

ИСПОЛЬЗОВАНИЕ МЕТОДОВ КОРПУСНОЙ ЛИНГВИСТИКИ
В КУРСЕ ИНОСТРАННОГО ЯЗЫКА ДЛЯ СПЕЦИАЛЬНЫХ ЦЕЛЕЙ

Актуальность. Обучение студентов технического профиля иностранному языку является одной из наиболее трудных, но в то же время и одной из самых востребованных задач современной высшей школы. Учебники иностранных языков, применяемые в ВУЗах, обнаруживают серьёзные недостатки, связанные с «универсальностью» таких пособий, отсутствием специальной направленности материалов, что, в свою очередь, приводит к невостребованности полученных в ходе занятий иностранным языком знаний выпускниками ВУЗов. Необходим поиск методов и инструментов изучения иностранных языков, способных помочь сократить разрыв между существующей системой обучения иностранному языку и действительными нуждами учащихся [1].

Одним из способов решения данной проблемы является использование методов корпусной лингвистики. Иностранные исследования показывают, что данные корпусов могут служить более ценным справочным материалом для учащихся, чем словари или заранее подготовленные преподавателем списки слов и выражений. Например, Д. Гаскель и Т. Кобб показали, что использование учащимися корпуса для самостоятельного исправления своих ошибок в письменных работах приводит к положительным результатам [2]. Однако, необходимо отметить, что корпусные методы могут быть эффективны только в том случае, если корпус является достаточно репрезентативным, т.е. содержит достаточное количество примеров использования языка в определенном дискурсе, представляющем интерес для учащегося [3]. Таким образом, можно говорить о необходимости разработки корпусов для специальных целей, проблема которых на настоящий момент не решена ни для английского, ни для немецкого языка. Более того, анализ существующих публикаций выявил почти полное отсутствие материалов по работе с корпусами на немецком языке [4].

Цели и задачи. Санкт-Петербургский университет им. Петра Великого совместно с Ганноверским университетом им. Лейбница разрабатывают трехязычный корпус *the Kod.ING corpus*, представляющий собой собрание текстов диссертаций на немецком, русском и английском языках по различным техническим специальностям, который после завершения проекта будет доступен в онлайн-режиме для целевых пользователей обоих университетов. Данный корпус призван стать инструментом совершенствования существующих программ обучения студентов вышеназванных университетов по иностранному языку. Целью нашего исследования является изучение возможностей немецкоязычной части данного корпуса и оценка потенциала его использования в качестве инструмента преподавателя немецкого языка. Это обусловлено не только наибольшей разработанностью данной части корпуса, но и растущим интересом к немецкому языку среди учащихся ВУЗов России в связи с развитием международных программ обучения, в частности, предусматривающих прохождение обучения и стажировок в Германии.

Современные исследования в области корпусной лингвистики показывают, что, вопреки сложившемуся мнению, носители языка, как в устной, так и в письменной речи, тяготеют к использованию одних и тех же линейно организованных сочетаний лексических единиц [5]. Такого рода сочетания не являются фразеологизмами, а представляют собой часто повторяющиеся цепочки слов, не всегда обладающие конструктивной законченностью и/или общим для всех единиц, выходящих в их состав, значением. Например, «per cent of the», «for the first time», «at the end of the», «it has been (shown/observed/argued) that». Наиболее разработана тема таких словосочетаний

англоязычными исследователями. Тем не менее, в англоязычной литературе не существует общепринятого термина и единогласия в определении таких словосочетаний. В рамках нашего исследования мы будем использовать предложенный Д. Бибером термин «lexical bundles», который раскрывается им как «последовательность слов, часто встречающуюся в определенном стиле речи» [6]. В отечественных публикациях линейно организованным сочетаниям лексических единиц почти не уделяется внимания и, следовательно, отсутствует общепринятое название для данного явления. Поэтому вслед за Н.В. Денисовой, Е.С. Петровой мы будем использовать термин-кальку «лексические пучки» [7].

Одной из **основных задач** на пути достижения цели нашего исследования является нахождение лексических пучков на материале немецкого языка. Первым шагом на пути выявления таких сочетаний, является выделение наиболее часто употребляемых в корпусе лексических единиц. По утверждению М. МакКарти составление списка таких единиц уже может предоставить достаточно информации для выделения характерных особенностей дискурса корпуса [5].

Материалом для нашего исследования послужила часть корпуса the Kod.ING corpus, содержащая собрание полных текстов диссертаций (с частеречевой разметкой) по специальностям «Электротехника», «Гражданское строительство» и «Машиностроение» на немецком языке (далее – Корпус). Все диссертации были разделены на три группы, которые, в свою очередь были озаглавлены нами «Электротехника» (Эл), «Гражданское строительство» (ГС) и «Машиностроение» (М). Группа «Гражданское строительство» состоит из 43-х, группа «Электротехника» - из 35-ти, группа «Машиностроение» - из 127-ми диссертаций. С помощью инструмента «WordList» программы для составления конкордансов «AntConc», разработанной профессором Университета Васэда Лоренсом Энтони и находящейся в свободном доступе (<http://www.laurenceanthony.net/software/antconc/>), нами были составлены списки наиболее частотной лексики данных групп диссертаций по отдельности, а также список наиболее часто встречающихся лексических единиц во всех трех группах. Каждый список состоит из 100 лексических единиц. Проведенное нами сравнение данных списков имело своей **целью** определение наиболее часто встречающихся видов лексических единиц, а также оценку перспектив их использования для выявления лексических пучков в данном корпусе.

