16.09.2016).

**Literature**

1. *Andreyev A.A.* (2013), Internet-tekhnologii i modeli obucheniya v srede Internet. M.: MIPK.
2. *Babanskaya O.M.* (2015), Mekhanizmy vklyucheniya MOOK v obrazovatel’nye programmy vysshego obrazovaniya: opyt Tomskogo gosudarstvennogo universiteta. Razvitie yedinoy obrazovatel’noy informatsionnoy sredy: materialy XIV Mezhdunar- odnoy nauchno-prakticheskoy konferentsii, s. 92–94.
3. *Romanova G.M., Apukhtin A.V., Belonozhkin Yu.N.* (2011), Razrabotka kogni- tivno-produktivnoy modeli i metodiki realizatsii pedagogiki sotsial’nogo konstruktsion- izma v oblasti podgotovki kadrov dlya sfery turizma i gostepriimstva. Promezhutochnyi otchet po NIR. Sochi.
4. *Geykhman L.K., Stavtseva I.V.* (2015), Novye tekhnologii v professional’nom obrazovanii: uchebnoe posobie. Chelyabinsk: Izdatel’skiy tsentr YuUrGU.
5. *Noskova A.* (2016), Onlayn obrazovanie ili ofﬂayn: kto kogo? URL: http:// edumarket.ru/library/coaching/31917/ (16.09.2016).
6. *Tikhomirova E.* (2016), Zhivoe obuchenie: Chto takoe e-learning i kak zastavit’ ego rabotat’. M.: Al’pina Pablisher.
7. *Daphne Koller* (2012), What we’re learning from online education. URL: http:// [www.ted.com/talks/daphne\_koller\_what\_we\_re\_learning\_from\_online\_education.html](http://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.html) (16.09.2016).
8. *Dixon Sh., Shewell J., Haraway A.M., Cinco J.* Arizona State University MOOC “TESOL. Part I. How to teach English”. URL: [https://www.coursera.org/specializations/](http://www.coursera.org/specializations/) tesol (16.09.2016).
9. *Manos E., McLaughlin C.* Arizona State University MOOC Lesson Planning with the ELL in Mind. URL: [https://www.coursera.org/learn/ell-lesson/home/welcome](http://www.coursera.org/learn/ell-lesson/home/welcome) (16.09.2016).

***Т.В. Знаменская, А.А. Пакудина, О.Н. Ровбо***

**СОЗДАНИЕ ДИСТАНЦИОННОГО КУРСА ДЛЯ ОБУЧЕНИЯ СПЕЦИАЛЬНОЙ ЛЕКСИКЕ**

**В УСЛОВИЯХ ВУЗА: АНАЛИЗ И ПОИСКИ РЕШЕНИЙ**

**Аннотация.** В статье рассматривается возможность создания дистанционного курса по железнодорожной специальности с использованием разнообразных интернет-ресурсов и аудиовизуальных материалов, таких как визуальный словарь- тезаурус, видео по специальности, а также мобильные приложения, позволяющие закреплять и актуализировать лексику по данной специальности.

**Ключевые слова.** Иноязычная компетенция, специальная лексика, обучаю- щий дистанционный курс, визуальный словарь.

***T.V. Znamenskaia, A.A. Pakudina, O.N. Rovbo***

**DEVELOPING A PROFESSIONAL VOCABULARY DISTANCE LEARNING COURSE:**

**ANALYSIS AND FINDING SOLUTIONS**

**Abstract.** The article deals with the development of a professional vocabulary dis- tance learning course for engineering students. Different ways of teaching professional vocabulary online are examined, visual thesaurus being one of them. The article also examines Internet resources and mobile apps helping to learn English vocabulary online. **Keywords.** Foreign language competence, professional vocabulary, distance learn-

ing course, visual thesaurus/dictionary.

Одним из важных условий конкурентного преимущества молодого специалиста является способность эффективно осуществлять профес- сиональную коммуникацию на иностранном языке, поскольку «умение высказаться на заданную тему по специальности на иностранном языке служит своеобразным показателем иноязычной компетенции, без кото- рого сегодня невозможно представить современного технического спе- циалиста» [Молодых-Нагаева и др. 2015: 277]. Необходимым условием успешного осуществления данной коммуникации является владение специальной лексикой в выбранной профессиональной сфере, поэтому на занятиях по иностранному языку в техническом вузе большое внимание

© ЗНАМЕНСКАЯ Татьяна Владимировна. — Znamenskaia Tatiana. Петербургский госу- дарственный университет путей сообщения Императора Александра I, Россия. — Emperor Alexander I St. Petersburg State Transport University, Russia. E-mail: [donotstopmenow@mail.ru](mailto:donotstopmenow@mail.ru)

© ПАКУДИНА Анна Александровна. — Pakudina Anna. Петербургский государственный университет путей сообщения Императора Александра I, Россия. — Emperor Alexander I St. Petersburg State Transport University, Russia. E-mail: [Pogoran1@gmail.com](mailto:Pogoran1@gmail.com)

© РОВБО Ольга Николаевна. — Rovbo Olga. Петербургский государственный университет путей сообщения Императора Александра I, Россия. — Emperor Alexander I St. Petersburg State Transport University, Russia. E-mail: [rvbolga@gmail.com](mailto:rvbolga@gmail.com)

должно уделяться работе, направленной на изучение специальной лексики и умение применить ее в собственном высказывании.

В Петербургском государственном университете путей сообщения Императора Александра I изучение специальной лексики осуществля- ется в процессе выполнения лексических упражнений, чтения, перевода и обсуждения аутентичных текстов на профессиональную тематику, работы с соответствующими аудио- и видеоматериалами и т. п. Однако количество часов, выделяемое на овладение учащимися иностранным языком в аудитории, не позволяет изучить специальную лексику в том объеме, который впоследствии понадобится молодому специалисту для успешной профессиональной коммуникации. Таким образом, большое значение приобретают самостоятельная работа учащегося по овладению учебным материалом и, следовательно, создание таких электронных учебных ресурсов, которые будут не только доступны в любое удоб- ное для учащегося время, но также просты в использовании, полезны и интересны. В настоящее время существует множество интернет-сайтов и мобильных приложений, позволяющих изучать лексику иностранного языка самостоятельно. Современные курсы иностранных языков пред- лагают своим студентам помимо аудиторных занятий разнообразные, разработанные в соответствии с программой курсов электронные ресурсы для отработки различных аспектов языка, немаловажной составляющей которых являются задания на закрепление изучаемой лексики, а также пополнение словарного запаса новыми лексическими единицами. В то же время электронных ресурсов, дающих возможность изучать специальную лексику в той или иной профессиональной сфере, либо очень мало, либо просто не существует, поэтому в Петербургском государственном уни- верситете путей сообщения Императора Александра I создание подоб- ных ресурсов осуществляется преподавателями кафедры «Иностранные языки» совместно с Центром интернет/интранет технологий в обучении (ЦИИТО), действующим уже более десяти лет. Так, были созданы элек- тронные материалы для изучения лексики по таким специальностям, как

«Логистика», «Подвижной состав железных дорог (Вагоны)» и «Эконо- мика предприятий и организаций (строительство)».

Создание курса по специальности — достаточно трудоемкая и дли- тельная работа, требующая от его составителей (лингвистов и про- граммистов) не только поиска материала, но и, в идеале, создания некой универсальной модели дистанционного курса с разнообразными зада- ниями и качественным наполнением. Перед авторами данной статьи была поставлена задача преобразовать имеющееся методическое пособие для первого курса по одной из востребованных специальностей «Подъемно- транспортные и строительные машины» в обучающий дистанционный

курс таким образом, чтобы получившийся продукт мог быть использован как основа для курсов по другим техническим специальностям, пред- ставленным в ПГУПС Императора Александра I. Одним из основных требований является использование разнообразных аудиовизуальных средств и интернет-ресурсов.

Недостатком существующих методических пособий и разработок по некоторым железнодорожным специальностям является однообразие заданий, сводящихся к чтению и переводу текстов, а также выполнению послетекстовых лексических и грамматических упражнений, в то время как заданий, направленных на развитие таких важных при изучении иностранного языка навыков, как аудирование и говорение, очень мало. Поэтому одной из задач при создании курса по вышеуказанной специ- альности будет подбор и включение в курс аудио- и видеоматериалов. Отметим, что использование традиционных аудиоматериалов не всегда достигает своей цели. Студенты лучше воспринимают видеоинформацию как более привычную и удобную, так как в современном мире основными средствами массовой информации являются интернет и телевидение. Основное преимущество видеоматериалов в сравнении с аудиоматериа- лами состоит в том, что это приближенная к естественной форма комму- никации, когда мы видим ситуацию общения и собеседника. Кроме того, при просмотре видеоматериалов мы воспринимаем не только вербальную, но и невербальную информацию, такую как жесты и мимика говорящего. Благодаря этому восприятие информации в формате видео выше, чем в формате аудио [Harmer 1998: 108].

Развитие навыка говорения у учащихся планируется осуществлять на основе метода обучения, применяемого в некоторых мобильных при- ложениях для изучения английского языка, например, в приложениях “ABA English” и “English Central”. Обучение основано на том, что, прослушав фразу на английском языке, учащийся должен ее повторить в диктофон на мобильном устройстве, а затем сравнить получившийся результат с оригиналом. Мобильное приложение “ABA English” также предлагает учащемуся самому стать участником диалога и реагировать на предлагаемые приложением реплики, используя только что отра- ботанные фразы. Идея создания подобного задания с использованием специальной лексики представляется очень заманчивой и интересной. Вместе с тем эту идею очень трудно воплотить в жизнь, поскольку найти подходящий для обучения видеоматериал по данной специаль- ности довольно непросто.

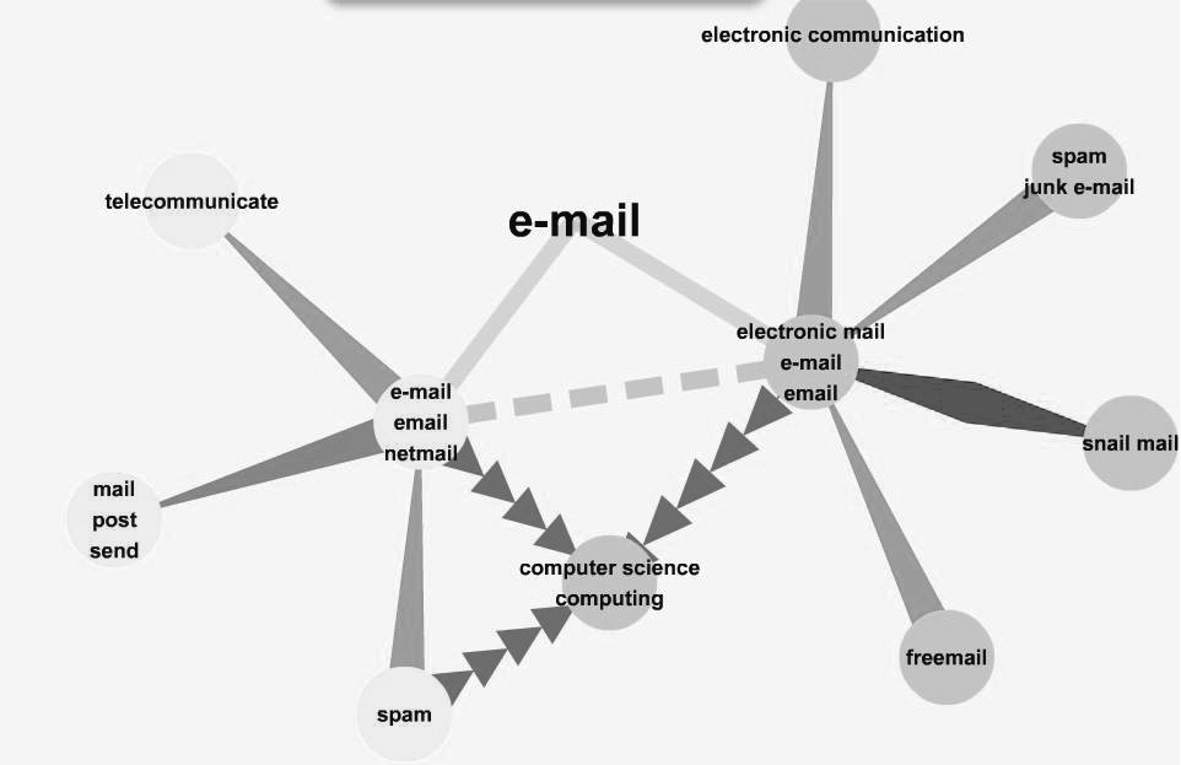
Для видеоматериала к электронному курсу по специальности

«Подъемно-транспортные машины» также планируется создать комплекс упражнений, направленных на введение, отработку и закрепление новой

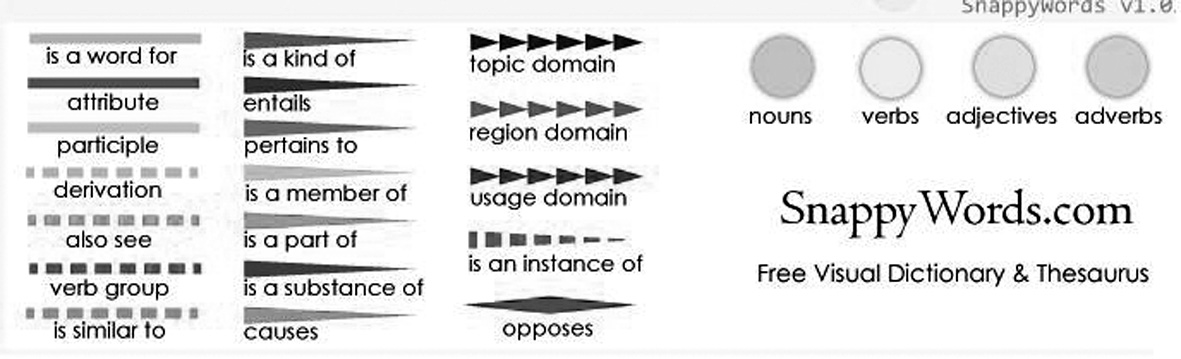
лексики. На первом этапе, до просмотра видеоматериала, вводится новая лексика. После просмотра видеоинформации необходимо проконтролиро- вать общее понимание содержания и закрепить новую лексику. Следую- щие задания можно использовать для этой цели:

* Match the synonyms.
* Match the antonyms.
* Translate the words.
* Find the Russian equivalents for the given words.
* Fill in the gaps.
* Др. стандартные лексические упражнения.

