

Применение стохастической теории обучения и корпусных технологий в компьютерной лингводидактике

Using Stochastic Learning Theory and Corpus Tools in CALL

В.Р. НЫММ¹
Voldemar Nymm¹
jetnomm@gmail.com

К.Р. Пиотровская¹
Xenia Piotrowska¹
krp62@mail.ru

С. НЫММ²
Sven Nõmm²
sven.nomm@ttu.ee

¹Российский государственный педагогический университет
им. А. И. Герцена,
Санкт-Петербург, Российская Федерация

²Таллинский технологический университет, Таллинн, Эстония

¹ Herzen State Pedagogical University of Russia,
Saint Petersburg, Russian Federation

² Department of Software Science School of Information Technology,
Tallinn University of Technology, Tallinn, Estonia

Abstract

The paper considers solving the CALL-problems in the framework of computing experiment and statistics methods. Special attention is drawn to developing the methodological, informational and organizational components of a CALL system. The managing and control algorithms for learning process are based on methods and models of stochastic learning theory. Education process is based on a single-element model. The system is supported by free corpus manager tools.

Keywords: *Computing experiment framework, free corpus manager tools, computer assisted language learning, stochastic learning theory, statistics methods*

Аннотация

Обсуждается проблема компьютерного обучения языку с точки зрения применения статистико-вероятностных методов и организации вычислительного эксперимента. Особое внимание уделено развитию методологии, информационной и организационным компонентам. Реализуемые программой алгоритмы организации

и управления процессом обучения основаны на моделях и методах стохастической теории обучения. Обучение строится на базе одноэлементной модели. Предлагается методика отбора и замены содержания обучения при поддержке свободными корпусными сервисами.

Ключевые слова: *Вычислительный эксперимент, свободные корпусные технологии, компьютерная лингводидактика, стохастическая теория обучения*

1 Введение

В науке хорошо известны примеры словосочетаний со “счастливой судьбой”. Возникнув для обозначения явления или понятия одной области исследований, они (то ли в силу предвосхищения будущего развития, то ли в силу каких-то других неизвестных причин) с годами расширяют свой ареал применения и включают в себя новую семантику понятия, расширяют объекты и области исследования и разработки. К числу таковых относится термин “компьютерная лингводидактика”, введенный в начале 90-х гг. в работах межрегиональной группы “Статистика Речи”, возглавляемой профессором Р.Г. Пиотровским, для обозначения широкого круга вопросов, связанных с использованием компьютерных технологий в учебном процессе [Piotrovskaya, 1991]. Поскольку специализация группы заключалась в создании систем автоматической переработки текста, то и компьютерные лингводидактические разработки планировалось создавать на больших лингвистических базах систем машинного перевода в виде лексико-грамматических справочных служб широкого профиля [Piotrovskaya 1993]. Одной из отличительных черт современного этапа развития компьютерной лингводидактики является не столько развитие элементной и программной компьютерной базы, сколько постоянный рост изменчивости лексического и лексико-грамматического состава профессионально ориентированных языков, обусловленного возникновением междисциплинарных областей знаний и расширением объемов информационного обмена. Это требует:

а) организации механизма постоянной актуализации содержания языкового обучения студентов в соответствии с их профессиональной ориентацией;

б) интенсификации процесса обучения на основе тех широких возможностей, которые предоставляют современные информационные технологии.

Решение этих задач обусловило всплеск интереса к статистическим методам исследования. Применение статистических методов к решению задач отбора содержания обучения не является чем-то новым. Идеи и методы применения статистико-вероятностных методов при выборе терминологических единиц профессионально ориентированных языков активно разрабатывались международной группой “Статистика речи” уже начиная с шестидесятых годов XX века. В значительной степени эти идеи и методы опережали свое время. Несмотря на то, что существовавшие на тот момент электронные средства не могли обеспечить сколько-нибудь существенной поддержки автоматизации при разработке новых учебников и многое приходилось делать, что называется “вручную”, были созданы хорошие профессионально-ориентированные учебные пособия [Golovacheva 1978, German-Prozorova et al. 1982]. В настоящее время, не утративший своей актуальности статистико-вероятностный подход [Lazdin 2013, Lazareva 2013, Sergeevnina 2011, Pavlovskaya et al. 2015], реализуется на новых, универсальных платформах, причем большинство западных исследователей трактуют наличие в современ-

ных системах ранее известных лингво-статистических моделей как элементы текст-майнинга [Scriver et al. 2017, Anthony et al. 2013].

