

РЕФЕРИРОВАНИЕ КАК СПОСОБ ИЗВЛЕЧЕНИЯ И ПРЕДСТАВЛЕНИЯ ОСНОВНОГО СОДЕРЖАНИЯ ТЕКСТА

Наталья Михайловна Нестерова

д. филол. н., профессор кафедры иностранных языков,
лингвистики и межкультурной коммуникации

Пермский национальный исследовательский политехнический университет
614990, Пермь, Комсомольский просп., 29a. nest-nat@yandex.ru

Наталья Александровна Герте

аспирант кафедры иностранных языков,
лингвистики и межкультурной коммуникации

Пермский национальный исследовательский политехнический университет
614990, Пермь, Комсомольский просп., 29a. gerte_nata@mail.ru

Статья посвящена рассмотрению процесса реферирования как одного из самых сложных видов речемыслительной деятельности. Особое внимание уделено исследованию основных методов реферирования, а также механизмов свертывания содержания исходного текста с целью его дальнейшего представления в тексте реферата. Авторами предлагается методика денотативного анализа текста, направленная на выявление его денотатной структуры. В работе также затрагивается вопрос об автоматизированном реферировании как одном из перспективных направлений компьютерной обработки текстов в условиях высокой информатизации общества. По мнению авторов, методика реферирования, основанная на денотативном анализе текста, может быть использована для формализации процесса реферирования.

Ключевые слова: реферирование; реферат; методы реферирования; семантическая адекватность; денотатная структура текста; автоматизированное реферирование.

Искусство реферирования как извлечения наиболее существенной информации из исходного текста и ее последующей презентации во вторичном (реферативном) тексте стало неотъемлемой частью современной жизни. Тем не менее реферирование имеет многовековую историю, уходящую корнями во времена Шумерской цивилизации, когда писцы создавали краткие записи на глиняных табличках. Далее, в эллинистический период и в эпоху распада Римской империи на пергаменте и свитках также делались краткие тезисы и описания документов. Тогда (предположительно) и возник термин *abstractus*. Первые реферативные журналы (РЖ) в Европе появились в XVII–XVIII вв., т.е. в эпоху, которая известна как «Эпоха открытий». Такие журналы позволяли своевременно информировать ученых о научных достижениях и наиболее значимых публикациях. Первыми журналами, публикующими рефераты, были «Philosophical Transactions», издающийся с 1665 г. Лондонским королевским обществом, итальянский журнал

«Giornali dei letterati» (1668–1680), немецкий журнал «Acta eruditorum» (1682). В XIX в. стали выходить в свет журналы, которые могут быть отнесены уже к собственно РЖ: «Pharmazeutisches Zentralblatt», переименованный впоследствии в «Chemisches Zentralblatt», «Fachzeitschrift für die gesarnte Papier, Pappen und Papierstoffindustrie», «Stahl und Eisen», «Science Abstracts».

В XX в. начинают издаваться реферативные журналы мирового значения: «Bulletin de l'Institut Pasteur» (Франция, 1903), «Agricultural Index» (США, 1916), «Berichte über die wissenschaftliche Biologie» (Германия, 1926), «Review of Applied Mycology» (Великобритания, 1922), «Review of Applied Entomology» (Великобритания, 1913) и др.; аналогичная картина наблюдается в области таких наук, как химия, физика: «Chemical Abstracts» (США, 1907), «Physikalische Berichte» (Германия, 1920), «Chimie et industrie» (Франция, 1918), «Geophysical Abstracts» (США, 1929) и др. [Соловьев, Гречихин, Здоров 1983].

В России первым реферативным центром стала Петербургская академия наук, а первыми русскими РЖ – «Краткое описание комментариев Академии наук» (1728), «Содержание ученых рассуждений» в четырех томах (1748–1754), академические сочинения, выбранные из 1-го тома «Деяний императорской Академии наук» под заглавием «Nova Acta» (1801). Первым теоретиком реферирования можно считать М. В. Ломоносова, который не только говорил о необходимости и пользе реферирования, но и сформулировал основные правила, которые должны были соблюдаться референтами-журналистами: «Всякий, кто берет на себя труд осведомлять публику о том, что содержится в новых сочинениях, должен прежде всего взвесить свои силы. Ведь он затевает трудную и очень сложную работу, при которой приходится докладывать не об обыкновенных вещах и не просто об общих местах, но схватывать то новое и существенное, что заключается в произведениях, создаваемых часто величайшими людьми. Высказывать при этом неточные и безвкусные суждения – значит сделать себя предметом презрения и насмешки; это значит уподобиться карлику, который хотел бы поднять горы» [Жанры информ. лит. 1983: 7]. Сам же термин «*реферат*» впервые появился в России в «Настольном словаре для справок по всем отраслям знаний» Ф. Г. Толля (1864), где он обозначался как «отношение, деловая записка, изложение дела вкратце» [Колесникова 2008: 32].

