

Карамнов А. С. Модель создания корпуса учебника английского языка / А. С. Карамнов // Научный диалог. – 2013. – № 2(14) : Педагогика. – С. 59–69.

УДК 372.881.111.1+81'33+371.671:81'243

Модель создания корпуса учебника английского языка

А. С. Карамнов

В статье рассматриваются принципы создания корпуса учебника английского языка, которые предлагается разделить на содержательные и конструирующие. Следование указанным принципам призвано обеспечить успешную реализацию стратегии построения корпуса учебника в соответствии с целью и задачами конкретного исследования.

Ключевые слова: учебник английского языка; корпусная лингвистика; языковой корпус; педагогический корпус.

Постановка проблемы. Корпусные технологии анализа и обработки массивов текстовых данных являются новым способом организации и проведения языковых исследований, в том числе экспертного характера [Мильруд и др., 2012]. Для того чтобы можно было осуществить надежную и валидную процедуру оценивания учебных материалов с помощью таких технологий, необходимо создать корпус учебника или учебного пособия, релевантного проблемной области конкретного исследования. При правильном применении метода формирования корпусов данных в практике проведения лингвистической экспертизы учебных материалов удастся повысить степень объективности получаемых результатов и существенно облегчить работу лингвистов.

Языковой корпус и процедура его формирования. Прежде чем приступить к созданию языкового корпуса учебника, необходимо определиться с тем, что это такое и какие факторы влияют на эту процедуру.

Лингвистический корпус – это объемное, критериально обоснованное и хранящееся в электронном виде собрание естественных случаев использования языка. Критериальная обоснованность корпуса заключается в том, что его содержание (языковое наполнение) никогда не случайно. Степень спланированности корпусного содержания меняется в зависимости от его объема; и даже многомиллионные корпуса, главным критерием создания которых является полнота репрезентации языковой системы, не смогут обеспечить присутствие абсолютно всех случаев языкового употребления. Критериальная обоснованность лингвистического корпуса находит свое отражение и в существовании различных типов корпусов – от фундаментальных, отражающих состояние языка в целом, а не каких-то его отдельных фрагментов и подсистем, до авторских, включающих коллекцию текстов одного автора.

В типологии языковых корпусов выделяют среди прочих педагогические и исследовательские корпуса.

Педагогические корпуса (pedagogic corpora) основаны на языке, используемом в учебном процессе, и могут включать в себя академические учебники и учебные пособия, скрипты записей коммуникации, происходящей во время уроков, а также любые другие письменные тексты или скрипты устного языка, которые учащиеся встречают в окружающей их образовательной среде [Bennett, 2010]. Следует пояснить, что термин «педагогический корпус» объединяет в себе содержательную область не только языковой педагогики, но и лингвистической дидактики.

Исследовательские корпуса широко используются в практике лингвистических экспертиз и предназначены преимущественно для

изучения различных аспектов функционирования языковой системы [Баранов, 2009]. Подтипами этих корпусов являются объектный корпус (тот, на который направлены действия, предусмотренные практикой данной экспертной области в соответствии с поставленными задачами) и различные фоновые (или референтные) корпуса (подмножества текстов, с которыми происходит сравнение объектного корпуса).

Таким образом, *корпус учебника представляет собой частный случай педагогического исследовательского корпуса*, поэтому процедура его создания будет определяться факторами, влияющими на создание тех более широких типов корпусов данных, к которым он относится.

Содержательные принципы создания корпуса учебника английского языка. Процесс построения корпуса учебника английского языка связан с проблемой выделения текста, который следует включить в его состав, поэтому здесь следует придерживаться нескольких содержательных принципов: принципа целостности, принципа соответствия проблемной области, а также принципа структурной направленности.

Принцип целостности заключается в ориентации на определенную категорию учебников в соответствии с исследовательскими требованиями. Западные учебники, например, в основе своей организации имеют определенный набор тем, вокруг которых строится вся учебная деятельность, подбираются задания, «подгоняется» содержание и т. д. Языковое наполнение таких многоуровневых курсов (elementary, pre-intermediate, intermediate, upper-intermediate, advanced) основывается на прототипическом подходе [Ghsoon, 2003], когда с каждым уровнем расширяется лексическое поле той или иной темы. На рисунке 1 дан пример реализации такого подхода в учебниках серии «Kid's Vox» [Nixon et al., 2008a; Nixon et al., 2008b]. В центральном круге задана тематика проблемной области,

в то время как последующие два звена содержат лексические единицы по теме «Время обедать», встречающиеся в учебниках для первого и второго классов соответственно.