Этот объект применения статистико-вероятностных методов связан с наличием в сети Интернет многочисленных корпусов текстов по самым разнообразным профессиональным тематикам и открытым сервисов, предназначенных для выделения устойчивых словосочетаний и контекстов их употребления, являются эффективным инструментарием для формирования содержания обучения профессионально ориентированных лингводидактических комплексов на основе лингво-статистических моделей и методов.

Альтернативный объект применения статистико-вероятностных методов относится к интенсификации языкового обучения на базе компьютерных средств и состоит в разработке эффективных алгоритмов управления процессом обучения. Основные трудности создания таких алгоритмов связаны с отсутствием адекватных знаний об этих процессах и достаточно опосредованных связях между основными критериями задачи (качеством обучения, временем обучения, устойчивостью полученных знаний). Для их преодоления междисциплинарная рабочая группа Российского государственного педагогического университета им. А.И. Герцена развивает исследования по проблеме компьютерного обучения языку, основные положения и результаты которых отражены в ряде работ [Nymm et al. 2017].

2 Постановка и пути исследования проблемы

Вне зависимости от конкретной задачи обучения, центральную роль в ее постановке играют взаимодействие трех перечисленных выше критериев. Поскольку затрачиваемое на обучение время является величиной, которую при любой организации процесса практически невозможно измерить, в качестве ее адекватного представления нами рассматривается суммарное количество стимулов упражнений, предъявленных обучаемому в ходе предшествующего обучения. Минимизацию введенного заменителя времени обучения естественно рассматривать как экстремальную цель, а качество обучения и устойчивость полученных знаний – как ограничения на решение задачи.

В качестве основного инструмента моделирования процесса обучения рассматривается вектор параметров состояния обучаемого, значение каждой i -ой координаты которого в любой момент времени выражает степень обученности обучаемого i -му языковому явлению. Его использование позволяет в определенном смысле формализовать понятие качества обучения. При этом структура функционала или вектор-функции, количественно представляющая это понятие и зависящая от типа решаемой задачи, синтезируется исследователем, а в роли аргументов выступают координаты вектора параметров состояния обучаемого. Это дает возможность представить ограничения на решение задачи в форме системы целевых неравенств.

Отсутствие возможности явного выражения связи между экстремальной целью и двумя другими критериями и, тем самым, найти точное решение оптимизационной задачи, переводит процесс ее решения в область построения стратегии, обеспечивающей максимально быструю сходимость координат вектора параметров состояния обучаемого к значениям, при которых выполнялись бы целевые неравенства. Каждая такая стратегия строится на основе априорных знаний и интуитивных представлениях о процессах обучения. Чем выше степень априорного знания, тем больше возможностей для построения алгоритма, близкого по своей осмысленности к “гипотетически оптимальному”.

Сравнение алгоритмов, реализующих различные стратегии, и функционирование механизма их постоянного развития осуществляется путем статистической обработки данных о реакциях обучаемых в ходе обучения и по его завершении.

3 Аппарат исследования

В качестве базового элемента исследований выступает программа компьютерного обучения и тестирования (CALL-программа). Она объединяет в одной оболочке функциональные модули, реализующие практически все основные типы упражнений типа *стимул - реакция*, на основе которых строятся современные учебники иностранного языка. Программа предназначена для использования студентами как при выполнении заданий учебного плана по изучению иностранного языка, так и при выполнении ими (в рамках курсовых и выпускных квалификационных работ) научных и научно-педагогических исследований по проблеме компьютерного языкового обучения. Для выполнения исследовательских функций программа снабжена средствами сбора данных о реакциях обучаемого в ходе обучения.