Кроме видеоматериалов основой курса станут также аутентичные тек- сты для чтения с большим количеством специальной лексики, взятые из журналов по данной предметной области. В одном из разработанных ранее дистанционных курсов «Подвижной состав железных дорог (Вагоны)» интерес представляет наличие визуального словаря. При наведении курсора на термин появляется «облачко» с переводом слова на русский язык, транскрипцией (также есть возможность услышать произношение слова) и фотографией объекта, скрывающегося за новым для учащегося термином. Таким образом, введение новой лексической единицы осущест- вляется тремя способами: визуально (соотнесение с понятием), переводом на родной язык, а также при помощи контекста.



*Рис 1.* Выдача запроса по слову ***e-mail***



*Рис. 2.* Обозначения, принятые для ресурса **“*Snappy Words visual English Dictionary”***

Тем не менее чтения текстов даже с очень хорошим визуальным сло- варем оказывается недостаточно для закрепления специальной лексики и дальнейшей ее актуализации. То или иное слово существует в связи с другими словами. Поэтому интересным представляется интернет- ресурс [www.snappywords.com,](http://www.snappywords.com/) визуализирующий парадигматические и (в меньшем объеме) синтагматические связи между словами. Данный ресурс представляет собой визуальный словарь и тезаурус. Он довольно прост в использовании. В поисковом окне печатается необходимое слово. При нажатии клавиши “Enter” на экране появляется схема взаимосвязей введенного слова с другими словами. Части речи обозначены различными цветами, а парадигматические отношения — с помощью разнообразных линий, соединяющих искомое слово с другими, а также другие слова друг с другом, что позволяет визуально увидеть синонимические, антонимиче- ские, гиперонимические и другие отношения.

Значение слова можно посмотреть, наведя на него курсор. Также можно расширить схему, дважды кликнув клавишей мыши по одному из «узлов». Таким образом, становится возможным увидеть еще больше связей.

Cоздание подобного словаря терминов по специальности и разработка учебных заданий, помогающих активировать эти связи, — одни из перво- степенных задач при создании курса.

Создание дистанционного курса по специальности обнаруживает ряд трудностей, связанных как с отбором материала для курса, так и с подбором и разработкой технической составляющей курса. На данный момент авторы статьи изучают существующие онлайн-курсы и интернет- ресурсы, обучающие лексике английского языка с целью выбрать наи- более эффективные учебные материалы, на базе которых можно соста- вить дистанционный курс, обучающий терминологии по технической специальности.

**Литература**

1. *Молодых-Нагаева Е.Г., Нордман И.Б.* (2015), Использование элементов различных образовательных парадигм при обучении профессиональной коммуни- кации на иностранном языке в вузе // Теория и практика общественного развития,

№ 21, с. 277–279.

1. *Harmer J.* (1998), How to Teach English. L.: Longman, p. 198.

**Literature**

1. *Molodykh-Nagaeva E.G., Nordman I.B*. (2015), Ispol’zovanie elementov razlichnykh obrazovatel’nykh paradigm pri obuchenii professional’noy kommunikatsii na inostrannom yazyke v vuze // Teoriya i praktika obshchestvennogo razvitiya, № 21, s. 277–279.
2. *Harmer J.* (1998), How to Teach English. L.: Longman, p. 198.

***А.В. Зубов***

**ОТ УЧЕБНОЙ ПРОГРАММЫ — К КОМПЬЮТЕРНОЙ ПРОГРАММЕ ОБУЧЕНИЯ ЯЗЫКАМ**

**Аннотация.** В статье описывается один из возможных подходов к процессу подключения педагогов-специалистов по иностранным языкам и созданию эффективных обучающих программ по иностранным языкам. Детально, на при- мере конкретной учебной программы «Практическая грамматика английского языка», раскрывается последовательность действий педагога-специалиста для того, чтобы выделить в этой учебной программе те разделы, темы, проблемы и задачи, в усвоении которых может быть использован компьютер. Показывается начало процесса создания обучающего сценария и указывается литература, где содержится детальное описание процесса создания обучающего сценария и сами обучающие сценарии.

**Ключевые слова.** Задачи, компьютер, обучающий, проблемы, программа, сценарий, темы.

***A.V. Zubov***

**FROM EDUCATIONAL PROGRAM — TO COMPUTER PROGRAM OF LANGUAGES TEACHING**

**Abstract.** The article describes one of possible approaches to the process of connecting educators in foreign languages to the creation of effective training programs in foreign languages. In detail, on the example of a speciﬁc educational program “Practical English Grammar”, reveals the sequences of teacher’s actions in order to derive from this teacher program those parts, themes, problems and tasks, in learning which can be used the computer. The author demonstrates how to build the training scenario for solution of each task and shows one of these scenarios and point to the scientiﬁc literatures where showed such scenarios.

**Key words.** Computer, problems, program, scenario, tasks, themes, training.

Стратегическая цель учебного предмета «Иностранный язык» заклю- чается в формировании «вторичной языковой личности» учащегося [Халеева 1990: 14–15], его способности осуществлять различные виды речевой деятельности (чтение, письмо, говорение, аудирование), языковые и речевые операции и действия на уровне фонетики, лексики и грамма- тики, готовности к межкультурному общению на базе усвоенных куль- туроведческих знаний и особенностей неречевого поведения носителей иностранного языка.

© ЗУБОВ Александр Васильевич. — Zubov Alexander. Минский государственный линг- вистический университет, Беларусь. — Minsk State Linguistic University, Belarus. E-mail: [proscien@msln.by](mailto:proscien@msln.by)

В вузах России и за рубежом накоплен богатый опыт компьютерного обучения иностранным языкам. Тем не менее многие преподаватели с осторожностью относятся к возможностям использования компьютер- ных средств обучения. Это связано с тем, что до сих пор:

1. не выработаны критерии полезности применения компьютеров непосредственно на занятиях для каждой возрастной группы обучаемых;
2. не налажено сотрудничество преподавателей с программистами, создателями обучающих средств по иностранным языкам;
3. не определены критерии оценки учебных программных средств;
4. нет системы подготовки преподавателей вузов, учителей и мето- дистов школ к проведению занятий с использованием компьютеров [Высоцкий 2000: 87].

В данной статье предложено решение первых двух из перечисленных задач.

Практика показывает, что при применении информационных обра- зовательных средств в вузах на первый план выдвигаются технические и организационные вопросы, а педагогическая подготовка преподавателей остается вне внимания руководителей. А неприятие преподавателями современных средств обучения и их психолого-педагогическая неподго- товленность к их использованию является главным тормозом широкого внедрения инноваций в учебный процесс [Казаченок и др. 2012: 5].

Все вышесказанное в полной мере относится к использованию инфор- мационных технологий в обучении как родных, так и иностранных языков, и факторы тревожности, связанные с ощущением незнания, неумения.

Педагогическая деятельность при использовании компьютеров в обучении языкам сильно отличается от традиционной, требует наличия специфических знаний, умений и навыков у преподавателя [Вечер и др. 2013: 16]. Для этого необходима специальная подготовка преподавателей иностранных языков [Вечер и др. 2013: 46; Доронина 1993: 73–76].

В данной статье предложен один из возможных подходов к процессу подключения педагогов-специалистов по иностранным языкам к созданию эффективных обучающих компьютерных программ [Zubov и др. 2014].

Пусть необходимо составить комплекс обучающих программ для усвоения знаний, предлагаемых учебной программой «Практическая грам- матика английского языка» [Программа педагогических институтов 1988]. Как показывает практика создания обучающих программ по иностранным языкам [Зубов 1989], для этого необходимо решить следующие 4 задачи:

1. провести проектирование состава выбранной учебной программы и выделить в ней разделы, темы, проблемы и задачи;
2. провести методическую проработку учебного материала каждой отдельной задачи, входящей в состав учебной программы и создать ее обучающий сценарий;
3. создать на основе обучающего сценария каждой отдельной задачи компьютерную обучающую программу и провести ее тестирование;
4. объединить созданные обучающие программы всех учебных задач в единый обучающий курс «Практическая грамматика английского языка». Первые две задачи могут быть выполнены опытным преподавателем, имеющим большой опыт преподавания английского языка и знакомым с методами использования компьютеров в обучении иностранным языкам. В процессе решения первой из приведенных выше задач прово- дится детальный анализ учебной программы «Практическая грамматика английского языка» и в ее составе выделяются *разделы, темы, проблемы*

и *задачи*.

Такой анализ позволит выделить в ней два *раздела*: «Морфология» и «Синтаксис».

Далее в составе каждого из выделенных разделов необходимо найти *темы*, в изучении которых эффективную помощь может оказать компью- тер.

В разделе «Морфология» рассматриваемой учебной программы выде- ляются следующие *темы*: имя существительное; глагол; имя прилагатель- ное; имя числительное; слова категории состояния; наречие; местоимение; модальные слова; частицы, предлоги; союзы; междометия.

В разделе «Синтаксис» таким же образом можно выделить такие *темы*: синтаксические связи слов; общие сведения о предложении; коммуникативные типы предложений; простое предложение; односо- ставное предложение; эллиптическое предложение; сложносочиненное предложение; сложноподчиненное предложение; чужая речь; пунктуа- ция.

Каждая из тем может быть представлена в виде последовательности определенных *проблем*. Так, в составе темы «Глагол» учебной программы

«Практическая грамматика английского языка» явно выделяются такие проблемы: «основные признаки и типы глаголов», «лицо и число глагола»,

«время и вид глагола», «наклонение», «модальные глаголы», «инфини- тив», «герундий», «причастие I», «причастие II».

Каждая из выделенных проблем может быть представлена в виде отдельной конкретной *задачи* или нескольких простых задач. В состав каждой задачи может входить как чисто теоретический материал, так и конкретные практические задания (упражнения, вопросы). Возможна организация структуры задач и на смешанной основе, когда теоретические вопросы чередуются с практическими заданиями по усвоению теорети- ческих правил.

Так, тема «Имя существительное» раздела «Морфология» учебной программы «Практическая грамматика английского языка» может быть представлена в виде следующей последовательности *проблем*:

* 1. образование множественного числа английского существительного;
  2. образование притяжательного падежа английского имени суще- ствительного;
  3. способы выражения рода английских имен существительных;
  4. основные функции артикля;
  5. особые функции артикля.

По каждой из этих проблем, представляющих в этом случае отдель- ные конкретные задачи, могут быть созданы компьютерные обучающие программы.

Продемонстрируем начало процесса создания обучающей программы по первой из перечисленных выше пяти задач «Образование множествен- ного числа английского существительного». Как было отмечено, прежде всего необходимо создать обучающий сценарий для такой обучающей про- граммы. Создание такого обучающего сценария требует от специалиста- педагога по английскому языку выполнения следующей последователь- ности действий [Зубов 2009: 20]:

* конкретизировать тип учебной задачи;
* определить параметры, по которым будет оцениваться результат обучения конкретной учебной задаче;
* разработать формальные процедуры оценки знаний по усвоению материалов задачи;
* подобрать учебный материал к задаче и определить порядок его предъявления;
* разработать модель действий обучаемого и компьютера при решении данной задачи и представить ее в виде обучающего сценария.

Детальное описание этих действий по отношению к задаче «Образо- вание множественного числа английских существительных» дано в работе [Зубов 2009: 20–31]. Здесь же в деталях раскрывается обучающий сцена- рий этой задачи [Зубов 2009: 34–38]. В этой же книге, на страницах 50–52, приводится фрагмент компьютерной обучающей программы, написанной на языке QBASIC по данному сценарию.

**Литература**

1. *Вечер Н.А., Гулевич Г.И., Дарошкевич В.П.* (2013), Образовательный про- цесс: методика, опыт, проблемы: сборник научно-методических статей, с. 44–46.
2. *Высоцкий О.Р.* (2000), Компьютер в образовании. Информатика и образо- вание, № 1, с. 87.
3. *Доронина О.В.* (1993), Страх перед компьютером: природа, профилактика, преодоление // Вопросы психологии, № 1, с. 68–78.
4. *Зубов А.В., Зубова И.И.* (2009), Методика применения информационных технологий в обучении иностранным языкам. М.
5. *Казаченок В.В., Фирас Таки А.* (2012), Особенности информатизации учебного процесса в современных условиях. Информатизация образования: еже- квартальный научно-методический журнал, № 2 (67), с. 3–13.
6. *Зубов А.В.* (1989), Методические указания по разработке и использованию обучающих программ по иностранным языкам. Ч. 1. Анализ учебной программы языковой дисциплины и выделение перечня учебных задач. Минск.
7. Программы педагогических институтов. Практическая грамматика англий- ского языка. Практическая фонетика английского языка. Для специальности

№ 2103 «Иностранные языки» (1988), № 13, с. 3–36.

1. *Халеева И.И.* (1990), Основы теории обучения пониманию иноязычной речи (подготовка переводчиков): автореф. дис. … д-ра пед. наук.
2. *Zubov A., Hajczuk R., Szymula R.* (2014), От учебной программы — к ком- пьютерной программе обучения языкам. Bialystok.

**Literature**

1. *Vecher N.A., Gulevich G.I., Daroshkevich V.P*. (2013), Obrazovatel’nyi protsess: metodika, opyt, problemy: sbornik nauchno-metodicheskikh statey, s. 44–46.
2. *Vysotskiy O.R*. (2000), Komp’yuter v obrazovanii. Informatika i obrazovanie,

№ 1, s. 87.

1. *Doronina O.V*. (1993), Strakh pered komp’yuterom: priroda, proﬁlaktika, pre- odolenie // Voprosy psikhologii, № 1, s. 68–78.
2. *Zubov A.V., Zubova I.I*. (2009), Metodika primeneniya informatsionnykh tekhnologiy v obuchenii inostrannym yazykam. M.
3. *Kazachenok V.V., Firas Taki A*. (2012), Osobennosti informatizatsii uchebnogo protsessa v sovremennykh usloviyakh. Informatizatsiya obrazovaniya: yezhekvartal’nyi nauchno-metodicheskiy zhurnal, № 2 (67), s. 3–13.
4. *Zubov A.V*. (1989), Metodicheskie ukazaniya po razrabotke i ispol’zovaniyu obuchayushchikh programm po inostrannym yazykam. Ch. 1. Analiz uchebnoy pro- grammy yazykovoy distsipliny i vydelenie perechnya uchebnykh zadach. Minsk.
5. Programmy pedagogicheskikh institutov. Prakticheskaya grammatika angli- yskogo yazyka. Prakticheskaya fonetika angliyskogo yazyka. Dlya spetsial’nosti

№ 2103 «Inostrannye yazyki» (1988), № 13, s. 3–36.