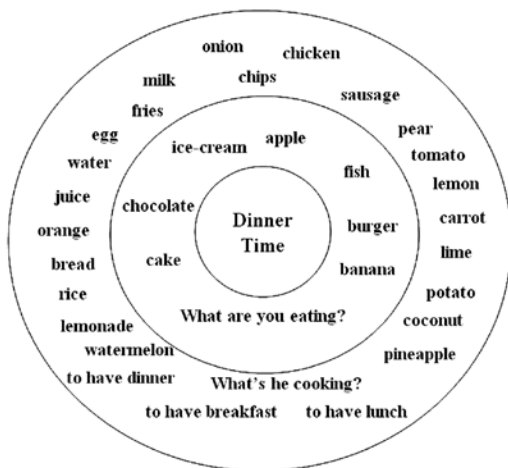


Рис. 1. Пример реализации прототипического подхода к лексической наполняемости учебников курса «Kid's Vox»

Таким образом, в соответствии с принципом целостности и в зависимости от поставленных в ходе исследования задач создаваемый корпус учебника может включать в себя как содержание книг всей уровневой серии, так и материал отдельных пособий. Если речь идет об анализе линейки УМК, то корпус может составляться на основе каждой отдельной книги, либо учебников одного звена (начального, среднего или старшего), либо всей линейки. Кроме того, может быть составлен корпус отдельных компонентов УМК (рабочей тетради, скриптов мультимедиа-приложений и т. д.) или всего комплекта.

Принцип соответствия проблемной области ориентирует создателя корпуса на определение границ языкового пространства проблемной области, которое планируется отобразить в корпусе.

При этом сообразно цели и задачам исследования корпус учебника английского языка может содержать все его текстовое наполнение, включая название самого издания и уроков, ссылки на словарные статьи, грамматический справочник и т. д., либо ограничиваться определенной выборкой данных.

Такая выборка может проводиться, например, по видам речевой деятельности, в результате чего будет сформирован корпус заданий и текстов на чтение, говорение, аудирование (включая аудио-скрипты), письмо. Другим критерием, по которому отбирается материал, является его соответствие компонентам коммуникативной компетенции (лингвистическому, дискурсивному, прагматическому, стратегическому, культурному) [Мильруд и др., 2012].

Принцип структурной направленности состоит в том, что стратегия создания языкового корпуса учебника ориентирована на компонентный состав книги. Структура учебника не раз рассматривалась в теории и практике обучения иностранным языкам, в частности в работах Д. Д. Зуева [Зуев, 1983], В. Г. Бейлинсона [Бейлинсон, 1986], Э. Г. Гельфман, М. А. Холодной [Гельфман и др., 2006].

Так, согласно указанному принципу может быть смоделирован: а) корпус основного текста учебника английского языка, который будет содержать материал, служащий ключевым источником учебной информации; б) корпус дополнительных текстов, который будет включать в себя материал, предназначенный для закрепления и углубления знаний (например, тексты для домашнего чтения); в) корпус пояснительных текстов, объединяющий различные примеры, справки, примечания, словарные статьи и другой материал, призванный обеспечить более полное понимание и усвоение информации.

Кроме этого, здесь можно говорить и о построении корпуса учебных заданий (задач и упражнений), корпуса аппарата организации усвоения (АОУ) (заголовков, планов, параграфов, образцов, памяток, подписей к рисункам и схемам, рубрик типа «Progress

Check» и т.п.), корпуса аппарата ориентировки (АО) (предисловий, введений, оглавлений, рубрик, указателей и проч.).

Итак, содержательные принципы создания корпуса учебника английского языка, подчеркивающие идею целостности учебного материала в многоуровневых и линейных курсах, идею соответствия корпусного содержания конкретной проблемной области, а также идею градуированной значимости структурных компонентов издания, исполняют роль ориентиров в планировании стратегии построения языкового корпуса.

Конструирующие принципы построения корпуса учебника английского языка. Процедура создания корпуса учебника английского языка основывается на ряде конструирующих принципов, среди которых можно выделить принцип жанровой обусловленности, принцип репрезентативности, принцип аутентичности, принцип анализируемости и принцип соблюдения авторских прав.

