

*На правах рукописи*

**ВАДЯЕВ Сергей Евгеньевич**

**ЛИНГВИСТИЧЕСКИЕ ПРИНЦИПЫ ПОСТРОЕНИЯ  
И ИСПОЛЬЗОВАНИЕ КОРПУСА ТЕКСТОВ ДЛЯ  
ИССЛЕДОВАНИЯ ОФИЦИАЛЬНО-ДЕЛОВОГО СТИЛЯ  
СОВРЕМЕННОГО НЕМЕЦКОГО ЯЗЫКА**

**(на материале электронного корпуса «DER»)**

Специальность 10.02.04 – Германские языки

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата филологических наук

Нижний Новгород – 2005

Работа выполнена на кафедре немецкой филологии Самарского государственного университета.

Научный руководитель: доктор филологических наук,  
профессор Сергей Иванович Дубинин

Официальные оппоненты: доктор филологических наук,  
профессор Ефим Григорьевич Вышкин

кандидат филологических наук,  
доцент Николай Евдокимович Бажайкин

Ведущая организация: Волгоградский государственный  
университет

Защита состоится «18» мая 2005 г. в 14 часов на заседании диссертационного совета К 212.163.01 по защите диссертаций на соискание ученой степени кандидата филологических наук при Нижегородском государственном лингвистическом университете им. Н.А.Добролюбова по адресу: 603115, г. Нижний Новгород, ул. Минина, д. 31а, научный читальный зал.

С диссертацией можно ознакомиться в библиотеке Нижегородского государственного лингвистического университета им. Н.А.Добролюбова.

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2005 г.

Ученый секретарь  
диссертационного совета



В.В. Денисова



тивная оценка их частотности. Только такой подход может обеспечить достоверное выявление функционально-коммуникативных особенностей данного ТТ и его актуальных характеристик в немецком литературном языке последних десятилетий.

Для сравнения квантифицируемых признаков текстов (например, для выявления их принадлежности к функциональному стилю или к типу текстов), а также для определения и сравнения лексического многообразия текстов в современной стилистике активно используются количественные методы. Их применение для стилистического анализа текстов имеет долгую традицию (У. Пипер, У. Шмитц, А.А. Пиотровская, А.Я. Шайкевич). Как правило, в качестве основы для таких исследований служит *корпус*, репрезентативная выборка текстов, составленная с учетом таких факторов, как функционально-стилистические особенности текстов, письменная или устная форма их существования. Общая коммуникативная направленность конкретных ТТ в значительной мере влияет на употребление языковых средств и их частотность. Таким образом, частотные характеристики могут служить в качестве одного из дифференциальных признаков для отнесения группы текстов к определенному функциональному стилю, их выделения в отдельный ТТ.

Использование количественных методов для стилистического анализа предполагает исследование обширных массивов текстов, поэтому для их применения большое значение приобретает подготовка репрезентативных выборок текстовых данных и разработка средств доступа к ним, для чего всё чаще используются *электронные (машинные) корпуса текстов*. Они позволяют более эффективно, надёжно и широко использовать количественные методы, в частности для выявления стилистических особенностей входящих в их состав текстов. При этом особое внимание уделяется выявлению тенденций частотности и дистрибуции отдельных языковых единиц и их групп, стохастических зависимостей между ними, а также соотношений между внутренними характеристиками текста и внешними по отношению к нему факторами (У. Шмитц, Х. Бергенхольц, Л. Хоффман, У. Пипер).

В рамках данной диссертации был разработан оригинальный электронный корпус текстов «*DER*» (*Deutsch-Russisches Parallelkorpus*), послуживший текстовой и инструментальной базой для выявления особенностей функционирования немецких модальных глаголов в текстах официально-делового стиля. Он представляет собой *параллельный корпус текстов* (далее ПКТ), включающий немецко- и русскоязычные варианты документов ООН общим объёмом

мом в 300 000 словоупотреблений. Использование ПКТ «DER», помимо преимуществ одноязычного корпуса при изучении отдельного языка, дает реальные возможности для проведения контрастивных изысканий, исследования проблем передачи различных языковых явлений в разных языках, поиска использующихся в переводческой практике эквивалентов, а также для составления и пополнения переводных словарей и обучения иностранному языку.

**Теоретической основой** исследования послужили работы отечественных и зарубежных лингвистов по функциональной стилистике (В. Фляйшер, Г. Михель, М.П. Брандес, Э.Г. Ризель, Е.И. Шендельс, В.И. Прворотов, М.Н. Кожина), грамматике (Г. Хельбиг, И. Буша, Г. Цифонун, Л. Хоффман, О.И. Москальская), а также корпусной лингвистике (Д.О. Добровольский, А.Н. Баранов, В.М. Андрищенко, М.Н. Михайлов, С. Йоханссон, Т. МакЭри, Э. Уилсон).

**Методологической основой** построения ПКТ «DER» стали концепции разработанные лингвистами при создании параллельного немецко-финского корпуса «Finde» университета г. Вюрцбург, ФРГ ([www.finde.de](http://www.finde.de)), англо-норвежского параллельного корпуса «ENPC» университета г. Осло, Норвегия (<http://www.hf.uio.no/iba/prosjekt/ENPCmanual.html>).

**Итогом** данной работы являются два теоретически и практически значимых результата, первый из которых состоит в обосновании необходимости и разработке принципов создания ПКТ «DER», а второй – в выявлении модальной специфики немецких текстов официально-делового стиля и особенностей функционирования в них модальных глаголов.