1. *Khaleyeva I.I*. (1990), Osnovy teorii obucheniya ponimaniyu inoyazychnoy rechi (podgotovka perevodchikov): avtoref. dis. … d-ra ped. nauk.
2. *Zubov A., Hajczuk R., Szymula R*. (2014), Ot uchebnoy programmy — k komp’yuternoy programme obucheniya yazykam. Bialystok.

***Л.И. Корнеева***

**ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ В ОБУЧЕНИИ ПЕРЕВОДЧИКОВ**

**Аннотация.** В статье исследованы современные возможности обучения лингвистов-переводчиков в области письменного перевода с использованием информационно-коммуникационных технологий. Отмечается особая важность сформированности информационно-технологической компетенции в профессио- нальной подготовке переводчика. Информационно-технологические навыки пере- водчика структурированы в статье в зависимости от их функционального назна- чения и подразделяются на лингвистические, коммуникативные навыки, навыки ведения информационного поиска и др. Также оцениваются возможности исполь- зования в обучении письменному переводу национальных электронных корпусов текстов, учебного интернет-форума.

**Ключевые слова.** Информационно-коммуникационные технологии, информационно-технологическая компетенция переводчика, информационно- технологические навыки, национальные электронные корпуса текстов, учебный интернет-форум.

***L.I. Korneeva***

**INFORMATION AND COMMUNICATION TECHNOLOGIES FOR TRAINING WOULD-BE TRANSLATORS**

**Abstract.** The article explores modern opportunities of using information and communication technologies in training language experts in translation. Reference is made to the critical importance of fully formed information and technological competency in the professional training of a translator. The article classiﬁes information and technological skills of a translator depending on their functions, and divides them into linguistic skills, communication skills, data retrieval skills etc. Assessment is also made of the opportunities of using national electronic text corpuses and educational Internet forums in teaching written translation.

**Keywords.** Information and communication technologies, information and technological competency of a translator, information and technological skills, national electronic text corpuses, educational Internet forum.

За последние 10–15 лет характер работы переводчика и требования к нему существенно изменились. В первую очередь изменения коснулись письменного перевода научно-технической, официальной и деловой доку- ментации. Сегодня, как правило, уже недостаточно просто перевести текст, пользуясь компьютером как пишущей машинкой. Заказчик ожидает от

© КОРНЕЕВА Лариса Ивановна. — Korneeva Larissa. Уральский федеральный университет, Екатеринбург, Россия. — Ural Federal University, Ekaterinburg, Russia. E-mail: [lorakorn@list.ru](mailto:lorakorn@list.ru)

переводчика, что оформление готового документа будет соответствовать внешнему виду оригинала настолько точно, насколько это возможно, и при этом удовлетворять принятым в данной стране стандартам.

Эти жесткие условия можно соблюсти лишь в том случае, если пере- водчик владеет современными информационно-коммуникационными технологиями. Арсенал письменного переводчика постоянно пополняется электронными инструментами (целая серия компьютерных программ и электронных словарей, национальных электронных корпусов текстов), которые облегчают поиск необходимой информации и соответственно сам процесс перевода, а также позволяют существенно повысить качество перевода.

К сожалению, учебному курсу «Технологии автоматизированного перевода» не уделяется достаточно времени в процессе подготовки письменных переводчиков. В результате выпускники вступают в мир письменного перевода, имея лишь весьма приблизительное представ- ление о тех возможностях, которые дают переводчику информационно- коммуникационные технологии, электронные инструменты.

Традиционно применительно к деятельности переводчика выделяют следующий ряд компетенций:

1. коммуникативная (лингвистическая, прагматическая и социолинг- вистическая);
2. специальная (базовая, предметная, дискурсивная, социокультурная, технологическая, стратегическая);
3. социальная;
4. личностная [Гребенщикова 2005].

Однако исследователи перевода и профессиональные переводчики отмечают, что современный переводчик должен обладать обширными знаниями в области информатики, знать основы и перспективы развития новых технологий, иметь практические навыки и умения использования информационно-компьютерных технологий в переводческой деятельно- сти, уметь совершать оптимальный выбор информационно-компьютерных технологий при принятии переводческих решений и т. д. Исходя из этого, в компетентностную модель переводчика, помимо вышеперечисленных компетенций, необходимо включить ***информационно-технологическую составляющую (компетенцию)****.*

Под информационно-технологической компетенцией переводчика мы понимаем готовность и способность переводчика использовать при переводе современные информационно-коммуникационные технологии, профессиональные знания (теоретические, процедурные, интегративные), навыки и умения (поисково-ориентировочные, аналитические, инстру- ментальные) в области современных средств связи.

К сожалению, хотя исследователи отмечают значимость информаци- онно-технологической компетенции переводчиков [Инютин 2006], общая система ее формирования у студентов-переводчиков исследуется фраг- ментарно.

Таким образом, формирование и развитие информационно-техноло- гической компетенции как у преподавателя перевода, так и у обучаю- щихся остается по-прежнему актуальной проблемой. Главные факторы, препятствующие внедрению современных технологий в обучение пере- воду, — психологическая неготовность самих преподавателей перевода и отсутствие в вузах современного программного обеспечения.

Информационно-технологическая компетенция переводчика пред- полагает владение соответствующими техническими средствами и в первую очередь электронными. Речь идет о том, что специалисты называют полуавтоматизированным или человеко-машинным перево- дом (computer-assisted translation, CAT). Автоматизированный перевод (*англ*. computer-assisted translation, computer-aided translation, CAT) — это вид перевода, когда переводчик переводит тексты с помощью компьютерных программ, предназначенных для того, чтобы облегчить процесс перевода.

Информационно-технологические навыки переводчика, необходимые для функционирования стандартного рабочего места переводчика, подраз- деляются в зависимости от их функционального назначения следующим образом:

* 1. навыки лингвистического, а точнее переводческого характера, пред- полагающие использование тезауруса, электронных словарей;
  2. коммуникативные навыки, связанные с приемом и передачей инфор- мации (получение и отправление сообщений по электронной почте, пере- дача файлов, участие в работе форумов и конференций);
  3. навык ведения информационного поиска с помощью электронных словарей, энциклопедий, поисковых машин в сети интернет, виртуальных библиотек и архивов, терминологических банков и т. д.;
  4. навыки, необходимые для повышения продуктивности и эффектив- ности работы (сканер, интерактивное голосовое меню, текстовый редактор и др.) [Шевчук 2010].

Информационно-компьютерные ресурсы переводчика условно можно разделить на две группы:

1. Альтернативные ресурсы, созданные в электронном и бумажном формате. Это — словари, глоссарии, энциклопедии, справочники, онлай- новые газетные и журнальные архивы, которые имеют альтернативу в виде бумажных носителей информации, но, в отличие от последних, позволяют быстрее вести поиск регулярно обновляемой информации.
2. Безальтернативные ресурсы — системы класса Translation Memory

«Переводческая память»; ресурсы машинного перевода, электронные корпуса текстов, терминологические базы данных и многие другие. Эти ресурсы не имеют альтернативы в виде печатных изданий, и поэтому их польза для целей перевода бесспорна и не вызывает сомнений.

В работе письменного переводчика, особенно при переводе с рус- ского языка на иностранный, важную роль играют электронные корпуса текстов. Большинство ошибок, которые допускает переводчик, являются не столько грамматическими или синтаксическими, сколько связанными с неправильным выбором конструкции, комбинации слов в предложе- нии или с незнанием норм сочетаемости отдельных слов в контексте. Начинающий переводчик, в силу отсутствия у него языковой интуиции, копирует синтаксические конструкции и словосочетания родного языка в тексте перевода (межъязыковая интерференция).

К сожалению, имеющиеся бумажные словари мало помогают прове- рить то или иное интуитивное предположение о приемлемости и неприем- лемости конкретной конструкции/словосочетания в тексте перевода в силу недостаточной репрезентативности выборки и ограниченности сведений о комбинаторных возможностях иностранной лексики. Однако указанная проблема решается при обращении к корпусу текстов иностранного языка, представленному в электронном виде, и автоматическому поиску данного словосочетания в корпусе.

Национальный корпус — это собрание текстов в электронной форме, представляющих данный язык на определенном этапе его существо- вания [Соснина 2005: 85]. Обязательное требование к национальному корпусу — его представительность. Это означает, что корпус должен содержать по возможности все типы письменных и устных текстов, представленных в данном языке (художественные тексты разных жанров, публицистические, учебные, научные, деловые, разговорные, диалект- ные и т. п.), и что все эти тексты должны входить в корпус по возмож- ности пропорционально их доле в языке соответствующего периода. Следует иметь в виду, что хорошая представительность достигается только при значительном объеме корпуса (десятки и сотни миллионов словоупотреблений).

Существующие национальные корпуса составлены с разной степенью полноты и обработанности. Общепризнанным образцом является, в част- ности, Британский национальный корпус (100 млн слов), на который ориентированы и многие другие корпуса. Тематика и жанровая направлен- ность текстов весьма разнообразна. Это газетные и журнальные статьи, художественная и научная литература, брошюры, доклады, а также записи бытовых разговоров, радиопередач, интервью, дискуссий и совещаний.

На материале текстовой выборки указанного типа составлен целый ряд компьютерных корпусов, например:

* корпуса The Bank of English (150 млн слов), British National Corpus (100 млн слов), American National Corpus (22 млн слов), The Australian Corpus of English (1 млн слов), Brown Corpus of American English (более 1 млн слов), The International Corpus of English (600 000 слов);
* Национальный корпус русского языка (более 20 млн слов);
* Чешский национальный корпус (100 млн слов);
* The Finnish Translation Corpus (7 млн слов) — корпус, составленный на материале финского языка в Савонлиннской школе перевода при Уни- верситете города Иоэнсу;
* корпуса немецкого языка IDS Korpora и IMS Stuttgart и другие.

Поиск лингвистической информации в корпусе осуществляется с помощью программы-конкордансера, которая выстраивает соответ- ствующий конкорданс (вертикальный список случаев употребления слова, расположенных в алфавитном порядке в электронном корпусе текстов, слово подается вместе с его левым и правым окружением) [Шевчук 2010]. Подобные электронные средства, быстро сканируя гигантский массив текстов в поисках тех или иных сочетаний слов и возможного сокращения слова, дают достаточно точные сведения об общепринятом контекстуаль- ном словоупотреблении и подсказывают переводчику правильную синтак- сическую конструкцию на иностранном языке. Они содержат огромное количество примеров, иллюстрирующих комбинаторные возможности того или иного слова в современном языке и показывающих его воз- можное место в предложении и прочие сведения. На наш взгляд, трудно найти какие-либо иные более эффективные ресурсы, которые могли бы помочь переводчику столь же быстро и надежно проверить свои догадки

* + сочетаемости данного слова.

Компьютерные корпуса текстов представляют большой научный инте- рес, так как они дают возможность проводить по ходу перевода экспери- ментальное исследование, с тем чтобы выявить, какие слова, выражения, грамматические конструкции и типы дискурса действительно использу- ются в настоящее время носителями языка, как часто и для каких целей. Важную роль при обучении письменному переводу играет учебный интернет-форум. Учебный интернет-форум — это инструмент для дис- танционного обмена учебной информацией между студентами и препо- давателями, успешно применяемый для организации самостоятельной работы студентов. Модератором учебного форума выступает, как правило, преподаватель практики перевода, регулярно «выкладывающий» задания и упражнения по письменному переводу. Выполнение заданий и обратная

связь с преподавателем являются обязательными для студентов.

Обобщая изложенное выше, можно сделать следующий вывод. Работа переводчика в современных условиях невозможна без использо- вания информационно-коммуникационных технологий, интерактивных методик и работы с электронными ресурсами. Соответственно, для каче- ственной подготовки лингвистов-переводчиков требуют своего решения следующие проблемы: оснащение выпускающих кафедр необходимым оборудованием; повышение квалификации преподавателей переводческих дисциплин в области современных информационно-коммуникационных технологий; разработка методов и методик обучения переводчиков с использованием инновационных технологий; обеспечение преподавателя перевода необходимой учебно-методической литературой.

**Литература**

1. *Гребенщикова А.В*. (2005), Формирование профессиональной компетент- ности будущих переводчиков средствами информационно-коммуникационных технологий: дис. … канд. пед. наук. Челябинск.
2. *Инютин Н.Г.* (2006), Формирование информационно-технологической компетенции будущего «переводчика в сфере профессиональной коммуникации»: дис. … канд. пед. наук. Новгород, 320 с.
3. *Соснина Е.П.* (2005), Параллельные корпусы и обучение переводу. Про- блемы обучения переводу в языковом вузе. М.: МГЛУ, с. 85–92.
4. *Шевчук В.Н.* (2010), Электронные ресурсы переводчика: справочные мате- риалы для начинающего переводчика. М.: Либрайт, 136 с.

**Literature**

1. *Grebenshchikova A.V*. (2005), Formirovanie professional’noy kompetent- nosti budushchikh perevodchikov sredstvami informatsionno-kommunikatsionnykh tekhnologiy: dis. … kand. ped. nauk. Chelyabinsk.
2. *Inyutin N.G.* (2006), Formirovanie informatsionno-tekhnologicheskoy kompe- tentsii budushchego «perevodchika v sfere professional’noy kommunikatsii»: dis. … kand. ped. nauk. Novgorod, 320 s.
3. *Sosnina E.P*. (2005), Parallel’nye korpusy i obuchenie perevodu. Problemy obucheniya perevodu v yazykovom vuze. M.: MGLU, s. 85–92.
4. *Shevchuk V.N*. (2010), Elektronnye resursy perevodchika: spravochnye materialy dlya nachinayushchego perevodchika. M.: Librayt, 136 s.