Принцип жанровой обусловленности основан на понимании того, что содержание корпуса не может быть случайным и отбирается в соответствии с определенными критериями. Это особенно важно при проведении узкопрофильных исследований. Например, если целью является изучение вводных слов в академической речи, то важно обращаться к корпусу, который будет состоять только из фрагментов академических текстов.

Принцип репрезентативности является главным для обеспечения допустимости использования корпуса текстов. В статистике под репрезентативностью понимается соответствие характеристик, полученных в результате выборочного наблюдения, показателям, характеризующим всю генеральную совокупность. В корпусной лингвистике можно говорить о количественной и качественной моделях репрезентативности.

Количественная модель репрезентативности строится на основе частоты явлений в проблемной области. Это означает, что соз-

даваемый корпус должен отражать все свойства проблемной области (например, язык учебника английского языка) в определенной пропорции. Пропорциональная репрезентативность может быть достигнута пропорциональным сужением проблемной области или с помощью стратифицированной выборки [Баранов, 2009]. Пропорциональное сужение подразумевает вероятность того, что какие-то явления проблемной области со сравнительно низкой частотой употребления не попадут в корпус данных. При стратифицированной выборке учитываются все содержательные компоненты проблемной области, но при этом каждый из них представлен одинаковым объемом данных. Существуют и другие вероятностные методики оценки репрезентативности, в частности модель накопления и коррекции, при использовании которой частота отбираемых примеров употребления определенного явления при последовательном накоплении материала должна приближаться к относительной частоте его встречаемости в проблемной области, что графически будет изображаться движением кривой к некоторому пределу.

Качественная модель репрезентативности основывается на параметризации проблемной области, которая задает верхнюю и нижнюю границу всех возможных сочетаний корпусных характеристик. Реализация этой модели, скорее всего, не даст стопроцентного результата, но позволит максимально учесть все составляющие проблемной области. Так, при параметризации в процессе создания

$$\text{учебное задание} = \left\{ \begin{array}{l} 0001 \\ 0010 \\ 0100 \\ 0011 \\ 0110 \\ 1100 \\ 0111 \\ 1110 \\ 1111 \end{array} \right\}$$

Рис. 2. Матрица сочетаний значений параметров корпуса

корпуса формируется список релевантных параметров, сочетание значений которых образует матрицу (рис. 2).

Затем на каждую возможную комбинацию отыскивается пример, который и становится частью корпуса.

Например, если применить такую модель репрезентативности к корпусу заданий для учебника английского языка, то к его релевантным параметрам можно будет отнести следующие:

- а) целевая установка (задача, упражнение);
- б) вид речевой деятельности (чтение, письмо, говорение, аудирование, перевод);
- в) синтаксическая структура задания (повелительное предложение с глаголом в форме 2-го лица, повелительное предложение в форме 1-го и 3-го лица, повествовательное предложение, вопросительное предложение и т. д.);
- г) наличие в составе задания парентезы-отсылки (например, the pictures of Exercise 5 can help you) и другие возможные параметры.

Так, например, задание “Read and say which story is true to the picture” займет одну из строк в матрице со следующим набором параметров: <задача; на чтение и говорение; повелительное предложение с глаголом в форме 2-го лица; без парентезы-отсылки>.

Описанные выше репрезентативные модели, однако, используются гораздо реже сплошного отбора данных ввиду их технологической усложненности, так как вероятностные модели подразумевают непосредственные количественные измерения, а качественные модели нередко требуют участия экспертных групп для оценки существенности параметров проблемной области [Баранов, 2009]. При сплошном отборе в корпус вносятся все соответствующие его назначению и целям исследования данные, что имеет свои сложности в последующей оценке репрезентативности корпуса такой модели.

Принцип аутентичности диктует включение в состав корпуса только аутентичных материалов. В связи с этим стоит отметить, что понятие аутентичности часто довольно широко интерпретируется. Так, содержание учебника может быть аутентичным по своему происхождению, когда он написан носителями языка, аутентичным по своим свойствам, когда написанное автором-носителем языка не отличается от написанного в естественно-языковой среде, и аутентичным по своим функциям, когда материал естественно вписывается в учебный процесс независимо от авторства [Millrood, 2010]. Поэтому главным критерием аутентичности текста учебника следует считать то, что его языковое наполнение не было специально придумано с единственной целью наполнения корпуса.

Принцип анализируемости подразумевает возможность автоматизированного анализа корпусных данных программными средствами. Для этого корпус должен быть сохранен в одном из электронных форматов (.txt, .rtf, .html и проч.) в зависимости от корпусного менеджера (программы компьютерной поддержки), который будет использоваться для выполнения задач по обработке текстовых массивов.