**Актуальность** данного диссертационного исследования определяется, с одной стороны, необходимостью выявления специфики функционирования модальных глаголов в текстах различных функциональных стилей немецкого языка, а, с другой стороны, ограниченным количеством параллельных корпусов текстов вообще и для пары языков немецкий и русский в частности, а также недостаточной разработкой концепции и технологии их составления.

**Научная новизна** предлагаемой диссертации состоит в том, что в ней детально описана лингвистическая концепция и разработана технология построения немецко-русского параллельного корпуса текстов, а также наглядно представлены возможности использования одного из субкорпусов ПКТ для стилистического анализа. В рамках работы детально представлены принципы построения ПКТ, а также осуществлено практическое его приложение для

выявления особенностей употребления модальных глаголов в одном из типов текстов официально-делового стиля современного немецкого языка (официальные доклады Генсекретаря ООН).

**Теоретическая значимость** исследования состоит в обосновании зависимости между частотностью средств реализации категории модальности (на примере модальных глаголов немецкого языка) и основной коммуникативной функции текстов официально-делового стиля (на примере документов ООН). Данная работа рассматривает также широкий круг проблем, касающихся особенностей официально-делового стиля в современном немецком языке, места документов ООН как самостоятельного типа текстов в динамичной системе функциональных стилей современного немецкого литературного языка. Несмотря на то, что официально-деловой стиль является многоплановым феноменом и в спектре политических типов текстов проявляет в последние десятилетия значительную динамику, связанную с расширением международной функции немецкого языка (А.И. Домашнев), за последнее время появилось не так много работ, посвященных данной тематике<sup>1</sup>. Кроме того, в диссертации подробно разрабатывается теоретическая база для моделирования ПКТ, а также обосновываются возможности применения ПКТ «DER» для анализа стилистических характеристик ТТ.

**Материалом исследования** послужила подборка электронных документов ООН ([www.un.org](http://www.un.org)) на немецком и русском языках, входящих в состав ПКТ «DER». Первоначальную основу корпуса составляют ежегодные доклады Генерального Секретаря о работе ООН за 1997-2001 гг. Общий объем текстов насчитывает около 900 страниц (примерно 300 тыс. словоупотреблений). ПКТ «DER» может служить не только в качестве источника статистических данных или для поиска эквивалентов, но и в качестве базиса для исследований явлений одного из языков корпуса. Так, например, немецкие тексты ПКТ «DER» могут служить в качестве самостоятельного объекта специализированного исследования, что наглядно продемонстрировано в рамках данной работы.

<sup>1</sup> См., в частности, диссертационные исследования. Комлева Е.В. Лингвистические особенности выражения побудительности в текстах официально-деловой прозы (на материале современного немецкого языка), АКД. - СПб.: РГПУ им. Герцена, 2003; Исакова Л.Д. Соотношение типологических тенденций и особенностей проявления функционального стиля в современных германских языках, АДД. - М.: МГЛУ, 2001.

**Цель данного исследования** заключается в выявлении модальной специфики немецких текстов официально-делового стиля и особенностей функционирования в них модальных глаголов на базе репрезентативного корпуса текстов. Для решения поставленной задачи была разработана оригинальная технология составления немецко-русского ПКТ «DER» и продемонстрированы возможности его оптимального применения для выявления стилистических особенностей текстов официально-делового стиля (определение параметров поиска и обработка его результатов).

На первом этапе работы ставилась **общая задача** разработать технологию автоматической обработки текстов для ПКТ, а также средств их представления и последующего языкового анализа. На втором этапе планировалось выявить на основе данных **квантитативного и качествентивного стилистического анализа немецкого субкорпуса ПКТ «DER» закономерности частотно-распределительных характеристик немецких модальных глаголов и их зависимости от общей коммуникативно-функциональной направленности представленного в ПКТ «DER» типа текстов** (ежегодные доклады Генерального Секретаря).

**Практическая ценность исследования:** ПКТ «DER» уже в своём первоначальном виде может служить в качестве основы и инструментария для дальнейших стилистических, лексикологических, грамматических и контрастных исследований, в частности, позволяя:

- осуществлять *стилометрический анализ* различных стилистически релевантных явлений на основе субкорпусов ПКТ;
- проводить сравнительный анализ особенностей немецкого и русского языков, относящихся к лексическому составу, морфологии и синтаксису (статистический и сравнительный анализ лексических структур, синтаксических конструкций в параллельных текстах и др.);
- осуществлять поиск переводных соответствий для слов и коллокаций, а также примеров употребления лексических единиц и их перевода на другой язык в зависимости от контекста употребления;
- проводить исследования в области стилистики (рассмотрение способов передачи стилистических особенностей в текстах на разных языках, приемов и стилистических средств, при помощи которых в языке перевода передается содержание оригинального текста).

В качестве инструментария были избраны следующие **методы исследования**: квантитативный метод для определения частотности квантифицируемых единиц в корпусе и выявления соотношений между внутренними характеристиками текста и внешними по отношению к нему факторами; качественный метод для интерпретации полученных статистических данных, для семантического и грамматического анализа употреблений исследуемых языковых единиц в текстах корпуса, а также методы и новейшие технологии корпусной лингвистики для реализации и обработки ПКТ «DER».