Процесс компьютерного обучения реализуется в ходе последовательных сеансов и завершается тестированием. При этом все функции управления процессом обучения, включая:

- а) выбор порции упражнений для каждого очередного сеанса и каждого очередного стимула для предъявления студенту в ходе проведения сеанса,
- б) определение момента завершения каждого сеанса и момента завершения процесса обучения в целом, осуществляется программой.

Более того, тот факт, что доля программного кода, реализующего алгоритм управления процессом (при любой его сложности), достаточно невелика относительно общего объема кода программного продукта в целом, дает возможность одновременно использовать и нормально сопровождать несколько программ с разными алгоритмами управления,

Тем самым, роль обучаемого в ходе сеанса сводится к выполнению заданий, предлагаемых ему программой, и ожиданию сообщения о том, что сеанс обучения успешно завершён. Реализуемые программой алгоритмы управления процессом обучения основаны на моделях и методах стохастической теории обучения [Atkinson et al. 1969, Bush et al. 1973]. Достаточно полное описание одного из таких алгоритмов приведено в [Nyumm et al. 2015].

Проверка эффективности этого алгоритма оценивалась в ходе эксперимента по изучению 100 языковых клише общенаучной английской лексики (устойчивые словосочетания, вводные конструкции и т.п.). Участниками эксперимента были 50 студентов филологического факультета. Обучение выполнялось в ход самостоятельной работы студентов. Интервал времени между сеансами колебался у разных студентов в пределах от 24 до 72 часов. После обработки данных были получены следующие результаты. Среднее число предъявлений стимулов (по группе в целом), которое потребовалось для завершения обучения по одному упражнению, оказалось равным 4,64 (разброс значений: от 2.72 до 7.81, стандартное отклонение – 1.57.). Среднее число сеансов, в ходе которых эти предъявления реализовывались, оказалось достаточно малым – всего 1.91 (разброс значений: от 1.51 до 2.37, стандартное отклонение – 0.32). По сути дела, это означает, что несмотря на существенные различия в числе требуемых для этого предъяв-

лений, для выучивания правильного ответа практически всем обучаемым требовалось не более двух сеансов. Среднее значение доли правильных ответов обучаемых в ходе контрольного тестирования, проведенного по всему материалу после завершения обучения, так же представляется достаточно оптимистичным – 0.779 (разброс значений: от 0.61 до 0.9, стандартное отклонение – 0.0832).

Методология исследования. В методологическом плане программа ориентирована на использование в рамках вычислительного эксперимента. Под термином *вычислительный эксперимент* здесь понимается нами как *“циклический процесс, реализующий эволюционную стратегию построения модели исследуемого объекта или процесса, каждую итерацию которой составляют три последовательно реализуемые стадии: анализ, синтез и оценка, а проверка адекватности модели, синтезированной в ходе каждой очередной итерации, осуществляется путем сравнения результатов натуральных экспериментов с результатами, полученными с помощью компьютера путем вычислений по модели”* [Belyaeva et al. 2013].

Идея вычислительного эксперимента исходит из того, что, приступая к моделированию объекта или процесса, мы всегда имеем, пусть и приближенное, представление о нем, которое далее можно последовательно улучшать на основе данных, получаемых в ходе экспериментов. В нашем случае вычислительный эксперимент рассматривается как механизм поиска эффективных алгоритмов обучения различным аспектам грамматики и лексики иностранного языка. Для использования программы в рамках вычислительного эксперимента ее структура была спроектирована так, чтобы обеспечить достаточно простую (т.е. без сколько-нибудь существенных перестроек программы) замену одних алгоритмов управления процессом обучения другими.