Принцип соблюдения авторских прав является важной составляющей процесса создания корпуса данных, особенно если его содержание предполагает дальнейшее распространение в виде раздаточного материала или публикации. В связи с этим необходимо быть в курсе правил предоставления разрешения на использование материала Интернет-сайтов и издательств, брать письменные разрешения участников учебного процесса и любых лиц, чья работа планируется быть использованной при формировании языкового корпуса.

Конструирующие принципы построения корпуса учебника английского языка, таким образом, объединяют в себе ряд требований, выполнение которых формирует стратегию этого построения. Поэтому важно найти правильный баланс между полнотой явлений

проблемной области, представленной в корпусе, и их репрезентативностью, упорядочиванием текстовых данных, форматами их хранения и наличием программных средств их обработки. Не менее важно учитывать и принципы жанровой обусловленности и соблюдения авторских прав.

Заключение. Для решения задач в сфере оценивания аутентичности и релевантности содержания учебников английского языка могут оказаться эффективными инструменты корпусного языкового анализа. Процедура построения корпуса учебника английского языка основывается на содержательных принципах целостности, соответствия проблемной области, структурной направленности и конструирующих принципах жанровой обусловленности, репрезентативности, аутентичности, анализируемости и соблюдения авторских прав. На этих основаниях и формируется стратегия создания корпуса учебника английского языка с учетом цели конкретного исследования и поставленных задач.

Литература

1. Баранов А. Н. Лингвистическая экспертиза текста : теория и практика : учебное пособие / А. Н. Баранов. – 2-е изд. – Москва : Флинта : Наука, 2009. – 592 с.
2. Бейлинсон В. Г. Арсенал образования : характеристика, подготовка, конструирование учебных изданий / В. Г. Бейлинсон. – Москва : Книга, 1986. – 286 с.
3. Гельфман Э. Г. Психодидактика школьного учебника : интеллектуальное воспитание учащихся / Э. Г. Гельфман, М. А. Холодная. – Санкт-Петербург : Питер, 2006. – 384 с.
4. Зуев Д. Д. Школьный учебник / Д. Д. Зуев. – Москва : Педагогика, 1983. – 218 с.
5. Мильруд Р. П. Введение в корпусную лингвистику : учебное пособие / Р. П. Мильруд, П. В. Сысоев, А. В. Кашеева, А. С. Карамнов. – Тамбов : Издательский дом ТГУ им. Г. Р. Державина, 2012. – 129 с.
6. Мильруд Р. П. Компетентность учителя иностранного языка / Р. П. Мильруд, А. С. Карамнов // Иностранные языки в школе. – 2012. – № 1. – С. 11–17.

7. *Bennett G. R. Using Corpora in the Language Learning Classroom : Corpus Linguistics for Teachers / G. R. Bennett. – Ann Arbor, MI : University of Michigan Press, 2010. – 134 p.*

8. *Ghsoon R. English Coursebooks : Prototype Texts and Basic Vocabulary Norms / R. Ghsoon // ELT Journal. – 2003. – Vol. 57 (3). – P. 260–268.*

9. *Millrood R. Theory of Language Teaching : Linguistics, Didactics, Pedagogy / R. Millrood. – Saarbrücken : Lambert Academic Publishing, 2010. – 144 p.*

10. *Nixon K. Kid's Box : Pupil's Book 1 / K. Nixon, M. Tomlinson. – Cambridge : Cambridge University Press, 2008a. – 79 p.*

11. *Nixon K. Kid's Box : Pupil's Book 2 / K. Nixon, M. Tomlinson. – Cambridge : Cambridge University Press, 2008b. – 96 p.*

© Карамнов А. С., 2013

Model of English Textbook Corpus Creation

A. Karamnov

The article covers principles of creating an English textbook corpus. It is suggested that they should be divided into informative and constructive. The mentioned principles following serves to provide successful realization of the textbook corpus design strategy in accordance with the case study purposes.

Key words: English textbook; corpus linguistics; language corpora; pedagogic corpus.

Карамнов Антон Сергеевич, аспирант кафедры лингвистики и лингводидактики, Тамбовский государственный университет имени Г.Р. Державина (Тамбов), k.a.s.89@mail.ru.

Карамнов, А., post-graduate student, Department of Linguistics and Lingvodiadactics, Tambov State University named after G. R. Derzhavin (Tambov), k.a.s.89@mail.ru.