**Объектом** данного исследования является стилистическая система современного немецкого литературного языка, а в качестве **предмета** исследования выступают частотно-распределительные характеристики немецких модальных глаголов в текстах официально-делового стиля (документы ООН).

#### **Положения, выносимые на защиту:**

1. Важнейшим лингвистическим принципом построения корпуса текстов является обеспечение его репрезентативности. Значительное увеличение объема исследуемого корпуса не означает пропорционального роста количества уникальных лексем в нём (т.е. «автоматического» повышения репрезентативности). Динамика роста количества словоформ падает вместе с ростом количества словоупотреблений в корпусе, поэтому его репрезентативность должна обеспечиваться не столько за счет увеличения объёма, сколько за счет тщательной выборки текстов.
2. Статистические данные о структурных элементах текста могут служить в качестве надежной основы для выявления параллельных сегментов в текстах на немецком и русском языках и использоваться при разработке параллельных корпусов текстов.
3. Документы ООН занимают особое место среди политических типов текстов в составе официально-делового стиля современного немецкого языка, что обусловлено их *внешнеполитическим* характером, особым статусом их эмитента, усиливающейся интернационализацией немецкого языка.
4. Между *общей коммуникативно-функциональной направленностью* текста и основными средствами выражения *модальности* существует непосредственная зависимость, выражающаяся для модальных глаголов немецкого языка в изменении их частотно-распределительных характе-



ристик в зависимости от функционально-стилевой принадлежности текста.

5. *Речевой фон* употребления отдельных модальных глаголов немецкого языка и выбор синтаксических конструкций, в которых они употребляются, тесно связаны с *общей коммуникативно-функциональной направленностью* текстов определенного типа.

**Апробация результатов исследования:** основные положения работы были представлены в виде докладов на Международных научных конференциях «*Korpuslinguistik deutsch: synchron – diachron – kontrastiv*» (2003 г.) и «*Europatag*» (2004 г.) в университете г. Вюрцбург (ФРГ), «*Компаративистика: современная теория и практика*» (2004 г.) в Самарском государственном педагогическом университете, на конференции «*Русско-немецкие культурно-исторические контакты*» в Тольяттинском государственном университете (2004 г.), обсуждались на заседаниях кафедры немецкой филологии СамГУ.

**Структура и объем диссертации:** диссертация (объем 160 стр.) состоит из введения, трех глав, заключения, списка использованной научной литературы (180 единиц), списка входящих в ПКТ «*DER*» текстов, глоссария и приложения, которое включает CD-ROM с электронной версией ПКТ «*DER*» для Windows.

### Основное содержание работы

Во введении обосновывается актуальность исследования, определяется цель работы, её теоретическая значимость и практическая ценность, определяются методы исследования, характеризуется его материал, а также формулируются положения, выносимые на защиту.

В первой главе «*Корпус текстов – новый тип источника лингвистического исследования*» даётся определение понятия «корпус текстов», теоретически обосновывается использование электронных корпусов текстов в современной лингвистике (германистике), проводится критический анализ существующего опыта в области составления корпусов текстов, рассматриваются основные их типы и их назначение.

Использование машинных корпусов текстов открыло новые возможности для проведения лингвистических исследований, а разнообразные современные технологии позволяют существенно увеличить объем и повысить эффективность обработки языковых данных. Сам по себе корпус текстов не несет

новой информации о языке, являясь «копилкой», куда складываются тексты, репрезентирующие язык в своем естественном употреблении. При помощи компьютера можно эффективно управлять содержащейся в корпусе информацией и ее представлением для исследователя. Полученные на основании анализа корпуса данные являются релевантными в первую очередь для используемого корпуса, но, тем не менее, они могут быть обобщены для оценки явлений в языке в целом.

Термин «корпус» является неоднозначным и служит для обозначения самых разнообразных объектов, нередко утрачивая строгость в употреблении (ср., «корпус высказываний/примеров употребления» и др.). Общим является то, что во всех случаях речь идет о некоей подборке языкового материала. Как правило, имеется в виду подборка полных текстов или их достаточно крупных фрагментов<sup>2</sup>. Корпус текстов следует отличать от подборки примеров, поскольку в данном случае из используемых текстов более или менее произвольно извлекаются отдельные «цитаты» и условие содержательной когерентности не выполняется (К. Бауш, Т. МакЭнери, Э. Уилсон).

Нет единства мнений и по поводу того, должен ли корпус содержать только тексты, которые являются результатом спонтанных, естественных ситуаций коммуникации, или данный термин применим и к текстам, искусственно возникшим вне акта языковой коммуникации. Поскольку в дискуссии об эмпирических данных, собираемых для анализа языка, центральную роль играет противопоставление корпуса и интуиции, неуместно обозначать термином «корпус» набор придуманных (возникших вне акта коммуникации) примеров. Высказывания, возникшие в результате коммуникативного намерения, по многим причинам нельзя приравнивать к примерам, которые искусственно создает в рамках лингвистической дискуссии считающий себя «компетентным» носитель языка. Неоднозначными являются также данные, полученные в результате неспонтанного акта коммуникации (М.Н. Михайлов, Х. Бергенхольц, И. Мугдан).

Мы придерживаемся следующего определения: *корпусом* является ограниченное количество языкового материала, состоящего исключительно из (письменных или устных) текстов или их частей. Под текстом в данном слу-

<sup>2</sup> Андрущенко В. М. Автоматическая обработка текста // Лингвистический энциклопедический словарь / Ред. В.Н.Ярцева - М.: Советская энциклопедия, 1990 - С. 14-15.