В совокупности со встроенными в программу средствами сбора информации о ходе процессов обучения, которыми программа снабжена, реализация этого условия обеспечивает механизм параллельного развития как самой программы, так и проводимых с ее помощью исследований:

а) каждое организованное (т.е. проводимое либо в рамках выполнения студентами заданий учебного плана, либо в исследовательских целях) использование программы рассматривается как элемент массового эксперимента;

б) получаемые в ходе каждой итерации технологического цикла вычислительного эксперимента данные о ходе обучения обрабатываются и преобразуются в новые знания о процессах компьютерного обучения иностранным языкам;

в) на основе полученных знаний ставятся новые задачи исследования, создаются новые и корректируются существующие алгоритмы управления процессом обучения, создаются новые базы обучающей информации и корректируется состав и объем существующих;

г) производится замена одних алгоритмов управления другими и осуществляется переход к следующей итерации технологического цикла. Таким образом, в любой момент времени программа работает на основе алгоритма, соответствующего текущим знаниям обучаемого и текущим представлениям о процессах компьютерного обучения.

Отбор содержания обучения. Наряду с организацией процесса обучения и управлением, другой важной составляющей эффективности является оперативный отбор и замена содержания обучения. Оптимизация процесса формирования иноязычной или языковой профессиональной компетенции сопряжена, с одной стороны, со все ускоря-

ющимися темпами информационного обмена, который ведет к изменчивости лексико-грамматического состава профессионально ориентированного языка, с другой, - с появлением новых открытых сервисов, позволяющих осуществлять автоматизированное статистическое исследование текста и систем, помогающих оперативно формировать базы языковых упражнений в области лексики и грамматики. Объединенная группа исследователей из Японии и Новой Зеландии создала ряд открытых и надежных сервисов, под названием AntLab, которые активно используются в корпусных исследованиях [Anthony et al. 2015, Nation et al. 2016]. На наш взгляд, часть из этих сервисов полезна для организации квантитативного отбора содержания обучения для CALL-программы. Целевое назначение отобранных программ, входящих в семейство программ AntLab, отражено в таблице 1.

Таблица 1: Программное обеспечение для отбора содержания обучения на базе программ семейства AntLab

№ п/п	Название программы	Целевое назначение
1.	AntConc	конкордансер
2.	AntPConc	организатор параллельных текстов
3.	AntMover	анализатор семантических блоков текста
4.	AntWordProfiler	профилирование словаря и оценка сложности текста
5.	ProtAnt	инструмент для поиска файла прототипа
6.	TagAnt	инструмент для частеречного теггирования текста

Предварительно проводится обход наиболее известных в мире библиотек и наукометрических баз данных с помощью программы AntCorGen с целью отбора исследовательского материала. Затем программами AntFileConverter, EndCodeAnt и VariAnt, проводится перевод текстов в нужный формат, проверка правописания и удаление незначимых для анализа символов. С помощью программы AntConc строится частотный словарь, выделяются ключевые слова и выявляются устойчивые словосочетания с ключевыми словами. Объем исследуемых текстов далее наращивается путем использования программы ProtAnt по введенным ключевым словам, 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 необходимо 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 производится фрагментация и структурно-семантическая разметка текстов по предложениям с сохранением каждого предложения в отдельный файл. Далее программой ProtAnt снова обрабатываются уже файлы отдельных предложений и выявляются те из них, которые по лексическому или грамматическому составу подходят для кодирования упражнений в CALL – программу.

Организация вычислительного эксперимента. В организационном плане реализация вычислительного эксперимента опирается на метод /it learning by doing (обучение в деле) и рассматривается как часть высокотехнологичной образовательной среды,

в рамках которой осуществляется интеграция учебной и созидательной деятельности студентов в области информационных технологий, направленная на создание компонентов информационного и методического обеспечения систем языкового компьютерного обучения, с одной стороны, с другой, – на то, чтобы в процессе этой деятельности студенты научились:

а) использовать программу языкового компьютерного обучения и тестирования для выполнения заданий учебного плана;

б) выполнять на их базе научные и научно-педагогические исследования;

в) обрабатывать полученные данные о ходе обучения, интерпретировать результаты.