***А.А. Коханик***

**АНАЛИЗ САЙТОВ ДЛЯ ИЗУЧЕНИЯ**

**ИНОСТРАННОГО ЯЗЫКА ПРИ РАЗВИТИИ ЛЕКСИЧЕСКИХ НАВЫКОВ У ОБУЧАЕМЫХ**

**Аннотация.** В связи с растущей популярностью информационно- коммуникационных технологий (ИКТ) их внедрение в процесс обучения ино- странному языку является закономерным процессом. Однако следует помнить о том, что использование любых технологий должно отвечать целям и задачам занятия. В докладе представлены результаты анализа сайтов для изучения ино- странного языка по нескольким параметрам, выделенным в ходе опроса препо- давателей. Кроме того, даны методические рекомендации по возможным формам работы с проанализированными сайтами.

**Ключевые слова.** Лексические навыки, сайты для изучения иностранного языка, параметры, анализ.

***A.A. Kokhanik***

## ANALYSIS OF WEBSITES FOR LEARNING A FOREIGN LANGUAGE WHILE DEVELOPING LEARNERS’ LEXICAL SKILLS

**Abstract.** Due to the growing popularity of web 2.0 technologies their inclusion into the process of teaching foreign languages is a natural consequence. However, one should remember that the use of any technology must correspond to the aims of the lesson. The report states the results of the analysis of websites for learning a foreign lan- guage according to several parameters singled out after questioning teachers. Moreover, methodic recommendations on possible forms of using the analysed websites are given. **Keywords.** Lexical skills, websites for learning a foreign language, parameters,

analysis.

Формирование лексических навыков является неотъемлемой частью обучения иностранным языкам. Многие ученые в своих работах отме- чают важность этого языкового аспекта [Пассов и др. 2010; Пассов и др. 2002; Соловова 2006; Шамов 2007: 19–21; Шатилов 1986]. Для того чтобы сделать процесс обучения лексике более эффективным и увлекательным, мы считаем целесообразным использовать на занятиях информационно- коммуникационные технологии (ИКТ), в частности, сайты для изучения иностранного языка (мы сфокусировали внимание на английском языке), так как они отличаются общей доступностью, красочностью, высокой интерактивностью и разнообразием форм и упражнений.

© КОХАНИК Анастасия Андреевна. — Kokhanik Anastasiia. Национальный исследова- тельский университет Высшая школа экономики, Россия. — National Research University Higher School of Economics, Russia. E-mail: [anakokhanik@gmail.com](mailto:anakokhanik@gmail.com)

Целью нашего анализа сайтов для изучения английского языка явля- ется определение этапа, на котором можно эффективно использовать тот или иной ресурс, а задача состоит в выделении плюсов и минусов анализируемых ресурсов. Для того чтобы отобрать материал для анализа, мы провели опрос среди молодых преподавателей и магистрантов, кото- рые проходят курс «Теория и методика преподавания»1. По результатам данного опроса, которые представлены в нашей статье [Коханик 2016: 152–155], мы выделили 5 сайтов и 8 параметров, согласно которым следует отбирать ресурсы для занятий, по мнению респондентов.

В данном докладе представлены результаты анализа сайтов:

1. ESL Gold 2 (85%);
2. Better English 3 (77%);
3. English Wizz 4 (77%).

Мы руководствовались следующими параметрами:

* + - удобство использования (понятный интерфейс, четкие инструкции);
    - тематическое разнообразие;
    - классификация по темам;
    - отработка лексики на разных уровнях;
    - наличие разных видов упражнений;
    - контроль (как он осуществляется);
    - соответствие заданий уровню владения ИЯ учащимся;
    - возможность преподавателю заранее изучить лексическое наполне- ние упражнений.

При анализе сайта под номером 1 ESL Gold были рассмотрены разделы “Words and Phrases” уровня “Low Intermediate — Words and Phrases” (1), High Intermediate (2), “High Intermediate — Quizzes from The Internet TESL Journal” (3). Данные уровни языковой подготовки были выбраны после проведенного опроса [Коханик 2016: 152–155].

Первый раздел “Low Intermediate — Words and Phrases” предлагает упражнения по совмещению слов с их дефинициями на основе лек- сических единиц, выбранных из небольшого текста. Положительным моментом является то, что слова представлены в контексте. Тем не менее отработка происходит лишь на уровне одной лексической единицы.

Во втором разделе представлено 10 упражнений на сопоставление слова и дефиниции или слов с их синонимами, а также упражнения на выделение лишнего слова из предложенного ряда. Таким образом, упражнения довольно однообразны и отличаются только по лексиче- скому наполнению. Их не следует выполнять одно за другим, однако

1 <http://goo.gl/forms/qSk5vM1wfj>

2 <http://www.eslgold.com/vocabulary/high_intermediate.html>

3 <http://www.better-english.com/vocabulary.htm>

4 <http://www.latcomm.com/englishwizz/>

их можно использовать на этапе контроля, разделив упражнения между обучаемыми.

Большим разнообразием видов упражнений отличается раздел “High Intermediate — Quizzes from The Internet TESL Journal”, который состав- лен из заданий, предложенных пользователями сайта. Здесь можно найти как вопросы множественного выбора, так и упражнения на совмещение синонимов/антонимов и т. п. Таким образом, сложность заданий немного растет с повышением уровня, однако отработка лексических навыков остается на уровне слова.

Стоит отметить, что инструкции по выполнению заданий даны доста- точно четко, однако в целом не хватает интерактивности, в отдельных слу- чаях ответы приходится записывать отдельно, так как их нельзя выбрать на экране. В результате верные ответы выводятся списком при нажатии на соответствующую ссылку или даются ниже. На этапе тренировки было бы более целесообразно лишь указать, где обучаемые ошиблись, чтобы дать им возможность еще раз подумать над данными вопросами, однако на этапе проверки знаний предложенный вариант контроля может подойти.

Несмотря на то что упражнения на сайте расклассифицированы по тематическим блокам, само тематическое разнообразие недостаточно богато. Например, во втором разделе дана лишь академическая лексика. В разделе “High Intermediate — Quizzes from The Internet TESL Journal” представлено больше тем (например, “Beverages”, “Plant or Animal”, “Clothes”), однако иногда названия упражнений даются по типу задания, например “Matching Quiz”, “Odd Man Out” и т. д. Отсутствие единого принципа наименования затрудняет работу с сайтом.

Тем не менее задания соответствуют уровню владения иностранным языком учащимися и у преподавателя есть возможность заранее изучить лексическое наполнение упражнений. Данный ресурс целесообразно использовать на этапе контроля полученных знаний, при условии, что лексический материал соответствует пройденному. А к разделу “Low Intermediate — Words and Phrases” можно обратиться при введении новой лексики, так как слова здесь даны в контексте определенной ситуации.

Второй сайт в нашем списке Better English довольно удобен в исполь- зовании, инструкции даны точно и ясно. Несмотря на то, что лексический раздел носит название Business English Vocabulary Exercises, в упражне- ниях представлена не только деловая лексика, но и общеупотребительная (Travel, Happy Idioms, Abbreviations, Jokes и т. д.), а все задания распреде- лены по темам. Положительным моментом является и то, что отработка идет на разных уровнях: от слов и фразовых глаголов до идиом, словосо- четаний и выражений.

Тем не менее виды упражнений не отличаются разнообразием. Пре- имущественно представлены закрытые вопросы множественного выбора. Также есть задания по заполнению пропусков в предложениях, но и здесь даны варианты ответов.

Контроль осуществляется интерактивно: рядом с выбранным ответом появляется либо YES/, либо NO/, сверху на экране показан процент верного выполнения задания, а при каждой ошибке он снижается. Если обучаемый выбрал неверный ответ, у него есть возможность подумать еще раз и выбрать другую опцию. В некоторых упражнениях (напр., Travel) обучаемый сначала проставляет свои варианты, а потом смо- трит процент верных и неверных ответов, но и здесь есть возможность исправить свои ошибки. Такая форма контроля довольно удобна и ори- ентирована на ученика, так как реакция на ответ дается своевременно. Кроме того, в некоторых разделах вместе с верным ответом также появ- ляется короткое объяснение. Например, в разделе Jokes при проверке ответа “Terrible” на вопрос “My dog has no nose” “How does it smell?” на экране появляется сноска: «smell» can mean the action of smelling or the smell it makes. Однако при выполнении определенных заданий (Business Expressions 1, 2, 3) нет возможности интерактивно выбрать вариант, так как присутствует только клавиша Answer, которая сразу показывает правильный ответ.

На сайте не заявлен необходимый уровень языковой подготовки уча-

щихся и задания предложены разного уровня, от Intermediate до Advanced. Таким образом, учитель должен заранее изучить упражнения, чтобы понять, подходят ли они уровню определенной группы учащихся. Сделать это не составляет трудностей, так как лексическое наполнение не меняется и есть возможность заранее его проанализировать.

Стоит отметить, что к данному сайту можно прибегнуть на этапе контроля, если лексический материал упражнений соответствует прой- денному, или на этапе тренировки, например, перед выполнением какого- либо коммуникативного задания, чтобы вспомнить пройденный материал. На данных этапах обучаемые уже знакомы с лексикой, и работу с упраж- нениями на сайте можно представить как соревнование, кто быстрее выполнит задание, у кого процент верных вариантов окажется наиболее высоким и т. д. Можно дать упражнения с сайта и в начале изучения определенной темы, чтобы ученики выписывали интересные выражения, но есть опасность, что при большом количестве незнакомых выражений учащиеся растеряются и их мотивация понизится.

Сайт под номером 3 English Wizz, оформленный в виде игры- викторины, отличается красочным оформлением и понятным интер- фейсом. Перед началом игры ведущий объясняет правила и дает четкие инструкции. В игре представлено 6 тематических категорий, каждая из

которых содержит по 4 вопроса разной сложности. Лексика отрабатыва- ется на разных уровнях, от слов и фразовых глаголов до коллокаций. Тем не менее в игре представлен только один вид упражнений — вопросы множественного выбора (multiple choice questions), минус которых состоит в том, что учащийся может выбрать правильный вариант наугад. Верный ответ показывается сразу, что является плюсом, так как подобная инфор- мация актуальна для учащегося. Кроме того, за положительный ответ даются очки, за неверный «наказывают» напарника по игре. Такой подход повышает мотивацию обучаемых.

В самом начале можно выбрать уровень языковой подготовки: Beginner, Intermediate или Advanced. Однако вопросы, например, на уровне Intermediate достаточно сложные и скорее подходят для Upper-Intermediate или Advanced. Большим минусом данного сайта является его непред- сказуемость — при каждом открытии игры категории и вопросы даются в произвольном порядке, и у преподавателя нет возможности заранее изучить лексическое наполнение упражнений.

В целом следует отметить, что данный сайт направлен на общее раз- витие обучаемого, его ориентирование в разных тематических блоках. Этот ресурс можно использовать в конце прохождения тематического раздела, либо в начале, в качестве разминки, или же дать учащимся уста- новку выписать лексические единицы, которые показались им интерес- ными. Более того, повышенная сложность может мотивировать учащихся на дальнейшее обучение и развивать в них стремление узнавать новые лексические явления.

В заключение отметим, что, несмотря на достоинства (удобство в использовании, четкие инструкции, классификация по темам) и недо- статки (однообразие упражнений) каждого из ресурсов, при грамотном планировании занятия работа с сайтами будет помогать преподавателю при формировании лексических навыков учащихся. Кроме того, при работе с ИКТ у обучаемых повысится мотивация, что будет способство- вать лучшему усвоению материала.

**Литература**

* 1. *Коханик А.А.* (2016), Определение популярности информационно- коммуникационных технологий при развитии лексических навыков. Язык: кате- гории, функции, речевое действие: материалы девятой научной конференции с международным участием, 14–15 апреля 2016 г., Москва—Коломна, вып. 9, ч. II, с. 152–155.
  2. *Пассов Е.И., Кузовлева Н.Е.* (2010), Основы коммуникативной теории и технологии иноязычного образования. М.: Русский язык, 568 с.
  3. *Пассов Е.И., Кузнецова Е.С.* (2002), Формирование лексических навыков: учебное пособие. Воронеж: НОУ «Интерлингва», 40 с.
  4. *Соловова Е.Н.* (2006), Методика обучения иностранным языкам. Базовый курс лекций. Пособие для студентов педагогических вузов и учителей. 4-е изд. М.: Просвещение, 239 с.
  5. *Шамов А.Н.* (2007), Лексические навыки устной речи и чтения — основа семантической компетенции обучаемых: Иностранные языки в школе, № 4, с. 19–21.
  6. *Шатилов С.Ф.* (1986), Методика обучения немецкому языку в средней школе: учеб. пособие для студ. пед. ин-тов. М.: Просвещение, с. 20.

**Literature**

1. *Kokhanik A.A.* (2016), Opredelenie populyarnosti informatsionno-kommunikat- sionnykh tekhnologiy pri razvitii leksicheskikh navykov. Yazyk: kategorii, funktsii, rechevoe deystvie: materialy devyatoy nauchnoy konferentsii s mezhdunarodnym uchastiem, 14–15 aprelya 2016 g., Moskva—Kolomna, vyp. 9, ch. II, s. 152–155.
2. *Passov E.I., Kuzovleva N.E*. (2010), Osnovy kommunikativnoy teorii i tekhnologii inoyazychnogo obrazovaniya. M.: Russkiy yazyk, 568 s.
3. *Passov E.I., Kuznetsova E.S*. (2002), Formirovanie leksicheskikh navykov: uchebnoe posobie. Voronezh: NOU «Interlingva», 40 s.
4. *Solovova E.N.* (2006), Metodika obucheniya inostrannym yazykam. Bazovyi kurs lektsiy. Posobie dlya studentov pedagogicheskikh vuzov i uchiteley. 4-e izd. M.: Prosveshchenie, 239 s.
5. *Shamov A.N.* (2007), Leksicheskie navyki ustnoy rechi i chteniya — osnova semanticheskoy kompetentsii obuchaemykh: Inostrannye yazyki v shkole, № 4, s. 19–21.
6. *Shatilov S.F*. (1986), Metodika obucheniya nemetskomu yazyku v sredney shkole: ucheb. posobie dlya stud. ped. in-tov. M.: Prosveshchenie, s. 20.