чае понимается относительно самостоятельная, содержательно связанная последовательность естественных языковых высказываний, а под частью текста – его связанный, сплошной фрагмент.

Принципы построения любого корпуса текстов напрямую зависят от его практического назначения, что определяет принципы сбора и представления данных в корпусе. Выделяются следующие основные типы корпусов: *общие, специализированные, сопоставительные, учебные, педагогические, исторические (диахронические), мониторинговые и параллельные корпуса* (С. Ханстон). Тип корпуса задаёт параметры исследования и определяет эвристическую значимость полученных при этом данных.

Во второй главе «Разработка и построение электронного корпуса текстов „DER“» формулируются основные принципы составления параллельных корпусов. В качестве основных рассматриваются проблемы автоматического определения границ структурных единиц текста (абзацев и предложений) и различные подходы к автоматической стыковке параллельных текстов. Кроме того, в данной главе характеризуется программная оболочка для представления текстов ПКТ «DER» и возможные способы поиска в них.

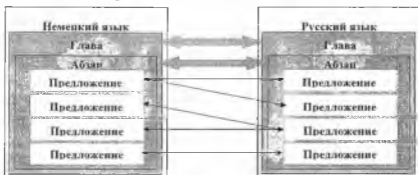
При составлении ПКТ, в отличие от одноязычных и сравнительных корпусов текстов, следует учитывать фактор устойчивых межкультурных связей. Множество текстов исходного языка, пригодных для включения в ПКТ, составляют лишь тексты, которые были переведены на второй язык. Подбор текстов для сбалансированного русско-немецкого ПКТ является сложной задачей, что связано с проблемой обнаружения текстов, аналогичных по тематике, жанру и относящихся к одному периоду времени. Хотя для пары языков *немецкий / русский* существует большое количество переводов, различия в культурных традициях приводят к дисбалансу в структуре субкорпусов, а получить «зеркальные» субкорпуса оригинальных и переводных текстов для ПКТ часто проблематично.

Важнейшей проблемой при составлении ПКТ является *стыковка* параллельных текстов (*alignment*). Решение данной задачи для ПКТ «DER» было разделено на два этапа: выявление границ структурных элементов текста и определение параллельных сегментов. Основой для выявления параллельных сегментов в текстах на немецком и русском языках служили статистические данные, отражающие структурное сходство между оригиналом и переводом, проявляющееся на разных уровнях текста в различной степени. Для стыковки

текстов ПКТ «DER» была создана специализированная программа, использующая для определения параллельных сегментов статистико-стохастический подход, что обеспечило высокую надежность результатов стыковки при минимальных временных затратах. Как показали результаты работы, статистические данные о структурных элементах текста могут служить в качестве надежной основы для выявления параллельных сегментов в текстах на немецком и русском языках. При стыковке параллельных текстов сначала стыковались более крупные сегменты текста, а затем происходила проверка правильности и начиналась обработка малых сегментов на следующем уровне. Такой подход позволил сократить количество ошибок в процессе автоматической стыковки (см. рис. 1).

Рисунок 1

Схема процесса стыковки текстов в ПКТ «DER»



Для наглядного представления текстов ПКТ «DER» и осуществления эффективного поиска в них была создана специальная программная оболочка (см. рис. 2). Программа предназначена для работы с тагированными параллельными текстами с лингвистической разметкой в формате XML, соответствующей международному стандарту TEI (TEI Lite, версия для XML, см.: [www.tei-c.org](http://www.tei-c.org)). Стандарт разметки TEI представляет собой набор предложений по кодировке электронных текстов различных типов именно для лингвистических целей. Особое значение при разработке программной оболочки придавалось следующим функциям: эффективный поиск в текстах корпуса, возможность различного представления его результатов (параллельный кон-



Кроме богатых поисковых возможностей, ПКТ «DER» позволяет получать различные статистические данные. Они могут использоваться как для проведения контрастных исследований, так и для изучения отдельных языков корпуса, а также служить для выявления особенностей текстов, входящих в него. Так, программная оболочка ПКТ «DER» позволяет определить соотношение длины текста в словоупотреблениях к количеству словоформ (type/token ratio, TTR), которое нередко используется в качестве одного из критериев богатства лексического состава представленных в корпусе текстов. Соотношение TTR может значительно изменяться в зависимости от типов текстов, входящих в корпус. Однако это не единственный фактор, который следует учитывать при создании корпуса. На примере ПКТ «DER» было установлено, что *значительное увеличение объема корпуса не ведет к прямо пропорциональному росту количества уникальных лексем в корпусе*. Как видно из данных, представленных на диаграмме 1, динамика увеличения количества словоформ падает вместе с ростом количества словоупотреблений в корпусе. При увеличении объема (т.е. количества словоупотреблений) немецкой части ПКТ «DER» более чем в 8 раз количество словоформ увеличивается примерно в 3,5 раза. Это позволило ограничить первоначальный состав ПКТ «DER» шестью крупными текстами на немецком языке на основе принципа оптимальной достаточности выборки.

Диаграмма 1

Динамика увеличения соотношения количества словоупотреблений к количеству словоформ (types-tokens ratio) при увеличении объема ПКТ «DER»



Даже при значительном увеличении объема корпуса (т.е. количества словоупотреблений) количество словоформ во всем корпусе повышается лишь не-

значительно. Это позволяет сделать вывод о том, что даже значительное повышение объема корпуса не означает пропорционального роста количества уникальных лексем, встречающихся в текстах корпуса, что было важно для определения *оптимального* объёма исследуемого ПКТ «DER». Тщательная выборка текстов позволяет на основе достаточно компактного корпуса получать достоверные данные для всех текстов данного типа, отвечающих заданным параметрам отбора.