г) создавать (на основе существующего программного обеспечения) собственные системы компьютерного обучения.

Такой подход, с одной стороны, придает обучению четко выраженный прикладной характер, позволяет продемонстрировать применение информационных технологий для выполнения полной технологической цепочки операций моделирования в области, относящейся к сфере основной специализации студентов, с другой, – дает возможность использовать созидательную практическую деятельность студентов для реализации двух важных функций вычислительного эксперимента:

а) создания, тестирования и корректировки баз обучающей и тестовой информации по иностранным языкам;

б) выполнения лабораторных экспериментов, в роли участников которых выступают сами студенты.

Обе эти функции реализуются в ходе практических занятий и самостоятельной работы магистрантов первого курса в рамках дисциплины “Информационные технологии в профессиональной деятельности”. Первая из перечисленных функций позволяет избежать достаточно типичных ситуаций, когда при полной готовности к выполнению эксперимента практически невозможно найти участников (тем более мотивированных) для его реализации. Еще более очевидна важность второй функции. Как показывают результаты внедрения подхода, интеллектуальный ресурс, который студенты составляют в совокупности, и временной ресурс, выделяемый учебным планом на самостоятельную работу студентов, достаточны для создания и актуализации электронных баз обучающей и тестовой информации.

4 Выводы

Нами рассмотрены принципиальные моменты подхода к исследованию проблемы компьютерного языкового обучения, развиваемого междисциплинарной рабочей группой университета. Методологической основой подхода является вычислительный эксперимент, в рамках которого реализуется механизм постоянного саморазвития системы компьютерного обучения как в части программного, так и в части информационного обеспечения.

Многообразии грамматических структур, степень сходства или различия средств выражения одних и тех же явлений в русском и иностранных языках, лексические различия и многое другое разбивает проблему на множество задач, каждая из которых является предметом отдельного исследования. Их решение, с одной стороны, требует значительных интеллектуальных ресурсов, с другой, – является прекрасной основой

для формирования на базе метода *learning by doing* (обучение в деле) научно исследовательских компетенций студентов в рамках выпускных квалификационных работ и диссертационных исследований.

Дальнейшие шаги по развитию системы предполагают:

а) совершенствование алгоритмов управления процессом обучения и расширение программного обеспечения системы до уровня сетевой технологии;

б) разработку процедур выделения (на базе текст-майнинговых сервисов и корпусных технологий) метаинформационных структур (маркеров и шаблонов разной степени детализации), характерных для текстов той или иной профессиональной (например, математической) направленности;

в) формирование на основе этих процедур баз метаинформационных структур (*phrase banks*) и упражнений по обучению использованию структур в рамках системы языкового компьютерного обучения.

Список литературы

[Anthony et al. 2015] Anthony L., Baker P. (2015) ProtAnt: A Freeware Tool for Automated Prototypical Text Detection. In F. Formato and A. Hardie (Eds). Proceedings of Corpus Linguistics 2015 Lancaster, UK: Lancaster University. Pp. 24-26.

[Anthony et al. 2013] Anthony L., Bowen M. (2013) The Language of Mathematics: A Corpus-based Analysis of Research Article Writing in a Neglected Field. Asian ESP Journal, 2013, 9(2). Pp. 5-25.

[Atkinson et al. 1969] Atkinson R., Bower G., Krothers E. (1969) Vvedenie v matematicheskuyu teoriyu obucheniya, M.: Mir, 1969. – 487 s. [An introduction to mathematical learning theory] (In Russian) = Аткинсон Р., Бауэр Г., Кротерс Э. Введение в математическую теорию обучения, М.: Мир, 1969. – 487 с.

[Belyaeva et al. 2013] Belyaeva L.N., Dzhepa T.L., Zak G.N., Kamshilova O.N., Nymm V.R., Razumova V.V. (2013) Avtomatizirovannoe rabochee mesto filologa v strukture obrazovatel'nogo prostranstva vuza. Monografiya [Philologist's automatic workstation in a university educational space structure. Monograph] SPb: ООО «Knizhnyi Dom». – 127 s. (In Russian) = Беляева Л.Н., Джепа Т.Л., Зак Г.Н., Камшилова О.Н., Нымм В.Р., Разумова В.В. Автоматизированное рабочее место филолога в структуре образовательного пространства вуза. Монография СПб: ООО «Книжный Дом», 2013. – 127 с.