Важнейшим шагом в развитии реферирования в нашей стране стало создание в 1952 г. Всесоюзного института научной и технической информации АН СССР (сейчас это Всероссийский институт научной и технической информации РАН), который начал выпускать всем хорошо известные тематические РЖ ВИНТИ. Кроме этих реферативных журналов выпускаются и РЖ ИНИОН РАН, печатающие рефераты по общественным и гуманитарным наукам.

Однако нужно отметить, что, несмотря на долгую историю и востребованность, реферирование в основном развивалось эмпирически. И только во второй половине прошлого века ввиду стремительного роста значения вторичных документов в информационном обслуживании и в связи с появлением реферативных журналов была осознана необходимость создания теоретических основ и методик процесса реферирования с целью его совершенствования и улучшения качества рефератов. Появились стандарты и были выработаны требования к рефератам как официальным информационным документам. Сам же реферат был определен как «краткое точное из-

ложение содержания документа, включающее основные фактические сведения и выводы, без дополнительной интерпретации или критических замечаний автора реферата» [ГОСТ 7.9–95]. Это определение практически совпадает с дефиницией В. И. Соловьева, автора одной из известных методик реферирования, который определил реферат как «семантически адекватное, ограниченное малым объемом и вместе с тем возможно полное изложение основного содержания первичного документа» [Соловьев, Гречихин, Здоров 1983: 225]. Исходя из этих определений, главными требованиями к качеству реферата следует считать *объективность, полноту отражения содержания первоисточника и семантическую адекватность первичного и вторичного текстов.*

Рост популярности и востребованности реферативных текстов сделал необходимой и разработку оптимальной и научно обоснованной методики составления рефератов. Так появились первые теоретические и методические работы, было предложено несколько методов реферирования, которые можно разделить на две большие группы. Первую составляют методы, основанные на анкетировании. К ним относятся фасетное и поаспектное реферирование. Вторая группа представляет собой методы, в основе которых лежит анализ смысловой структуры первоисточника.

Из методов первой группы наиболее известно *поаспектное* реферирование, предложенное В. И. Соловьевым [Соловьев 1971]. Данный метод предполагает семантический анализ текста с помощью его разбиения на отдельные содержательные аспекты (сетки). К таким аспектам можно отнести проблему и тему исследования, область знания, цель работы и полученные результаты, методы и условия проведенного исследования. По мнению автора данной методики, заданный перечень аспектов, на которые должен ориентироваться референт, позволяет обеспечить как семантическую адекватность первичного и вторичного текстов, так и эффективное восприятие информации [Соловьев, Гречихин, Здоров 1983: 246]. Соответственно, такой подход позволяет, с одной стороны, сделать процесс реферирования более формализованным, а с другой – стандартизировать сами рефераты.

Другой метод реферирования был предложен А. И. Жолковой, которая обозначила его как фасетное реферирование [Жолкова 1970]. Он представляет собой анализ содержания конкретной сферы научной деятельности, заключающийся в накладывании на текст сетки фасетов, соответствующей числу категорий, отражающих специфи-

ку данной области в различных ее аспектах. Основным понятием данной методики необходимо считать элементарные логические сообщения, которые создают табличный (структурированный) реферат [Жолкова 1985]. Автор фасетного реферирования справедливо считает, что табличная форма реферата – это один из способов индексирования документов и структуризации фактографической информации. (Табличный реферат не был предназначен для замены реферата традиционного типа.) В целом, можно сказать, что фасетное реферирование в некоторой степени близко поаспектному, так как обе методики основываются на «накладывании» сетки определенных аспектов или фасетов на оригинальный текст.

Представляется, что вышеназванные методы реферирования можно рассматривать как своего рода анкетирование: референту предлагается «анкета» с «вопросами», которые помогут ему найти в тексте первоисточника соответствующую информацию. Главным достоинством анкетных видов реферирования является ограниченность рамками априорно заданной схемы, которая отражает структуру предметной области, потребности специалистов и другие прагматические задачи. С другой стороны, у референта могут возникнуть некоторые трудности. Как утверждает Д. И. Блюменау, основным недостатком этих методов является обращение к интуиции референта, который должен выделить «существенную информацию», а также стремление к детальности самой методики, что может значительно усложнить процесс реферирования [Блюменау 1982: 81]. Мы же полагаем, что методы, основанные на анкетировании, в большей степени ориентированы на создание профильных, специализированных рефератов, поскольку в них отражается именно та информация, которая удовлетворяет потребности специалистов конкретной области знания.