***С.А. Кучина***

**ЭЛЕКТРОННАЯ ХУДОЖЕСТВЕННАЯ ЛИТЕРАТУРА: ЛИНГВОДИДАКТИЧЕСКИЙ АСПЕКТ**

**Аннотация.** В статье рассматривается электронное интерактивное литератур- ное произведение А. Кэмпбелл “Nightingale’s Playground” в аспекте его интегра- ции в программу обучения по иностранному (английскому) языку; представлены различные типы языковых упражнений, задания на тренировку навыков чтения, перевода и письма. Электронное литературное художественное произведение “Nightingale’s Playground” А. Кэмпбелл может быть адаптировано и интегрировано в учебный процесс не только как средство тренировки устных и письменных видов речевой деятельности на английском языке, но и в качестве ресурса, способствую- щего формированию у обучающихся навыков медиаграмотности.

**Ключевые слова.** Электронная литература, интерактивность, медиаграмот- ность, нарратив.

***S.A. Kuchina***

**ELECTRONIC LITERATURE: LINGUODIDACTICAL ASPECT**

**Abstract.** The article deals with the electronic interactive novel “Nightingale’s Playground” by A. Campbell and the key aspects of its integration in learning English practice. There is a set of authoring language and translation exercises, the reading com- prehension tasks and the assignments aimed at improving writing skills in the article. It is possible to use “Nightingale’s Playground” by A. Campbell not only as a resource for training writing and speech skills but also as a tool for improving digital literacy.

**Keywords.** Electronic literature, interactivity, digital literacy, narrative.

Электронная литература (electronic literature / eliterature) представляет собой художественные нарративы, берущие свое начало в электронной информационной среде. «Специфическими механизмами создания, функционирования и обработки данных в электронном тексте считаются самопорождающие модели текста, мультимедийные объекты, анимация, интерактивные процессы (активное участие читателя в процессе написа- ния или обсуждения готового электронного литературного произведения)» [Кучина 2015: 119]. Принципиально новый синтетический формат худо- жественного текста представляет собой электронный нарратив с интегри- рованными аудио, видео, анимационными и графическими элементами. Использование паралингвистических компонентов в электронном тексте придает ему зрительную и чувственную наглядность и является неисся-

© КУЧИНА Светлана Анатольевна. — Kuchina Svetlana. Новосибирский государственный технический университет, Россия. — Novosibirsk State Technical University, Russia. E-mail: [svkuchina@yandex.ru](mailto:svkuchina@yandex.ru)

каемым источником информационного насыщения и расширения прагма- тического потенциала.

Современная электронная художественная литература может быть успешно интегрирована в учебный процесс в качестве аутентичного ресурса для изучения иностранных языков и литературы1. Электронное художественное интерактивное произведение “Nightingale’s Playground” А. Кампбелл [Campbell 2010] состоит из четырех взаимосвязанных друг с другом частей. Первые две части представляют собой интерактивный художественный текст, именно на их основе разрабатывается “Nightin- gale’s Playground: pedagogy project”. Третья и четвертая части повество- вания — это виртуальная тетрадь Алекса Найтингейла (одного из героев электронного повествования) и его электронная книга с текстовыми, а также видео- и аудиофрагментами.

Общая сюжетная линия всех компонентов электронного нарратива свя- зана с размышлениями главного героя (Карла). Оставшись в одиночестве, он рассматривает содержимое своего старого чемодана, вещи-безделушки напоминают ему школьные годы, и Карл решает посетить встречу выпуск- ников в надежде увидеть старого школьного друга Алекса Найтингейла, но, к сожалению, никто из одноклассников не знает Алекса. Далее Карл вспоминает, что Алекс исчез еще во время их обучения в школе, и что с ним произошло точно, никто не знает. Вторая часть повествования, представленная в виде стилизованной 3D-игры с нарративным ком- понентом, в сюжетном плане она продолжает начатые в первой части поиски Алекса Найтингейла, но при этом локус повествования пере- мещается в квартиру Алекса. Здесь читатель, путешествуя в виртуаль- ном пространстве заброшенной квартиры, обнаруживает фрагменты текста в самых неожиданных местах жилища героя. Текстовые находки повторяют в идейном смысле основное предположение, артикулиру- емое главным героем в первой части повествования, о том, что Алексу Найтингейлу, а может быть, и самому Карлу удалось проскочить сквозь иллюзию повседневной реальности и они оба являются объектами вооб- ражения друг друга. Однако по поводу исчезновения Алекса в электрон- ном повествовании актуализирована и более тривиальная идея. В любом случае читатель заинтригован до конца повествования во второй части “Nightingale’s Playground” (до посещения Карлом последней комнаты в квартире Алекса).

1 Так, известен опыт реализации образовательного проекта “Inanimate Alice: Pedagogy Project”, подготовленного на основе электронного произведения К. Пилинжер и К. Джозе- фа “Inanimate Alice” [Inanimate Alice project]. Педагогический проект представляет собой серию детально разработанных планов занятий для преподавателя и рабочие материалы для обучающихся, нацеленные на тренировку навыков восприятия и интерпретации элек- тронного контента (текст, звук, анимация, графика, видео).

Образовательный проект к электронному произведению “Nightingale’s Playground” А. Кэмпбелл включает три основных модуля и один вводный, в котором представлена основная лексика и синтаксические конструкции, необходимые для успешного освоения учащимися тем основных моду- лей. Лексические единицы вводного урока могут быть распределены по следующим лексико-сематическим группам: ЛСГ 1: существительные, описывающие школьную жизнь и интересы подростков; ЛСГ 2: прила- гательные, описывающие свойства характера; ЛСГ 3: прилагательные, представляющие характеристики предметов, используемых в школьной жизни; ЛСГ 4: глаголы со значением передвижения; ЛСГ 5: глаголы со значением разрушения, разъединения целого; ЛСГ 6: существительные, описывающие нарративную концепцию повествования.

Лексический материал вводного модуля отрабатывается с помощью интерактивных упражнений, представленных в электронной интерактив- ной образовательной среде по иностранным языкам eLang НГТУ [eLang]. К основным типам заданий, используемым во водном модуле, относятся:

* 1. Установление соответствий (слово-перевод; слово-синоним; слово- антоним).
  2. Множественный выбор (выбор правильного ответа; выбор несколь- ких правильных ответов).
  3. Заполнение пропусков (в тексте или в диалоге).
  4. Восстановление последовательности (букв в слове; слов в пред- ложении).
  5. Открытый вопрос (ввод ответа).
  6. Аудио/видеотекст и задания к нему.

Главная задача основных модулей “Nightingale’s Playground: pedagogy project”: познакомить обучающихся с электронным нарративом как прин- ципиально новым литературным медиаобъектом и научить анализировать элементы (текст, звук, аудио, видео, графику) электронного повествова- ния. Основными компонентами модулей “Nightingale’s Playground: peda- gogy project” являются: план занятия для преподавателя (“Lesson Plan”) и материалы для самостоятельной работы студентов (“Student Recourse Pack”). Все планы занятий для преподавателей в “Nightingale’s Playground: pedagogy project” имеют сходную структуру, в которой последовательно представлены следующие части:

1. Тема раздела.
2. Теоретические материалы (список литературы со ссылками на источник, необходимые обучающимся для успешного освоения теории и практики раздела).
3. Цели обучения (а также навыки и умения, которые будут сфор- мированы у обучающихся в результате успешного освоения материалов раздела).
4. Общий план занятий.
5. Пошаговая инструкция для преподавателя.
6. Дополнительные материалы:
   * итоговые вопросы по темам раздела, предполагающие контроль освоения теоретических разделов модуля и навыков анализа электрон- ного художественного текста (summarizing);
   * дополнительные вопросы к материалам занятия (extending). Так, например, тема первого модуля в “Nightingale’s Playground: peda-

gogy project” — “Connection between story and medium” (Связь повество- вания и электронной медиасреды)”. В качестве теоретического материала студентам предлагается следующее исследование в области электронной литературы: Уордрип-Фруин Н. “Five elements of digital literature” [Wardrip- Fruin 2010].

Целью первого модуля “Connection between story and medium” явля- ется знакомство слушателей курса с основными элементами электронного повествования, а также формирование базовых навыков анализа электрон- ного нарратива (умение выделить основные компоненты и назвать их функции в повествовании). Общий план занятий предполагает прочтение студентами исследовательской работы в области электронного повествова- ния, которое осуществляется в микрогруппах и дополняется коллективным обсуждением прочитанного материала. Пошаговая инструкция реализует все направления общего плана занятия первого модуля с применением таких педагогических технологий, как элементы критического мышления:

* + - кластер (графическая организация понятия: electronic narrative);
    - прием INSERT (интерактивная система записи для эффективного чтения теоретических материалов, заявленных в материалах модуля);
    - participatory reading (процесс чтения и обсуждения интерактивного художественного текста, в котором реципиент принимает активное уча- стие).

Ряд упражнений модуля “Connection between story and medium” наце- лен на проверку понимания прочитанного (Reading Comprehension Tasks):

* + - заполнение пропусков в тексте;
    - текст и задания к нему;
    - восстановление последовательности (абзацев в тексте).

Итоговые вопросы модуля рассчитаны на проверку усвоения теорети- ческого и практического материала занятия:

* + - *Who are the main characters of “Nightingale’s Playground”?*
    - *Describe Carl’s inner world and his feelings about his school life and school mates?*
    - *Where does the main action of the story take place?*

Данные вопросы предполагают устные ответы непосредственно на занятии. Дополнительные вопросы по материалам занятия направлены

на осмысление структуры электронного повествования в аспекте его вос- приятия читателем. Например:

* + - * *How do you think the creators attempt to guide the readers through the evolving narrative?*
      * *Do you have your own suggestions for making the path through the nar- rative clearer?*

Материалы для самостоятельной работы студентов в данном модуле представляют собой анкету-эссе для слушателей курса со следующими вопросами-темами:

*What I did; What I enjoyed; What I found difﬁcult; What really worked.* Онлайн-анкетирование студентов планируется проводить на базе сервиса Webanketa [Webanketa 2015].

Помимо обучения языку, интеграция электронной литературы в обра- зовательное пространство может преследовать различные цели, в том числе и повышение медиаграмотности, актуализация межпредметных навыков. В образовательном процессе “Nightingale’s Playground: pedagogy project” может быть представлен в двух основных вариантах:

* + - * спецкурс на английском языке по электронной литературе;
      * несколько образовательных модулей в рамках основного курса по иностранному языку.

Педагогический проект “Nightingale’s Playground: pedagogy project” рассчитан на студентов с уровнем владения иностранным (английским) языком не ниже Intermediate. В частности, может быть использован в рам- ках курса «Практический курс иностранного языка» (аспект «Комменти- рованное чтение»), который является базовым в цикле профессиональных дисциплин учебного плана направления «Лингвистика», «Педагогическое образование» (квалификация «бакалавр»).

Все задания образовательного проекта “Nightingale’s Playground: peda- gogy project” после завершения разработки предполагают публикацию в следующих форматах:

* + - * печатном (учебное пособие);
      * электронном (интерактивный учебный курс по иностранному языку в электронной интерактивной образовательной среде по иностранным языкам eLang НГТУ).

**Литература**

1. Электронная интерактивная образовательная среда НГТУ по иностран- ным языкам eLang (2012–2014), available at: <http://elang.nstu.ru/>(дата обращения 15.09.2016).
2. *Кучина С.А.* (2015), Электронная литература: жанровые, семантические и структурные особенности. Филологические науки. Вопросы теории и практики, с. 119–123.
3. *Campbell A.* (2010), Nightingale’s Playground, available at: [http://nightingales-](http://nightingales-/) playground.com/ (дата обращения 15.09.2016).
4. Inanimate Alice project (2005–2016), available at: http://www.inanimatealice. edu.au/ (дата обращения 15.09.2016).
5. *Wardrip-Fruin N.* (2010), Five elements of digital literature, available at: https:// games.soe.ucsc.edu/sites/default/ﬁles/nwf-BC5-FiveElementsOfDigitalLiterature.pdf (дата обращения 15.09.2016).
6. Webanketa (2015), available at: <http://webanketa.com/ru/>(дата обращения 15.09.2016).

**Literature**

1. Elektronnaya interaktivnaya obrazovatel’naya sreda NGTU po inostrannym yazykam eLang (2012–2014), available at: <http://elang.nstu.ru/>(data obrashcheniya 15.09.2016).
2. *Kuchina S.A*. (2015), Elektronnaya literatura: zhanrovye, semanticheskie i strukturnye osobennosti. Filologicheskie nauki. Voprosy teorii i praktiki, s. 119–123.
3. *Campbell A.* (2010), Nightingale’s Playground, available at: [http://nightin-](http://nightin-/) galesplayground.com/ (data obrashcheniya 15.09.2016).
4. Inanimate Alice project (2005–2016), available at: http://www.inanimatealice. edu.au/ (data obrashcheniya 15.09.2016).
5. *Wardrip-Fruin N.* (2010), Five elements of digital literature, available at: https:// games.soe.ucsc.edu/sites/default/ﬁles/nwf-BC5-FiveElementsOfDigitalLiterature.pdf (data obrashcheniya 15.09.2016).
6. Webanketa (2015), available at: <http://webanketa.com/ru/>(data obrashcheniya 15.09.2016).

***Т.Н. Носкова, Т.Б. Павлова, О.В. Яковлева***

**СОЦИАЛЬНЫЕ МЕДИА: НОВЫЕ ОБРАЗОВАТЕЛЬНЫЕ ПРАКТИКИ И КОМПЕТЕНЦИИ1**

**Аннотация.** Анализ среды социальных медиа, в которых активна молодежь, выявляет ее потенциал в построении новых образовательных практик. Эффек- тивное достижение поставленных целей в них требует формирования особых компетенций педагогов. Раскрыт процесс формирования искомых компетенций. **Ключевые слова.** Социальные медиа, профессиональный опыт, профессио-

нальные компетенции, образовательная практика.

***T.N. Noskova, T.B. Pavlova, O.V. Yakovleva***

**SOCIAL MEDIA: NEW EDUCATIONAL PRACTICES AND COMPETENCES**

**Abstract.** Analysis of the social media environment, where young people are active, results in identifying its potential in building new educational practices. The effective achievement of set goals requires the of speciﬁc teachers’ competences. The process of the required competencies development is disclosed in the paper.