[Bush et al. 1973] Bush R., Mosteller F. (1973) Srvnjenje vos'mi modelei [Comparison of eight models] // Matematicheskie metody v social'nyh naukah. [Mathematical methods in social sciences] M.: Izd. Fizmatgiz, – 300 s. (In Russian) = Буш Р., Мостеллер Ф. Сравнение восьми моделей // Математические методы в социальных науках. М.: Изд. Физматгиз, 1973. – 300 с.

[German-Prozorova et al. 1982] German-Prozorova L. P., Ivanov V. G. Uchebnik angliiskogo yazyka: dlya studentov vysshih inzhenernyh morskikh uchilish. [English for students of higher engineering maritime school] Moskva : Vysshaya shkola, 1982. - 367 s. (In Russian)

= Герман-Прозорова Л. П., Иванов В. Г. Учебник английского языка: для студентов высших инженерных морских училищ. Москва : Высшая школа, 1982. – 367 с.

[Golovacheva 1978] Golovacheva A.K. (1978) Chastotnyi kurs uskorennoho obucheniya angliiskomu yazyku po profilyu radioelektroniki. [Frequency course for English accelerated training for radio electronics profile] L.: Izd-vo LGU, – 239 s. (In Russian) = Головачева А.К. Частотный курс ускоренного обучения английскому языку по профилю радиоэлектроники. Л.: Изд-во ЛГУ, 1978. – 239 с.

[Lazareva 2013] Lazareva M. (2013) Corpus-based foreign language learning. Pedagogic corpora and Web collaboration // Materialy Vtoroi mezhdunarodnoi nauchno-prakticheskoi konferencii «Aksiologicheskii aspekt sodержaniya nepreryvnoho inoyazychnoho obrazovaniya: problemy i resheniya» [Материалы Второй международной научно-практической конференции «Аксиологический аспект содержания непрерывного иноязычного образования: проблемы и решения»] [Axiological aspect of the continuous foreign language education content: problems and solutions. Proceedings of international scientific conference], Москва, 31 января - 2 февраля 2013 г. [Npreryvnoe pedagogicheskoe obrazovanie.ru] (In Russian), № 3. S. 96 .

[Lazdin 2013] Lazdin' T.A. (2013) Osnovy statisticheskoi optimizacii prepodavaniya inostrannyh yazykov [Bases of statistical optimization for foreign language teaching] // Vestnik Sankt-Peterburgskogo universiteta. Ser. 9. 2009. Vyp. 3. (In Russian) = Лаздинь Т.А. Основы статистической оптимизации преподавания иностранных языков // Вестник Санкт-Петербургского университета. Сер. 9 Вып. 3. 2009. С. 202-208

[Nation et al. 2016] Nation P., Anthony L. (2016) Measuring vocabulary size // Handbook of Research in Second Language Teaching and Learning, Volume III, E. Hinkel (Ed.) New York: Routledge. – 514 p.

[Nymm et al. 2015] Nymm V.R. Evolyucionnyi podhod k sozdaniyu i razvitiyu CALL-tehnologii na baze vychislitel'nogo eksperimenta [Evolutional approach to creation and development of CALL-technologies on the basis of a computing experiment] // Izvestiya Rossiiskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gercena [Proceedings of Herzen State Pedagogical University of Russia], № 173. SPb.: Izd. «Izvestiya RGPU», 2015. S. 137-147. (In Russian) = НЫММ В.Р. Эволюционный подход к созданию и развитию CALL-технологий на базе вычислительного эксперимента // Известия Российского государственного педагогического университета им. А.И. Герцена, № 173. СПб.: Изд. «Известия РГПУ», 2015. С. 137-147.