В третьей главе «Использование ПКТ «DER» для исследования особенностей официально-деловых текстов (на материале документов ООН)» даётся обзор использования количественных методов в стилистике, определяются актуальные проблемы исследования текстов официально-делового стиля, описываются основные характеристики входящих в ПКТ «DER» текстов. На основе представленных в корпусе текстов рассматриваются особенности функционирования модальных глаголов в текстах официально-делового стиля.

Документы ООН, входящие в состав ПКТ «DER», представляют интерес для изучения особенностей официально-делового функционального стиля в современном немецком языке и для опытного подтверждения предположений, касающихся особенностей таких текстов. Использование корпуса даёт возможность быстро и надёжно получить результаты обработки большого массива текстов, что обеспечивает получение более точных данных. На основе немецких текстов ПКТ «DER» были рассмотрены особенности частотно-распределительных характеристик модальных глаголов и определены наиболее употребительные значения модальных глаголов в текстах ООН.

Значения возможности, необходимости, разрешения, запрета и желания выражаются в немецком языке в основном при помощи модальных глаголов. Они являются одним из средств выражения модальности в немецком языке, которое вносит весомый вклад в формирование общей интенции всего текста, и играют важную роль в выражении его функционально-коммуникативных параметров. Согласно нашей гипотезе, они также имеют особое значение для реализации в текстах официально-делового стиля на немецком языке прагматической функции *долженствования* и выражения их *предписывающе-оценочного характера*. Опираясь на описанные в грамматиках немецкого языка значения модальных глаголов (Г. Хельбиг, И. Буша, Е.И. Шендельс) и

теорию речевых фонов<sup>5</sup> (Г. Цифонун, Л. Хоффман), были определены наиболее частотные значения модальных глаголов в немецких текстах ПКТ «DER», а полученные данные были соотнесены с общей коммуникативной интенцией официально-делового функционального стиля. ПКТ «DER» оптимально подходит для решения данной задачи, поскольку значения модальных глаголов в немецком тексте точно передаются в параллельных текстах на русском языке.

Решения поставленной задачи осуществлялось в три этапа. Сначала был произведен поиск всех модальных глаголов в немецком субкорпусе ПКТ «DER» и определена их частотность (для всего субкорпуса и каждого текста отдельно). На втором этапе был проведен анализ значений модальных глаголов в соответствии с теорией речевых фонов. На завершающем этапе было проведено сравнение полученной информации об употреблении каждого модального глагола с общей коммуникативно-функциональной направленностью текстов ПКТ «DER». Исходя из того, что тексты ООН являются типичными для официально-делового функционального стиля, мы предположили, что прагматическая функция *долженствования* и *директивный характер* данных текстов должны соответствующим образом отражаться в речевом фоне употребления модальных глаголов.

Для того, чтобы провести анализ употреблений модальных глаголов, необходимо было выявить в корпусе абсолютно все их употребления в контексте. ПКТ «DER» позволил быстро и эффективно выполнить эти сложные поисковые запросы. Для поиска всех употреблений модальных глаголов в субкорпусе были составлены специальные поисковые запросы с использованием «регулярных выражений». Например, поисковый запрос для глагола *dürfen* выглядит так:

```
\Wdürf(en|t)\W|\Wdarf(st)*\W|\Wdurft(e|est|en|et)\W|\WgedurftW
```

Программная оболочка ПКТ «DER» позволяет не только найти все употребления для модальных глаголов, но и выявить их частотность во всем корпусе и в каждом отдельном его тексте. В немецких текстах ПКТ «DER», содержащих 153 931 словоупотребление, было найдено 838 употреблений модальных глаголов *dürfen, können, wollen, mögen, sollen* и *müssen* (см. табл. 1).

<sup>5</sup> Семантика модальных глаголов описывается не как сумма отдельных значений, а в зависимости от «речевого фона» (Redehintergrund) – условий, с учетом которых удастся оценить положение вещей как возможное или необходимое. Речевой фон привносится говорящим/пишущим, воспринимается и изменяется слушающим/читающим. Это своеобразный ориентир акта коммуникации, который определяет его интенцию.



Распределение модальных глаголов по частотности в немецких текстах  
ПКТ «DER»

Модальные глаголы	können	sollen	müssen	wollen	mögen	dürfen
Доля от всех употреблений	44%	29%	20%	3%	2%	2%

Как показывает частотность употребления *dürfen*, *mögen* и *wollen* встречаются в текстах значительно реже, чем *können*, *müssen* и *sollen*. Однако показатели частотности для всего корпуса не дают исчерпывающих сведений о частотности каждого из модальных глаголов в отдельно взятом тексте. Можно предположить, что частотность отдельно взятого глагола не зависит от типа текста и, следовательно, может значительно колебаться в зависимости от конкретного экземпляра текста.