**Keywords.** Social media, professional experience, professional competences, educational practice.

Modern electronic space is the environment of new ways of interactions. A speciﬁc virtual world has been formed, and it is to some extent reﬂecting the social world. At the same time, this world has evolved a new society with its own norms and rules, based on the network ethics. For teachers it is important that the electronic space has become a habitat for the today’s youth — forerun- ners of the information society. This new space is different from the “face to face” classroom interactions. It takes place in a different establishment of social relations and interactions. Participants receive a large variety of information resources, social contacts and grounds for self-realization. At the same time,

1 ACKNOWLEDGMENTS: The research leading to these results has received, within the framework of the IRNet project, funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/ under REA grant agreement No: PIRSES-GA-2013-612536

© НОСКОВА Татьяна Николаевна. — Noskova Tatiana. Российский государственный педа- гогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [noskotatn@gmail.com](mailto:noskotatn@gmail.com)

© ПАВЛОВА Татьяна Борисовна. — Pavlova Tatiana. Российский государственный педа- гогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [pavtatbor@gmail.com](mailto:pavtatbor@gmail.com)

© ЯКОВЛЕВА Ольга Валерьевна. — Yakovleva Olga. Российский государственный педа- гогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University of Russia, Russia. E-mail: [ekzegeza@yandex.ru](mailto:ekzegeza@yandex.ru)

there are big risks to physical and mental health, and to the development of a growing person. Teachers need to “cultivate the culture” is the electronic space, developing new educational practices. But when developing new practices they should base not only of the accumulated professional and life experience, but also be deeply aware of the needs and demands of young people, active and initiative in the social media environment.

Social media is a variation of mass communication through the Internet, a historically new kind of communication. It is deﬁned as a group of Internet- based applications, ideologically and technologically Web 2.0 based, allowing participants to communicate in social networks and to create content in the process of exchanging it. Social media is a part of e-environment interactions in which a user opens a media channel to broadcast content in a certain direc- tion, addressed to a wide audience. This audience is based on self-organization and participants begin to interact with each other in the discussion of media messages and create new messages in this regard as a result of a network of cooperation. Message content producers enter into a relationship with the consumers of these messages and involve them in further content production. Participants of electronic interactions can share knowledge, experiences, opin- ions, news, videos, photos, music, links to websites, and so on to establish and develop social contacts.

Today, various education practices are based on social media implementa-

tion. Acquisition of various skills occurs in the remote interactions. Commu- nication includes not only texts but also images and multimedia. A network communication is formed by a variety of dynamic and interrelated resources that are created not only by experts, but also by students. Description of features of these interactions is the subject of research [Patarakin et al. 2015; Kommers et.al 2015].

The most vital questions are the following. Under what conditions new professional competence can be formed? How education specialists the new century can learn solving professional problems through electronic environ- ment? What are the characteristics of the environment where these competen- cies can be applied?

What are the generalized characteristics of new education practices based on social media? On the technical side, they include the use of Web 2.0 tech- nologies and social media channels. On the pedagogical side, these are specially organized education interaction, stimulating activity of electronic media users. On the psychological side, such new practices are creating a unique atmosphere of cooperation, unity of purpose and objectives of joint activities. From a com- municational point of view, it is important that online communities are bundles of interactions with students themselves being organizational centres, not just a teacher acting as an organizer of the process. However, all these features are external, formal and simply recognized. At the same time, there are certain

underlying factors that exhibit the novelty of pedagogical practices discussed here. These factors are associated with the formulation of new goals and objec- tives of education activities. Tasks that encourage students to go beyond the education standard, to ﬁnd their own meanings and motives in the education activity. Finding these new meanings can radically change a course of action. In general, comparing the new and traditional education practices, we can state the following. In the classroom education practices, the leading role belongs to a teacher, who sets out a content and organizes activities of students. Thematic content meets the education standards, it is the same for everyone and it is obligatory to assimilate for all students. Pedagogical technology of a classroom practice is built as a linear sequence of group activities, with the feedback, provided by the teacher, monitoring and evaluating all activities.

All this stimulates the implementation of education standards in the tight time framework of classroom teaching for the whole group of students.

When using social media as a set of arrangements, interactions and practices that aim to create learning environment based on the use of social services in order to provide users’ activity, the situation is changing. Current users of the network are themselves the authors of the network and can add their content — articles, photos, audio and video, leave comments, shape design of their pages, etc. Using the potential of social media requires a different course of education practices that allow evolving initiative, increasing activity through the choice, self-awareness of own meanings in the performed activities. Therefore, the aim of the new education activity is not just the implementation of training on the orders of a teacher, but the exertion of activities and initiatives in learning. In a networked environment, interacting with others, being an avid user of personally learned knowledge, a student carries out own search for new knowl- edge, development and competence, and achieves results in the joint distributed activities. Emotions, motivations, meanings of learning are changing [Pfaff- Rüdiger et al. 2016]. All this makes it possible to achieve new results, under certain pedagogical conditions during interactions through social media tools. Thus, through these new education practices the emergence of learning autonomy can be stimulated. According to such strategy, a learner builds his/ her own knowledge independently in a specially created and maintained by a teacher e-learning environment. In this environment, learners are able to transfer their knowledge and skills in the practice of life — to teach others, help them to develop their knowledge, participate in various projects, creative activities.

Therefore, in these practices moving beyond standard of education activities

and self-realization is stimulated.

Today, social media is implemented for different purposes, for various focus groups [Ricoy 2016], as resources for solving educational problems, as new communication formats, and for organizing feedback [Noskova et al. 2014, 2015; Mkttchian 2015].

Consequently, a modern teacher needs to have competences that help preparing children and youth to interact with a wide variety of social media. However, social media themselves are the basis for the complex resolution of this problem as they are involving young people in the educational interactions, allowing to realize new opportunities and eliminate risks of this medium.

Obviously, the new practices discussed in this paper do not replace the tradi- tional ones, but provide an opportunity for a new way to organize the education process. Therefore, it is advisable to build learning interactions both based on classroom education activity and electronic environment activity.

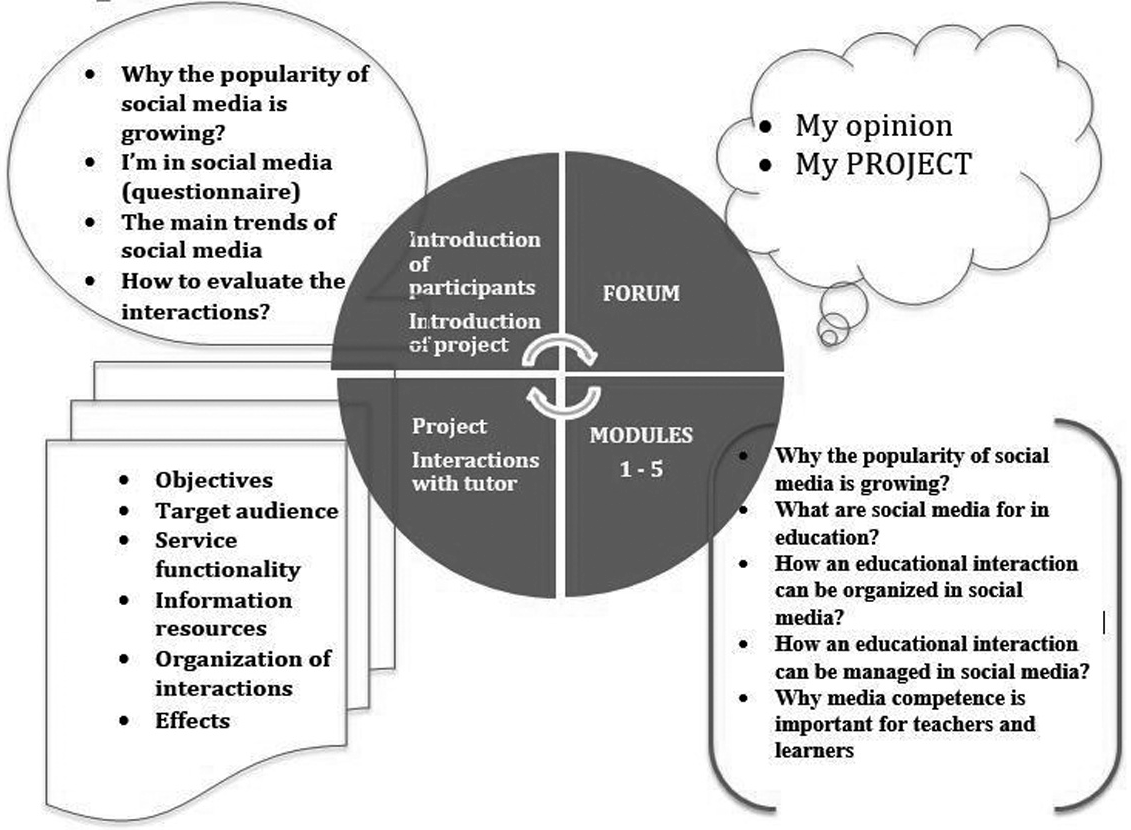
The Institute of Computer Science and Technology Education of Herzen University (Russia) has been implemented for three years jointly with the UNESCO Institute for Information Technologies in Education a course “Social Media and New Education Practices”. The course is addressed to master degree students, majoring in pedagogical education. The main goal of the course is to form competencies of teaching and solving educational problems in the modern media environment.

The relevance of the course is determined by several reasons:

* + - extensive use of social media in solving problems of education practice;
    - high motivation of young people to use social services;
    - renewed education standards demand revision of education practices. The course content includes the following issues:
    - new socio-psychological phenomenon in the global electronic environment;
    - social media: concept, social signiﬁcance and pedagogical potential;
    - media education problems;
    - the role of social media in the education environment: design of media environment to address speciﬁc education objectives;
    - Internet services of social media and virtual education methodology based on the interactions based on social media.

In the process of studying each module of the course students develop their pedagogical projects of an e-learning environment, through which, using the tools of social media, they are preparing for the decision of professional tasks related to the theme of master thesis. Master degree students from the Institute of Pedagogy and Psychology, Institute of Childhood, Faculty of Correctional Pedagogy joined the study of the course in the spring semester of 2016. General outline of the course is presented in Figure.

In education practice has become popular such a model, implemented on the basis of social media, which takes place parallel in a classroom and an e-learning environment. In social networks, examples that implement educa- tional activities based on social media are available. However, master degree students who studied the course “Social Media and New Education Practices” were able to reach a higher level — to design such education practices that



*Figure.* Social Media and New Education Practices. General course outline

evolve in a network environment with the interactions unfolding in electronic formats.

New education practices implemented based on social media is not just the expansion and enrichment of the traditional educational activities in auditorium. They can be considered the prototype of the future, when students themselves will demonstrate conscious educational initiative in ﬁnding solutions and adopting the goals and objectives of the educational activities for building competencies. In search and acquisition of electronic educational resources, they will start looking for network partners to work together, provide mutual support and share knowledge. Sharing knowledge helps to achieve the intended results better and to correct actions. They will seek for the assistance of teachers for the evaluation and correction of their knowledge and skills, building up a more optimal education path. These network communities of practice will stimulate the productive activity of students, their initiative based on the generated motivations to use acquired knowledge, to carry them into practice, exchanging them with other users, discuss, create, and act together.

Important competences will be formed in such organization of education

interactions of students — not only professional, but also social competences. The most important educational activity is personally meaningful; it has the

true sense and strengthened motivation. All this is necessary to the further education path, to the aspiration for lifelong learning in a dynamic knowledge society.

**Literature**

1. *Kommers P., Isaias P., Issa T.* (2015), Perspectives on social media: A Year- book.: Routledge.
2. *Mkttchian V., Bozhday A., Bershadsky A., Noskova T.* (2015), Development of a Global Policy of All-Pervading E-Learning, Based on Transparency, Strategy, and Model of Cyber Triple H-Avatar In Development of a Global Policy of All-Pervading E-Learning, Based on Transparency, Strategy, and Model of Cyber Triple H-Avatar, рр. 207–208.
3. *Noskova T., Pavlova T., Yakovleva O., Sharova N.* (2014), Communicative competence development for future teachers. The New Educational Review, 38 (14), рр. 189–199.
4. *Noskova T., Yakovleva O., Pavlova T., Smyrnova-Trybulska E.* (2015), Students in the information environment: a study of educational and extracurricular activities Int. J. Continuing Engineering Education and Life-Long Learning, Vol. 25, No. 4, pp. 394–410.
5. *Patarakin Y., Shilova O.* (2015), Concept of Learning Design for Collaborative Network Activity. Procedia — Social and Behavioral Sciences 214, 1083–1090.
6. *Pfaff-Rüdiger S. Riesmeyer C.* (2016), Moved into action. Media literacy as social process. Journal of Children and Media, 10 (2), pр. 164–172.
7. *Ricoy M.-C., Feliz T.* (2016), Twitter as a learning community in higher educa- tion. Educational Technology and Society, 19 (1), рp. 237–248.

***Н.К. Рябцева***

**СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПРИКЛАДНЫЕ ПРОБЛЕМЫ ПРЕПОДАВАНИЯ ИНОСТРАННЫХ ЯЗЫКОВ**

**Аннотация.** Статья посвящена обсуждению принципов и возможностей при- ложения новых информационных технологий к обучению иностранному языку. Особое внимание уделено графической визуализации интегрированных лингви- стических данных, в частности, представлению грамматических конструкций и контекстов употребления и сочетаемости тематически упорядоченной лексики. Статья основана на большом лингвистическом материале, в качестве которого выступает созданный автором Конкорданс «Активатор научной речи на англий- ском языке: Тематически упорядоченная общенаучная лексика с сжатыми контек- стами употребления», являющийся развитием идей ранее изданной автором книги

«Научная речь на английском языке» (М., 2013).

**Ключевые слова.** Тематически упорядоченная лексика, контексты употре- бления общенаучной лексики, графическая визуализация контента.

* 1. ***Riabtseva***

**NEW INFORMATION TECHNOLOGIES AND THEIR APPLICATION IN FOREIGN LANGUAGE LEARNING**

**Abstract.** The paper focuses on the principles and perspectives in applying new information technologies in foreign language learning and teaching. Special attention is paid to graphic visualization of integrated linguistic data, particularly to exposing construction grammar, as well as combinatory and collocational contexts of thematically arranged lexicon. The paper is based on extensive Concordance «Speech activator of academic writing: A thematically arranged lexicon in condensed collocational contexts» compiled by the author as an extension to already published *English for Scientiﬁc Pur- poses: A New Active Reference* (Moscow, Nauka-Flinta, 20013, 6th ed.).