Сводная таблица 1 содержит количественную информацию о лексических единицах, относящихся к различным частям речи как по каждой группе диссертаций так и по Корпусу в целом.

Таблица 1

	Арт.	Арт.+ предлог	Союз	Предлог	Гл.	Местоим.	Част.	Сущ.	Мод. гл.	Нар.	Прил.	Числ.
ГС	12	6	16	12	12	11	1	18	2	9	1	1
Эл	12	12	9	17	14	15	1	14	3	11	3	1
М	13	5	9	21	12	10	1	18	2	8	2	1
К	12	6	9	21	12	9	1	17	3	9	2	1

Так, можно утверждать, что числительные и частицы, представленные только лексемами «*zwei*» и «*nicht*» соответственно, а также модальные глаголы и прилагательные, являются наименее частотными лексическими единицами как отдельных групп, так и всего Корпуса в целом. Наиболее же часто встречаются слова, относящиеся к таким частям речи, как предлог, существительное, артикль и глагол. В отдельную графу нами

были вынесены лексические единицы, представляющие собой своего рода сокращения от артикля и предлога.

Каждая группа диссертаций обнаруживает свои особенности по частоте использования определенных групп лексических единиц. Например, в группе диссертаций «Гражданское строительство» 18 слов из 100 наиболее частотных являются существительными, а в двух других группах, а также в списке частотных лексем всего Корпуса такие слова встречаются в 2 раза реже.

Лексический состав всех четырех списков сходен. Существующие различия обусловлены особенностями тематики каждой группы диссертация. Например, слово «*Parameter*» (параметр) входит в число наиболее часто встречающихся лексем группы «Гражданское строительство», но не представлено в остальных группах, а также не вошло в число наиболее частотных слов всего Корпуса диссертаций. Аналогичным образом, слово «*Temperatur*» (температура) представлено только в списке группы «Машиностроение».

Для проверки работоспособности избранного нами подхода к определению лексических пучков с помощью инструмента «Clusters/N-Grams» программы «AntConc» нами были составлены списки наиболее часто повторяющихся последовательностей лексических единиц с элементами «*Parameter*» и «*Temperatur*». Проведенный анализ позволяет утверждать, что большинство таких последовательностей представляют собой двусловные сочетания, причем чаще прочих встречаются сочетания с артиклем, предлогом и глаголом, стоящими как перед, так и после существительного. Например, *für Parameter, Parameter von, Parameter werden, Temperatur des, ein Parameter*. Наибольший интерес представляют сочетания данных существительных с прилагательными (*folgende Parameter, maximale Temperatur*), а также трехсловные сочетания (*mit steigender Temperatur, die unbekanntes Parameter*), которые, учитывая их частую повторяемость в текстах корпуса, могут быть отнесены к лексическим пучкам.

Выводы. Проанализированные нами списки наиболее часто встречающихся лексических единиц в группах диссертаций по специальностям «Электротехника», «Гражданское строительство» и «Машиностроение», а также по всем трем группам в целом, обнаруживают определенные особенности частотной лексики данного корпуса, такие как преобладание слов таких частей речи как предлог, существительное, артикль и глагол, а также единообразие своего лексического состава. Отобранные нами лексические единицы могут быть использованы для нахождения лексических пучков в материалах Корпуса с последующим использованием этих результатов в учебных материалах в курсе немецкого языка для специальных целей.

ЛИТЕРАТУРА:

1. Горина О.Г. Использование технологий корпусной лингвистики для развития лексических навыков в профессионально-ориентированном общении на английском языке : дис. ... канд. пед. наук. СПб., 2014. - 242 с.
2. Gaskell D., Cobb T. Can learners use concordance feedback for writing errors // *System*. V.32б №3. – 2004. – P. 301–319.
3. Chang J.-Y. The use of general and specialized corpora as reference sources for academic English writing: A case study / J.-Y. Chang // *ReCALL* – 2014. – V. 26– № Special Issue 02–P. 243–259..
4. Boulton A. Learning outcomes from corpus consultation / A. Boulton // *Exploring New Paths in Language Pedagogy: Lexis and Corpus-Based Language Teaching*. – 2010. P.129–144.
5. McCarthy M. From Corpus To Classroom / M. McCarthy, A. O’Keeffe, R. Bott – , 2007.– 315p.
6. Biber D. If you look at ...: Lexical bundles in university teaching and textbooks / D. Biber, S. Conrad, V. Cortes // *Appl. Linguist.* – 2004. – V. 25 – № 3– P. 371–405.
7. Денисова Н.В., Петрова Е.С. Маркеры ряда: кластерность и вариантность (на материале английского языка) // *Вестник Челябинского государственного университета* – 2008. – Т. 30. С 44–49.

(Автоматическая нумерация списка литературы не допускается!)