На основании предположения о том, что значения частотно-распределительных характеристик модальных глаголов являются особенностью, свойственной определенному типу текстов (документам ООН, относящимся к официально-деловому функциональному стилю), была определена частотность модальных глаголов в каждом отдельно взятом тексте ПКТ «DER». Как видно из данных (см. табл. 2), глаголы *können*, *müssen* и *sollen* сохраняют высокую частотность в каждом из текстов корпуса. Это означает, что их стабильные частотно-распределительные характеристики являются отличительной чертой представленного в корпусе типа текстов, так как они практически не изменяются в отдельных текстах. Количество употреблений менее частотных глаголов (*dürfen*, *mögen* и *wollen*) нестабильно и колеблется в каждом отдельном тексте. На эти 3 глагола приходится всего 7% всех употреблений модальных глаголов в текстах корпуса. Столь низкая частотность данной группы глаголов и неустойчивость их частотно-распределительных характеристик в отдельных экземплярах текста свидетельствует о том, что они играют второстепенную роль в представленном ТТ, а их частотность не определяется особенностями представленных в ПКТ «DER» текстов.

Частотно-распределительные характеристики  
в текстах ПКТ «DER»

Модальный глагол	Тексты корпуса «DER» на немецком языке						Количество употреблений (все тексты)
	A-52-1	A-53-1	A-55-1	A-56-1	A-57-1	A-58-1	
können	43	67	64	73	62	72	381
sollen	29	46	50	35	30	50	240
müssen	25	26	37	31	19	28	166
wollen	0	8	6	6	0	3	23
mögen	4	0	3	1	2	6	16
dürfen	1	2	3	3	1	3	13
<i>Всего</i>	<i>102</i>	<i>149</i>	<i>163</i>	<i>149</i>	<i>114</i>	<i>162</i>	<i>839</i>

Для того, чтобы подтвердить, что приведенные частотно-распределительные характеристики модальных глаголов отражают особенность представленного в ПКТ «DER» типа текстов, был проведен анализ частотно-распределительных характеристик модальных глаголов в корпусе «Der deutsche Wortschatz» (далее «DDW», [www.wortschatz.uni-leipzig.de](http://www.wortschatz.uni-leipzig.de), проект университета г. Лейпцига). Данный корпус представляет собой базу данных, включающую тексты различных типов (в основном газетные и специализированные тексты). Корпус «DDW» содержит 35 млн. предложений (около 500 млн. словоупотреблений), что позволяет говорить о том, что он характеризует основной массив современной немецкой литературной письменной речи. В текстах корпуса «DDW» был произведен поиск всех употреблений модальных глаголов. Поскольку корпус «DDW» охватывает различные типы текстов широкой тематики и имеет значительный объем, данные о частотно-распределительных характеристиках модальных глаголов, полученные на его основе, могут считаться «масштабом» их употребления в современном немецком языке. В табл. 3 представлены результаты этого анализа.

Таблица 3

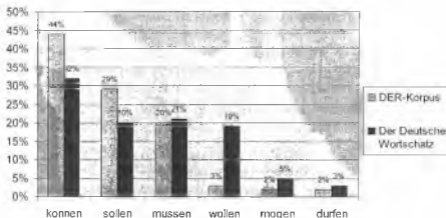
Распределение модальных глаголов по частотности в корпусе «DDW»

Модальные глаголы	können	sollen	müssen	wollen	mögen	dürfen
Доля от всех употреблений	32%	20%	21%	19%	5%	3%

Самым употребительным в корпусе «DDW» является *können*, что совпадает с данными, полученными на основе ПКТ «DER». Тем не менее, очевидно, что доля его употреблений в корпусе «DDW» заметно меньше (32%), чем в ПКТ «DER» (44%). Более показательными являются изменения частотно-распределительных характеристик *wollen*, *müssen* и *sollen*. Количество употреблений данных глаголов в корпусе «DDW» практически равно. Особенно значительным по сравнению с ПКТ «DER» (3%) является прирост употреблений *wollen* в корпусе «DDW» (19%). Количество употреблений *sollen* в «DDW» заметно ниже (20%), чем в немецких текстах ПКТ «DER» (29%) (см. диаграмму 2, отображающую различия частотно-распределительных характеристик модальных глаголов в немецких текстах ПКТ «DER» и в корпусе «DDW»)

Диаграмма 2

Распределение модальных глаголов по частотности  
в корпусе «DDW» и в ПКТ «DER»



При сравнении двух корпусов очевидно, что частотно-распределительные характеристики некоторых модальных глаголов в них значительно варьируются, что можно объяснить тем, что они содержат разные ТТ с различной коммуникативно-функциональной направленностью, а изменения в частотности употребления модальных глаголов являются отражением этих различий. Корпус «DDW» можно рассматривать как *общий*, поскольку он включает тексты различных типов. Перед таким корпусом ставится задача охватить не какую-то определенную проблемную область, а дать широкий срез наибольше-

го количества текстов самого различного типа. «DDW» имеет большой объем и может использоваться как базис для сравнения со специализированными корпусами. Таким *специализированным* одноязычным корпусом можно считать немецкую часть ПКТ «DER», т.к. он представляет собой подборку текстов определенного типа (тексты ООН, относящиеся к официально-деловому функциональному стилю).