**Keywords.** Thematically arranged lexicon, collocational contexts, graphic visu- alization of linguistic content.

Благодаря изобретению персонального компьютера и интернета, современные информационные цифровые технологии используются практически во всех областях деятельности человека. Важное значение они приобретают в преподавании иностранных языков. Они позволяют создавать бесконечное многообразие учебных, обучающих и справочных пособий, программ и тренажеров самых разных типов и видов, а также воплощать потребности преподавания иностранных языков в виде ценных

© РЯБЦЕВА Надежда Константиновна. — Riabtseva Nadezhda. Институт языкознания РАН, Россия. — Institute of Linguistics, Russian Academy of Sciences, Russia. E-mail: nadia\_ [riabceva@mail.ru](mailto:riabceva@mail.ru)

учебных материалов нового типа и создавать тем самым лингвистическое обеспечение инновационной образовательной среды — компьютерной лингводидактики (КЛД) [Беляева 2014]. Она включает в первую очередь особым образом организованные и автоматизированные конкордансы, тематические тезаурусы, словари-энциклопедии, разного рода «актива- торы речи», а также особые «лингвистические каталоги» и «визуализа- торы», вбирающие последние достижения теоретической, прикладной,

«лексикографической», компьютерной, квантитативной, когнитивной и корпусной лингвистики и мн. др. Так, современные информационные технологии позволяют значительно усовершенствовать лексикографиче- скую работу за счет использования таких порождаемых ими свойств, как

«модульность, гибкость, динамичность, сбалансированность, дружествен- ность» [Беляева 2010: 79]. При этом важнейшее значение имеет нагляд- ное, функционально и операционально ориентированное представление тематически и грамматически/«конструктивно» упорядоченной лексики, поскольку способствует ее активному усвоению, т. е. не только ее пони- манию, но и употреблению.

**Инновационные технологии в преподавании иностранного языка** В течение более чем двадцати лет разработки компьютерных про- грамм, способствующих обучению (иностранному) языку, назывались “computer-*assisted* language learning” (CALL: «обучение языку с *помощью* компьютера»). В середине 2000-х гг. подобные системы стали перепро- филироваться в “computer-*enhanced* language learning” (CELL: «повыше- ние *качества* обучения иностранному языку с помощью компьютера»), в частности, за счет использования разнообразных мультимодальных форм представления информации, интерактивности интерфейса и т. п. [Petrie 2005]. В настоящее время одним из перспективных прикладных направ- лений в разработке учебных пособий по иностранному языку выступает также интеграция в программные продукты лингвистической информации

различных типов и видов.

При этом главными дискурсивными и коммуникативными навыками и умениями, необходимыми для активного освоения иностранного языка, остаются: внимание к внутриязыковой лексико-грамматической сочетаемости, свойственной именно для данного иностранного языка, выделение и использование аутентичных прагматических и стилистиче- ских выразительных средств общения, которые употребляют носители языка в заданных социальных, коммуникативных, предметных, профес- сиональных и т. п. ситуациях, умение компактно и эффективно свернуть и развернуть свой или чужой текст (ср. *summary*), осознание отражения в речи культуры носителя языка, в частности, межперсональной «дис- танции» и др.

Так, в качестве важнейшего инновационного приложения новых информационных технологий к преподаванию иностранных языков и фор- мированию активных коммуникативных навыков выступает такой прием компьютерной лингводидактики, как графическая визуализация инте- грированного лингвистического контента: одновременное представление структуры грамматических конструкций, контекстов употребления темати- чески упорядоченной лексики и ее лексико-семантической сочетаемости. Для создания такого лингводидактического контента нами используется представительный объем лингвистических данных, в качестве которых выступает созданный автором Конкорданс «Активатор научной речи на английском языке: Тематически упорядоченная общенаучная лексика с сжатыми контекстами употребления». Он является развитием идей, воплощенных в изданной ранее книге [Рябцева 2013], и служит основой для создания соответствующего интерактивного интернет-ресурса.

Так, в [Рябцева 2014] были показаны принципы приложения новых

информационных технологий графической визуализации к данному интегрированному лингвистическому контенту (Конкордансу) в иссле- довательских и дидактических целях, в частности, к представлению грамматических конструкций, контекстов их употребления и сочетаемо- сти образующей их тематически упорядоченной общенаучной лексики на английском языке.

Отличительной особенностью созданного Конкорданса выступает принцип интегральности представления лингвистической информации, а также многофакторная тематическая квалификация общенаучной лексики. Последняя разбита на семантические классы: интенциональ- ную, ментальную, логическую, модальную, метатекстовую, оценочную,

«связочную» и др., которые далее разбиты на тематические классы и подклассы слов с приписанными им моделями их контекстуального использования. Благодаря этому Конкорданс эксплицирует семантические, тематические и «ситуативные» связи внутри общенаучной лексики, кото- рые были заложены в словарь-справочник [Рябцева 2013] имплицитно.

Эти связи проявляются в том, что на каждом шаге научного изложения может быть использован комплекс разнообразных общенаучных лингви- стических средств описания и рассуждения, которые взаимодействуют друг с другом, а также взаимодополняют и взаимозаменяют друг друга. Так, тематический класс лексики MENTAL (PROCEDURES, STATES, OPERATIONS, ATTITUDES, etc.) включает следующие тематические подклассы: *1.1. Research procedures; 1.2. Steps of research; 1.3. Research pro- cesses; 2.1. Prospecting; 2.2. Instrumentalia; 2.3. Priorities; 3.1. Mental states;*

*3.2. Mental apparatus; 3.3. Critical attitudes; 4.1. Inquiring; 4.2. Inquiries;*

*4.3. Qualiﬁcations; 5.1. Suppositions; 5.2. Mental impressions; 5.3. Prospects;*

* 1. *Disposals. Mental ontology; 6.2. Requirements: Mental abilities/ qualities/*

*procedures; 6.3. Attitudes and opinions*. Тематический подкласс *Mental appara- tus*, в свою очередь, включает следующую тематически связанную лексику: *argument, argumentation, conclusion, decision, expectation, insight, judgment, opinion, prediction* и др., каждый элемент которого естественным образом сочетается с модальной, интенциональной, оценочной и т. д. лексикой, ср.: *to advance/ appeal to/ convince by/ back up/ consider/ develop/ evaluate/ offer/ present/ produce/ propound/ put forward/ raise/ use/ (fail to) make/ be interpreted as a (circular/ cogent/ conclusive/ convincing/ detailed/ elegant/ general/ irrefutable/ logical/ persuasive/ rational/ solid/ sound/ strong/ valid)*

***argument*** *about/ against/ over/ (to the effect) that P.*

В целом можно сказать, что Конкорданс «Активатор научной речи на английском языке» совмещает в себе свойства гипертекста, тезауруса и тематического указателя, идеи грамматики конструкций, особенно- сти сочетаемости общенаучной лексики и принципов ее употребления в порождении научного текста на английском языке, позволяющие пользователю выбирать модель развития текста из ряда коммуникативно аналогичных вариантов.

### Перспективы и задачи

Научное направление «прикладные (лингвистические) проблемы пре- подавания иностранных языков», целью которого выступает разработка лингвистического и программного обеспечения различных пособий и материалов нового типа по обучению активному владению иностранным языком, предполагает не только создание инновационных методик и тех- нологий обучения (в частности, посредством наглядного и «объемного» представления лингвистических знаний), но и вовлечение обучающихся в формирование материалов обучения, в первую очередь при помощи современных информационных технологий, а также формированию принципов автоматического порождения упражнений и домашних заданий и мн. др. [Knoop et al. 2013].

Все это принципиально изменяет традиционную дидактическую область и придает ей современный креативный, «цифровой» и алгоритми- ческий характер, а также позволяет преодолевать отмечаемое специали- стами несоответствие уровня развития высоких технологий и качества образовательных программ, которые нуждаются в принципиальном обновлении, особенно за счет использования современных цифровых технологий как в общем, так и в профессиональном образовании. Так, вместо того чтобы, по старинке, заполнять пропуски в упражнениях учебника по иностранному языку подходящими элементами, учащиеся сами могут конкретизировать себе задания по работе с интернетом и установить, например, особенности сочетаемости общенаучной лексики со (своей) специальной лексикой в текстах определенной (например,

технической) тематики и произвести их систематизацию, превратив тем самым обыденное занятие в *edutainment* и *infotainment*. Подобные новые технологии составляют в настоящее время целое комплексное направление в педагогике и образовании под названием *Information and communication technologies (ICTs) in education and learning* («Информационные и комму- никационные технологии в образовании и обучении»).

В целом следует подчеркнуть, что современное общество и культура существуют в беспрецедентно быстро изменяющихся социальных и тех- нологических условиях, которые необходимо своевременно распознавать, квалифицировать и использовать, особенно в образовательных процессах, методических разработках, дидактических материалах и технологических инструментах, в частности, в обучении иностранному языку, поскольку оперативная разработка комплексных образовательных ресурсов, про- грамм и контентов, отличающихся инновационностью, технологичностью и мультимедийностью, позволяет сократить «лаг» между передовыми информационными технологиями, включая и мобильные, и традицион- ные методики обучения, в том числе и иностранному языку, и сделать образование более современным, интересным, актуальным, эффективным, действенным и развивающим, т. е. отвечающим современным требованиям “digital humanities” — цифровых гуманитарных наук [Mouri et al. 2015].

**Литература**

* + 1. *Беляева Л.Н.* (2010), Потенциал автоматизированной лексикографии и при- кладная лингвистика. Известия Российского гос. пед. ун-та им. А.И. Герцена,

№ 134, с. 70–79.

* + 1. *Беляева Л.Н.* (2014), Прикладная лингвистика и современная образователь- ная среда: проблемы и перспективы. Прикладная лингвистика в науке и образо- вании. Материалы международной конференции, Санкт-Петербург, 10–12 апреля 2014 г., с. 174–177.
    2. *Рябцева Н.К.* (2013), Научная речь на английском языке. Руководство по научному изложению. Словарь оборотов и сочетаемости общенаучной лексики: новый словарь-справочник активного типа (на английском языке). М.: Наука- Флинта, 6-е изд.
    3. *Рябцева Н.К.* (2014), Грамматика конструкций, словари сочетаемости, современные информационные технологии и преподавание иностранных языков. Лингвистика и методика преподавания иностранных языков: периодический сб. научных статей. Электронное научн. изд. М., ИЯз РАН, Вып. 6, с. 214–247. URL: [http://www.iling-ran.ru/library/sborniki/for\_lang/2014\_06/9.pdf.](http://www.iling-ran.ru/library/sborniki/for_lang/2014_06/9.pdf) (дата обращения 30.08.2016)
    4. *Рябцева Н.К.* (2015), Тенденция к визуализации в современном инфор- мационном пространстве, проблемы образования и инновационные технологии в преподавании иностранных языков. Лингвистика и методика преподавания иностранных языков: периодический сб. научных статей. Электронное научн.

изд. М.: ИЯз РАН, Вып. 7, с. 345–368. URL: <http://www.iling-ran.ru/library/sborniki/> for\_lang/2015\_07/13.pdf (дата обращения 30.08.2016).

* + 1. *Knoop S., Wilske S.* (2013), WordGap — Automatic Generation of Gap-Filling Vocabulary Exercises for Mobile Learning, Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning at NODALIDA, рp. 39–47.
    2. *Mouri K., Ogata H.* (2015), Ubiquitous learning analytics in the real-world language learning, Smart Learning Environments. Mouri and Ogata. 2015, 2:15. DOI: 10.1186/s40561-015-0023-x. URL: <http://slejournal.springeropen.com/articles/10.1186/> s40561-015-0023-x. (дата обращения 30.08.2016).
    3. *Petrie G. M.* (2005), Visuality and CALL research, CALL Research Perspectives. Ed. by Egbert, Joy L., Petrie Gina M. Lawrence Erlbaum Associates. URL: [http://lin-](http://lin-/) guistlist.org/pubs/books/get-book.cfm?BookID=16319 (дата обращения 30.08.2016).

**Literature**

1. *Belyaeva L.N.* (2010), Potentsial avtomatizirovannoy leksikograﬁi i prikladnaya ling¬vistika. Izvestiya Rossiyskogo gos. ped. un-ta im. A.I. Gertsena, № 134, s. 70–79.
2. *Belyaeva L.N.* (2014), Prikladnaya lingvistika i sovremennaya obrazovatel’naya sreda: problemy i perspektivy. Prikladnaya lingvistika v nauke i obrazovanii. Materialy mezhdunarodnoy konferentsii, Sankt-Peterburg, 10–12 aprelya 2014 g., s. 174–177.
3. *Ryabtseva N.K.* (2013), Nauchnaya rech’ na angliyskom yazyke. Rukovodstvo po nauchnomu izlozheniyu. Slovar’ oborotov i sochetaemosti obshchenauchnoy leksiki: novyi slovar’-spravochnik aktivnogo tipa (na angliyskom yazyke). M.: Nauka-Flinta, 6-e izd.
4. *Ryabtseva N.K*. (2014), Grammatika konstruktsiy, slovari sochetaemosti, sovre- mennye in¬formatsionnye tekhnologii i prepodavanie inostrannykh yazykov. Lingvis- tika i metodika prepodavaniya inostrannykh yazykov: periodicheskiy sb. nauchnykh statey. Elektronnoe nauchn. izd. M., IYaz RAN, Vyp. 6, s. 214–247. URL: [http://www.](http://www/) iling-ran.ru/library/sborniki/for\_lang/2014\_06/9.pdf. (data obrashcheniya 30.08.2016)
5. *Ryabtseva N.K*. (2015), Tendentsiya k vizualizatsii v sovremennom informatsion- nom prostranstve, problemy obrazovaniya i innovatsionnye tekhnologii v prepodavanii inostrannykh yazykov. Lingvistika i metodika prepodavaniya inostrannykh yazykov: periodicheskiy sb. nauchnykh statey. Elektronnoe nauchn. izd. M.: IYaz RAN, Vyp. 7,

s. 345–368. URL: <http://www.iling-ran.ru/library/sborniki/for_lang/2015_07/13.pdf> (data obrashcheniya 30.08.2016).