[Nymm et al. 2017] Nymm V.R., Piotrovskaya K.R., Makogon I.S., Ereemeeva Yu.P. Sozdanie i razvitie sredstv elektronnoi podderzhki obucheniya inostrannomu yazyku na baze vychislitel'nogo eksperimenta [Creation and development of electronic support tools for computing experiment-based foreign language learning] // Aktual'nye problemy sovremennoi prikladnoi lingvistiki, sbornik nauchnyh statei, posvyashennyi 80-letiyu doktora filologicheskikh nauk, professora A.V. Zubova. [Actual problems of theoretical and applied Linguistics. Proceedings of international scientific conference in honor of professor A.V. Zubov] Ministerstvo obrazovaniya Respubliki Belarus', Minskii gosudarstvennyi lingvisticheskii universitet. Minsk, 2017. S. 308-312. (In Russian) = НЫММ В.Р., Пиотровская К.Р., Макогон И.С., Еремеева Ю.П. Создание и развитие средств элек-

тронной поддержки обучения иностранному языку на базе вычислительного эксперимента // Актуальные проблемы современной прикладной лингвистики. сборник научных статей, посвященный 80-летию доктора филологических наук, профессора А.В. Зубова. Министерство образования Республики Беларусь, Минский государственный лингвистический университет. Минск, 2017. С. 308-312.

[Pavlovskaya et al. 2015] Pavlovskaya I.Yu., Gorina O.G. Korpusno-kognitivnye metody obucheniya leksicheskoi storone rechi na inostrannom yazyke [Corpora and cognitive methods of lexical training in a foreign language] // //Kognitivnye issledovaniya yazyka. [Cognitive Studies of Language] 2015. № 22. S. 552-554. (In Russian) = Павловская И.Ю., Горина О.Г. Корпусно-когнитивные методы обучения лексической стороне речи на иностранном языке //Когнитивные исследования языка, 2015. № 22. С. 552-554

[Piotrovskaya 1993] Piotrovskaya K.R. (1993) Modeli, programmnye i informacionnye sredstva uchebnogo ARM perevodchika. Dissertaciya na soiskanie uchenoi stepeni kandidata tehniceskikh nauk [Models. Program and information facilities of the AWS for a translator training. Thesis on candidate of technical sciences degree] (In Russian), Kiev, 1993. – 142 s. = Пиотровская К.Р. Модели, программные и информационные средства учебного АРМ переводчика. Диссертация на соискание ученой степени кандидата технических наук. Киев, 1993. – 142 с.

[Piotrovskaya, 1991] Piotrovskaya K.R. (1991) Sovremennaya komp'yuternaya lingvodidaktika [Modern computer linguodidactics] // Nauchno-tehnicheskaya informaciya (NTI). Seriya 2. Informacionnye processy i sistemy [Scientific and Technical Information/ Series 2. Information processes and systems] 1991. № 4. S. 26 (In Russian) = Пиотровская, К.Р. Современная компьютерная лингводидактика // Научно-техническая информация (НТИ). Серия 2. Информационные процессы и системы, № 4. С. 26

[Sergeevnina 2011] Sergeevnina V.M. (2011) Iz opyta sostavleniya uchebnyh terminologicheskikh slovarei dlya studentov neyazykovykh vuzov [Some experience in building educational terminological dictionaries for technical university students] // Vestnik Nizhegorodskogo universiteta, [Nizhegorodski University Herald]. S. 624-627 (In Russian) = Сергеевнина В.М. Из опыта составления учебных терминологических словарей для студентов неязыковых вузов // Вестник Нижегородского университета, 2011. С. 624-627

[Scrivner et al. 2017] Scrivner O. I., Trapido J., Lee. Text mining toolkit for digital corpora // Труды международной конференции “Корпусная лингвистика - 2017”, 27–30 июня 2017 г., [Proceedings of the International Conference “Corpora - 2017”, Saint Petersburg, June 27–29, 2017] Санкт-Петербург. S. 72-77