Следующим шагом была интерпретация статистических данных, полученных на основе двух корпусов. Наиболее очевидным и объяснимым с учетом функциональных особенностей текстов корпуса «DER» различием является низкая частотность *wollen* («DER» 3%, «DDW» 19%). Это связано с тем, что основным значением данного глагола является выражение *желание* или *намерения*. Выражения субъективных желаний и использование интрасубъективно-волитивного речевого фона мало соответствует коммуникативно-функциональной направленности текстов официально-делового стиля. Это также служит объяснением более низкой частотности *mögen* в немецкой части ПКТ «DER». Половину употреблений глагола *wollen* составляют случаи его использования в условных придаточных предложениях с союзом *wenn*. При этом практически все эти употребления относятся к телеологическому речевому фону, где *wollen* используется не столько для выражения субъективного желания, сколько для определения целей и выражения стремления эти цели достигнуть. Значение желания, намерения, выражаемое обычно глаголом *wollen*, отступает в таких случаях на второй план. Ср.:

Gezieltes Zugehen auf junge Menschen ist ausschlaggebend, *wenn* die Vereinten Nationen auch in Zukunft ihre Relevanz *bewahren wollen*. [Bericht des Generalsekretärs über die Tätigkeit der Vereinten Nationen 1998]

Модальный глагол *mögen* используется в текстах субкорпуса «DER» в основном (~80% употреблений) в форме сослагательного наклонения (Konjunktiv Präteritum). Эта форма служит для выражения значения желания и является синонимичной глаголу *wollen*. При этом желание, выраженное таким образом, носит смягченный, менее категоричный характер и относится к событиям в настоящем времени. В текстах ООН, представленных в ПКТ «DER», такое употребление глагола *mögen* используется в интрасубъективно-волитивном речевом фоне, например, для выражения пожеланий докладчика при его обращении к аудитории по конкретному поводу. Ср.:

Ich *möchte* diese Gelegenheit *wahrnehmen*, um die Mitgliedstaaten abermals aufzurufen, zu dem Treuhandsfonds für den Sondergerichtshof beizutra-

gen. [Bericht des Generalsekretars über die Tätigkeit der Vereinten Nationen 2003]

Основной характеристикой *müssen* и *sollen* является выражение различных форм долженствования. В корпусе «DDW» эти глаголы проявляют примерно одинаковую частотность (*sollen* – 20% , *müssen* – 21%). В немецкой части ПКТ «DER» наблюдается значительное преобладание количества употреблений *sollen* (29%) над *müssen* (20%). По нашему мнению, это связано с тем, что выражение прагматической функции долженствования носит в документах ООН менее категоричный, рекомендательный характер. Статус ООН и язык её дипломатических документов обуславливают общую тенденцию снижения категоричности выраженного модальными глаголами долженствования, при одновременном сохранении коммуникативной функции *долженствования* и *предписывающего, нормативного характера* данных текстов (доклады Генсекретаря).

Основным значением *sollen*, используемым в текстах ПКТ «DER», является значение менее категоричного долженствования, выражающее чужую по отношению к субъекту действия волю: на данное значение приходится порядка 70% от всех его употреблений. Особенность использования глагола *sollen* в докладах Генсекретаря ООН заключается в том, что субъект долженствования в высказываниях не называется конкретно или представляется лишь как средство выполнения чужой воли. Так, в следующем примере указывается на назначение гуманитарных программ, разработанных ООН, а глагол *sollen* не только подчеркивает необходимость выполнения какого-то действия, но и определяет субъект предложения как средство, необходимое для достижения поставленных целей.

Verschiedene Programme der Vereinten Nationen für "gute Staatsführung" sollen den einzelnen Regierungen dabei *behilflich sein*, jenes Gleichgewicht zu finden, das ihren Bedürfnissen am ehesten entspricht. [Bericht des Generalsekretars über die Tätigkeit der Vereinten Nationen 1997]

В отличие от глагола *sollen*, спектр значений, выражаемых *müssen*, значительно уже. Необходимость, выражаемая данным модальным глаголом, обусловлена самим субъектом действия, т.е. осознается им как долг, внутреннее побуждение и является следствием неизбежности действия или состояния. Достаточно часто (~30%) в корпусе «DER» встречается употребление глагола *müssen* в форме первого лица множественного числа. Долженствование, выраженное таким образом, объединяет субъекта высказывания (Генсекретаря)

с аудиторией, представляя тем самым должествование как общую необходимость, а не как чужую волю:

Diesen Herausforderungen *müssen wir uns stellen*, wenn wir das unentbehrliche Instrument der Friedenssicherung durch die Vereinten Nationen den Krisen anzupassen suchen, denen wir uns heute und in Zukunft gegenübersehen, und wenn wir an einem System der kollektiven Sicherheit für das einundzwanzigste Jahrhundert arbeiten. [Bericht des Generalsekretärs über die Tätigkeit der Vereinten Nationen 1997]

Заметное различие в употреблении модальных глаголов в корпусе «DDW» и в ПКТ «DER» заключается в значительной разнице в частотности глагола *können*. Исходя из особенностей представленных в корпусе «DDW» и ПКТ «DER» текстов, мы предположили, что при их сравнении нормативный характер текстов ООН должен проявиться в некотором снижении частотности *können* и более высокой частотности *dürfen*. *Dürfen* в основном служит для выражения «категоричного разрешения», которое более соответствует прескриптивной интенции официально-делового функционального стиля. Тем не менее, первоначальное предположение не подтвердилось и при сравнении частотно-распределительных характеристик *können* и *dürfen* в корпусах «DDW» и «DER» было обнаружено противоположное явление. По сравнению с корпусом «DDW» (32%) количество употреблений *können* в немецких текстах ПКТ «DER» (44%) заметно выше, в то время как количество употреблений *dürfen* снижается («DER» 2%, «DDW» 3%). Как показал анализ употреблений модального глагола *können* в ПКТ «DER», большинство из них относится к телеологическому речевому фону и служит для ориентации высказывания на определенную цель, а также для выражения возможности её достижения, что вполне соотносится с общей интенцией представленных в ПКТ «DER» текстов. Так, например, модальный глагол *können* в придаточных предложениях с союзом *dass/daß* (15% всех употреблений *können*) служит для выражения возможности достижения целей, определенных в придаточном предложении. Ср.:

Es *müssen* angemessene Mittel *aufgebracht werden*, um zu gewährleisten, *daß* die erforderlichen Sicherheitsvorkehrungen für das Personal der Vereinten Nationen *getroffen werden können*. [Bericht des Generalsekretärs über die Tätigkeit der Vereinten Nationen 1997]

Анализ употреблений *dürfen* даёт основание утверждать, что этот модальный глагол практически не используется в текстах ПКТ «DER» в своём основном значении для выражения «категоричного разрешения». Большинство употреблений *dürfen* (11 из 13) представляет собой сочетание модального гла-

гола с отрицанием, что по своей сути является конструкцией, синонимичной основному значению глагола *sollen*. Этот факт объясняет то, что практически все употребления данного модального глагола относятся к нормативному речевому фону. Преобладание в немецких текстах ООН данного значения может быть объяснено характером текстов: все высказывания с использованием *dürfen* либо относятся к регламентированию деятельности самой организации, либо выражают рекомендации для мирового сообщества:

*Solche Anstrengungen dürfen jedoch nicht davon ablenken, dass die Hauptverantwortung für die Verhütung von Konflikten bei den Mitgliedstaaten liegt. [Bericht des Generalsekretärs über die Tätigkeit der Vereinten Nationen 2000]*

Несмотря на то, что прагматическая функция долженствования играет важную роль в документах ООН, её выражение носит *некатегоричный, рекомендательный* характер. Статус ООН и язык её дипломатических документов отражают общую тенденцию снижения категоричности выраженного модальными глаголами долженствования, при одновременном сохранении коммуникативной функции *долженствования* и *предписывающего, нормативного характера* данных текстов. Большинство употреблений глагола *können* относится к телеологическому речевому фону и служит для ориентации высказывания на определенную цель, а также для выражения возможности её достижения, что соотносится с общей интенцией представленных в ПКТ «DER» текстов ООН. Волитивный речевой фон играет в докладах Генсекретаря ООН скорее второстепенную роль, т.к. он не соответствует общей функционально-коммуникативной направленности официально-делового стиля. Этим фактором обусловлена низкая частотность глаголов *wollen* и *mögen*.

Результаты анализа употреблений модальных глаголов в немецком субкорпусе ПКТ «DER» показали, что некоторые из них (*sollen* и *können*) используются в ряде типичных синтаксических конструкций. Примерно половина всех употреблений *sollen* приходится на определительные придаточные предложения (~28% всех употреблений глагола *sollen*), условные придаточные предложения с союзом *wenn* (~14%), придаточные предложения с местоименным наречием *wodurch* или с предлогом *durch* + *существительное/местоимение* (~7%). Примерно треть всех его употреблений *können* приходится на придаточные предложения с союзами *dass/daß* (15%), *damit* (5%) и *wie* (10%).

Каждый из модальных глаголов реализует спектр модальных значений, а в каждом конкретном употреблении актуализируется одно из них. При этом

значение модального глагола обуславливаются речевым фоном высказывания. Анализ употреблений модальных глаголов в немецких текстах ПКТ «DER» также показал, что различные значения модальных глаголов отчетливо передаются в параллельных текстах на русском языке, что может служить в качестве дополнительного дифференцирующего признака, отражающего речевой фон данного употребления и позволяющего оценить его зависимость от общей интенции текстов, представленных в ПКТ «DER».

В заключении обобщаются основные результаты исследования, представляются выводы относительно апробации критериев составления немецко-русского параллельного корпуса текстов и возможностей применения ПКТ в лингвистических исследованиях, а также намечаются возможные перспективы дальнейшего развития и более широкого использования ПКТ «DER».

Основные положения диссертации отражены в следующих публикациях автора:

1. Вадяев С.Е. Электронная лексикография и корпусная лингвистика // Аспекты становления и функционирования западногерманских языков. – Самара: Изд-во «Самарский университет», 2003. – С. 83 – 92
2. Вадяев С.Е. Использование машинных текстов в лингвистических исследованиях контрастивного характера // Русско-немецкие культурно-исторические контакты: Материалы международной научной конференции (22-23.09 2004 г.) – Тольятти: ТГУ, 2004. – С. 86 – 89.
3. Вадяев С.Е. Использование корпусов параллельных текстов в прикладной германистике // Слово, высказывание, дискурс. Международный межвузовский сборник научных статей. – Самара: Изд-во «Самарский университет», 2004 – С 104 – 121.
4. Вадяев С.Е. Машинная обработка текстов и количественные методы в современной германистике. Методические указания по курсу «Техническое обеспечение филологического исследования» – Самара: Изд-во «Самарский университет», 2005. – 40 с.

Подписано в печать 02 03 2005.

Формат 60x84 1/16. Бумага офсетная. Печать оперативная.

Объём 1 печ. л. Тираж 100 экз. Заказ № 1184

Издательство «Самарский университет», 443011, г. Самара, ул. Акад. Павлова, 1.  
Отпечатано на УОП СамГУ