1. *Knoop S., Wilske S.* (2013), WordGap — Automatic Generation of Gap-Filling Vocabulary Exercises for Mobile Learning, Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning at NODALIDA, рp. 39–47.
2. *Mouri K., Ogata H.* (2015), Ubiquitous learning analytics in the real-world language learning, Smart Learning Environments. Mouri and Ogata. 2015, 2:15. DOI: 10.1186/s40561-015-0023-x. URL: <http://slejournal.springeropen.com/articles/10.1186/> s40561-015-0023-x. (data obrashcheniya 30.08.2016).
3. *Petrie G. M.* (2005), Visuality and CALL research, CALL Research Perspec- tives. Ed. by Egbert, Joy L., Petrie Gina M. Lawrence Erlbaum Associates. URL: <http://linguistlist.org/pubs/books/get-book.cfm?BookID=16319>(data obrashcheniya 30.08.2016).

***Н.Л. Шубина, А.А. Виландеберк***

**ОБРАЗОВАТЕЛЬНЫЕ ПРОГРАММЫ**

**«АЛЬТЕРНАТИВНЫХ ФОРМАТОВ» В ВЫСШЕМ ОБРАЗОВАНИИ: КОНКУРЕНТНОЕ ПРЕИМУЩЕСТВО ИЛИ СОМНИТЕЛЬНЫЙ РИСК?**

**Аннотация.** В статье рассматривается сетевая образовательная программа как один из видов образовательной программы «альтернативного формата». Описан опыт проектирования сетевой образовательной программы в формате открытого образовательного ресурса: этапы ее создания и условия внедрения в образова- тельный процесс. Исследованы ее специфика и задачи.

**Ключевые слова.** Открытый образовательный ресурс, сетевая образователь- ная программа, сетевое взаимодействие, образовательная программа «альтерна- тивного формата».

* 1. ***Shubina, A.A. Vilandeberk***

**ALTERNATIVE EDUCATIONAL PROGRAMS**

**IN HIGHER EDUCATION: COMPETITIVE ADVANTAGE OR DOUBTFUL RISK?**

**Abstract.** The article describes alternative educational programs. One of the types of such program is network educational programs. The process of its designing and conditions of its implementation are presented. Also the speciﬁc characteristics and tasks are studied.

**Keywords.** Open educational recourse, network educational program, network communication, alternative educational programs.

В настоящее время одной из приоритетных задач в высшем образова- нии является проектирование и реализация программ «альтернативных форматов». Речь идет о создании принципиально новой методической системы и новой образовательной среды, отражающей новую идеологию образования — «Образование для всех». ЮНЕСКО рассматривает разра- ботку и реализацию такого рода программ как способ обеспечения каче- ственного образования для всех и как средство достижения общих целей социальной интеграции (48-я Международная конференция ЮНЕСКО по образованию, Женева, 2008).

Программы «альтернативных форматов» должны обеспечить инди- видуализацию, интерактивность обучения и моделирования специальной

© ШУБИНА Наталья Леонидовна. — Shubina Natalia. Российский государственный педа- гогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical University, Russia.

© ВИЛАНДЕБЕРК Анна Арнольдовна. — Vilandeberk Anna. Российский государственный педагогический университет им. А.И. Герцена, Россия. — Herzen State Pedagogical Univer- sity, Russia. E-mail: [magistr52@mail.ru](mailto:magistr52@mail.ru)

обучающей среды, а их разработка обусловлена необходимостью инте- грации людей с ОВЗ в общеобразовательную среду. Увеличение доли неформального образования свидетельствует о необходимости нового концептуального понимания открытого образования, в основе которого лежит расширение доступности, академической свободы и самостоятель- ного выбора образовательного маршрута и способов обучения. Реализация Федеральных целевых программ развития образования предполагает мобильное внедрение дополнительных образовательных программ, направленных на развитие креативной и конструктивной личности. Приоритетность разработки новых моделей программ «альтернативных форматов» обусловлена их финансовой рентабельностью как для обучаю- щихся, так и для образовательных организаций.

Проектирование и реализация такого рода программ меняет понима- ние идеологии современного высшего образования. Однако следует при- знать наличие рисков, возникающих при их реализации, что обусловлено привлечением к проектированию программ, их методическому обеспе- чению и информационно-организационному сопровождению разработ- чиков с опытом традиционного проектирования. Формальный перенос принципов формирования содержания программ, структурирования обра- зовательного процесса, инструментария диагностики качества освоения программы, отсутствие четкого понимания возможностей использования образовательных технологий и наличие стереотипов в управлении про- цессами обучения приводит к негативным результатам при попытках внедрения таких программ, а также и к дискредитации отношения про- фессионального сообщества к пониманию «нового образования».

Не менее сложным и противоречивым становится понимание новых

моделей образовательных программ. Например, до сих пор отсутствует единое понимание назначения, целей, задач и, самое главное, способов реализации **образовательных программ в сетевом формате** (так назы- ваемых сетевых программ) [Виландеберк и др. 2015]. Как показывает наш практический опыт, сетевую программу следует рассматривать как открытый образовательный ресурс нового типа, направленный на стиму- лирование личностного потенциала обучающихся, наращивание мето- дического потенциала всех участников сетевого сообщества в условиях непрерывного образования. Учитывая, что в настоящее время не опреде- лен методический потенциал сетевых программ, не выработаны единые принципы организации их архитектуры, трансформации информации в учебный материал, при разработке и внедрении сетевых программ следует опираться на опыт проектирования и разработки открытых обра- зовательных ресурсов. Следует обратить внимание на решение первооче- редных задач: формирование сетевого информационного пространства как зоны педагогической рефлексии; подготовка преподавателей, гото-

вых работать в новых условиях открытого и дистанционного обучения; определение принципов и методического потенциала образовательного контента, открытых образовательных ресурсов и образовательных техно- логий в новом формате; разработка принципов стандартизации сетевого взаимодействия, сетевых образовательных продуктов и критериев их оценки; разработка методических экспертных рекомендаций для препо- давателей и студентов по осуществлению образовательной деятельности в новой среде.

Одним из ключевых этапов для успешной разработки и реализации сетевых программ является поиск сетевых партнеров, готовых и спо- собных работать в условиях единого образовательного пространства на принципах взаимопартнерства. Этот процесс требует временных затрат, сформированных и устойчивых связей между сетевыми партнерами, которые проявляются в умении вести профессиональный диалог, объек- тивно оценивать свои возможности и способности, готовности к обмену актуальной информацией на взаимовыгодных началах.

Как свидетельствует опыт, можно говорить о наличии устойчивых факторов, сдерживающих развитие сетевого взаимодействия: не реализо- ваны на практике принципы формирования открытого образовательного пространства и механизмы развития межвузовской академической мобиль- ности студентов и преподавателей; не выработаны принципы взаимо- выгодного сотрудничества членов сетевого сообщества; не определены правовые аспекты сетевого взаимодействия и принципы гармонизации образовательных программ; не разработано профессиональное программ- ное обеспечение реализации сетевых программ; не создана новая система подготовки и переподготовки кадров.

Сетевая программа как открытый образовательный ресурс отличается от традиционных программ на всех этапах проектирования, реализации, корректировки и экспертизы качества программы. Сетевая программа должна ориентироваться на создание основы для индивидуальной само- стоятельной реализации образовательной деятельности, обеспечивающей социальную и профессиональную успешность, развитие творческих способностей, саморазвитие и самосовершенствование; обеспечить непре- рывное взаимодействие сетевых партнеров для постоянного совершен- ствования и самосовершенствования профессиональных компетенций; обеспечить реализацию запросов пользователя, а именно достижение целей-ориентиров, целей-результатов, целей-возможностей; ориентиро- ваться на подготовку специалиста, способного самостоятельно определять траекторию профессионального роста.

Для эффективной реализации «альтернативных программ» необхо- димо соблюдение целого комплекса условий.

1. Наличие соответствующей открытой образовательной среды, для которой характерны свобода выбора образовательной траектории в рам- ках одного образовательного учреждения, комбинаторная возможность выбора образовательных учреждений в рамках одной образовательной программы.
2. Наличие инновационного потенциала среды, обеспечивающего поддержку активности, креативности и самостоятельности обучаемого и обучающего (оптимальное распределение познавательной деятель- ности).
3. Соблюдение принципов адаптивности, гибкости и адекватности сетевой программы, формирующих основу нелинейного образовательного процесса.
4. Организация образовательного процесса на базе кластер-технологий [Виландеберк и др. 2011].
5. Проектирование сетевой программы на основе технологий моду- ляризации.
6. Наличие инновационного потенциала открытого образовательного ресурса.

Проектирование сетевой образовательной программы отличается нелинейным характером, что предполагает отсутствие «пошагового» плана разработки программы. Проектирование такой программы осу- ществляется в рамках сетевого взаимодействия рабочих групп, которые осуществляют свою деятельность в автономном и параллельном режиме, сокращая временные расходы на проектирование программы, а также осуществляя своеобразный «мониторинг мнений». Ведущая роль при- надлежит экспертной группе, которая должна провести «кастинг» идей, подходов и способов их осуществления. Проектный коллектив состоит из специалистов разного профиля, представляющих различные образо- вательные учреждения.

Как показывает опыт *проектирования сетевых образовательных про- грамм*, представляется возможным выделить следующие этапы:

1. Разработка блока целей и задач сетевой программы на базе индивидуально-ориентированного и деятельностного подходов (креатив- ность, самостоятельность, ответственность).
2. Разработка моделей сетевой программы и путей ее реализации.
3. Разработка содержательной составляющей разных моделей сетевой образовательной программы.
4. Разработка архитектуры сетевой программы и ее информационно- методического сопровождения.
5. Разработка информационно-методической поддержки сетевой про- граммы (рекомендации для менеджмента и преподавателей).
6. Отбор специальных технологий работы в формате сетевой про- граммы.
7. Разработка сетевого этикета участников образовательной про- граммы.
8. Разработка алгоритма действия пользователя при формировании индивидуальной образовательной программы: постановка конкретных целей пользователя и определение способов их корректировки и/или изменения.
9. Формирование состава менеджмента реализации сетевой про- граммы и ее информационно-методической поддержки.

Главной отличительной характеристикой сетевой программы является тот факт, что «условия, в которых проходит обучение, становятся контек- стом, где применяются знания, а характер знаний и учебные материалы становятся профессиональными инструментами, при помощи которых и осуществляется рабочий процесс [Виландеберк и др. 2010].

Сетевая образовательная программа как открытый образовательный ресурс направлена на решение следующих задач:

1. *Формирование образовательного сообщества* — «сообщества практики» [Атвелл], в рамках которого участники сообщества делятся своим пониманием различных вопросов и конкретным профессиональ- ным опытом.
2. *Стимулирование личностного потенциала* обучающегося, которое становится возможным в процессе интеллектуального обучения за счет постоянного общения в «группе равных».
3. *Развитие умения командной работы*, реализуемое в процессе постоянной проектной деятельности, распределения функций, пред- ставления результатов коллективной работы и их объективной коллек- тивной и индивидуальной оценки. «Предоставление учащимся доступа к информационно-коммуникационным технологиям является решающим фактором не только для претворения в жизнь их целей и планов, но и для подготовки успешных граждан, способных приносить реальную пользу государству и обществу» [Дендев (ред.) 2013].
4. *Наращивание методического потенциала* всех участников сете- вого сообщества в условиях непрерывного и многоэпизодного обучения на протяжении всей профессиональной деятельности за счет открытого доступа к конкретным практическим решениям, новым тенденциям и тех- нологиям. «Социальное программное обеспечение и социальные медиа позволяют увеличить интеллектуальный потенциал обучающегося за счет использования внешних источников знания и распространения процессов познания. «Как следствие, образование неизбежно будет ориентировано на актуальные и практически значимые вопросы: сейчас уже недоста-

точно просто позволить учащимся изучать то, что другие освоили ранее» [Дендев (ред.) 2013].

Таким образом, только при выполнении всех вышеперечисленных условий образовательная программа альтернативного формата будет способствовать реализации основных задач современного образования — развитию академической мобильности, укреплению сотрудничества и межвузовских связей в сфере высшего образования, повышению каче- ства и привлекательности высшего образования.

**Литература**

1. *Атвелл Г.* Персональная среда обучения. URL: goo.gl/fJhpLC (дата обра- щения 10.08.2016).
2. *Виландеберк А.А., Шубина Н.Л.* (2015), Образовательные программы в сете- вом формате: анализ преимуществ и рисков. Научно-методический комментарий для преподавателей. СПб.
3. *Виландеберк А.А., Шубина Н.Л.* (2011), Формирование образовательного кластера: сетевое взаимодействие и академическая мобильность. СПб.
4. *Виландеберк А.А., Шубина Н.Л.* (2010), Проектирование модульной обра- зовательной программы на базе кластер-технологий. СПб.
5. *Б. Дендев* (ред.) (2013), Информационные и коммуникационные технологии в образовании: монография. М.

**Literature**

1. *Atvell G*. Personal’naya sreda obucheniya. URL: goo.gl/fJhpLC (data obrash- cheniya 10.08.2016).
2. *Vilandeberk A.A., Shubina N.L*. (2015), Obrazovatel’nye programmy v setevom formate: analiz preimushchestv i riskov. Nauchno-metodicheskiy kommentariy dlya prepodavateley. SPb.
3. *Vilandeberk A.A., Shubina N.L*. (2011), Formirovanie obrazovatel’nogo klastera: setevoe vzaimodeystvie i akademicheskaya mobil’nost’. SPb.
4. *Vilandeberk A.A., Shubina N.L.* (2010), Proektirovanie modul’noy obrazovatel’noy programmy na baze klaster-tekhnologiy. SPb.
5. *B. Dendev* (red.) (2013), Informatsionnye i kommunikatsionnye tekhnologii v obrazovanii: monograﬁya. M.

*Научное издание*

**ПРИКЛАДНАЯ ЛИНГВИСТИКА В НАУКЕ И ОБРАЗОВАНИИ**

**ALPAC REPORT — ПОЛВЕКА ПОСЛЕ РАЗГРОМА**

**Труды VIII Международной научной конференции 24–26 ноября 2016 г.**

Публикуется в авторской редакции с оригинал-макета, предоставленного авторами

ООО «Книжный Дом», лицензия № 05377 от 16.07.2001 191186, Санкт-Петербург, М. Конюшенная ул., д. 5

Подписано в печать 18.11.2016 . Формат 60×84/16. Бумага офсетная.

Объем 20,5 печ. л. Тираж 200 экз. Заказ №

Отпечатано в типографии ООО «Инжиниринг Сервис» 190020, Санкт-Петербург, ул. Циолковского, д. 13