Другую большую группу методов реферирования составляют методы выделения содержания первоисточника на основе анализа реферлируемого текста. К данной группе можно отнести, например, логико-психологический анализ, предложенный А. А. Вейзе [Вейзе 1978]. К этой же группе относится и так называемое информационно-классификационное реферирование, в основу которого лег критерий точности информирования. Толчком для его развития послужил метод поаспектного реферирования, а именно разделение информации по степени значимости и полезности для потребителей. Предложенный метод реферирования опирается на содержательный и логический анализ первоисточника, кото-

рый позволяет оценивать важность и приоритетность одних элементов информации относительно других, что делает методику достаточно универсальной. В связи с этим данный подход стал широко применяться в информационных центрах и достаточно популярен в настоящее время.

Однако независимо от того, какой метод реферирования выбран референтом, основной задачей является установление семантической адекватности оригинала и реферата. Согласно приведенным выше определениям реферата содержательно адекватным является такой реферат, в котором в сокращенной форме, но точно, без искажений и интерпретации, воспроизводится *основное содержание* первичного документа. Это означает, что вопрос семантической адекватности непосредственно связан с проблемой «основного содержания» текста. Эта проблема может быть решена на основе анализа внутренних закономерностей процесса понимания, который является ведущим в любом виде смыслового преобразования, в том числе и при реферировании. Необходимо подчеркнуть, что понимание при реферировании должно быть полным, иначе процесс свертывания, как подчеркивал А. И. Новиков, может иметь случайный (стохастический) характер. Полное же понимание предполагает оценку текста в целом, что требует установления связей между всеми смысловыми блоками текста и соотнесения частей с целым, чтобы не было «разрывов» и «пустот» [Нестерова 2011]. Результаты экспериментальных данных показывают, что смысловое восприятие текста и его понимание есть переход от воспринимаемых языковых средств к единицам внутреннего языка. Этот переход сопровождается выделением из текста ключевых элементов, которые А. Н. Соколов называет «смысловыми вехами». Такими вехами, или «опорными пунктами», могут быть любые слова, содержащиеся в тексте, однако именно они представляют во внутренней речи большие смысловые отрезки, которые объединяются в «целостное мыслительное образование», несущее смысл первичного документа [Новиков, Нестерова 1991: 41]. Таким образом, понимание – это сложный аналитико-синтетический процесс, включающий в себя анализ (расчленение) и синтез (объединение), что и делает его одновременно и процессом смыслового свертывания.

Как уже было сказано, в психолингвистической теории текста, основы которой были заложены работами Н. И. Жинкина и А. И. Новикова, критерием и условием понимания текста является переход от его внешней формы к внутренней, т. е. к денотатной структуре текста, представ-

ляющей совокупность иерархически связанных между собой денотатов. Такой подход обеспечивает возможность эксплицитно выразить содержание в виде денотатного графа, методика построения которого была разработана А. И. Новиковым. Построение денотатного графа, т.е. «перекодирование линейного текста в целостную, иконическую схему» [Жинкин 1978: 51], ведет к полному преобразованию языковой формы текста, в процессе которого происходит определенный отбор денотатов, устанавливается их иерархия и устраняются те, которые не имеют отношения к моделируемой ситуации. Таким образом, денотатный граф – это свернутое эксплицитное отображение структуры содержания текста, которой могут соответствовать различные языковые формы [Новиков, Нестерова 1991: 61]. В этом случае реферирование можно представить как переход $T_1 - ДС - T_2$: сначала происходит переход от внешней языковой структуры текста к структуре его содержания (денотатной структуре), а затем переход от денотатной структуры к новой внешней форме вторичного текста. Стоит также отметить, что выделенные этапы взаимозависимы и взаимообусловлены: уровень понимания (как результат первого этапа) определяет вторичный текст с точки зрения его адекватности оригиналу. Установка на определенный вид смыслового преобразования влияет на процесс осмысления, и каждый вид такого преобразования имеет свои внутренние закономерности и специфику. Реферирование, в частности, характеризуется особыми отношениями между T_1 и T_2 , которые можно обозначить как отношения «коммуникативной неравноценности» [Новиков, Нестерова 1991: 55]. Другими словами, в процессе реферирования происходит содержательное уподобление текстов при полном отсутствии формального. Важно отметить, что возможность преобразования текста и его сжатия заложена уже в самом тексте благодаря такому его онтологическому свойству, как компрессионность [Новиков 2007]. Именно это свойство текста позволяет его сворачивать с различной степенью компрессии в соответствии с целью реферирования, которая может быть разной в зависимости от вида реферата.

В настоящее время наряду с общими информативными рефератами очень востребованными являются и так называемые специализированные рефераты, ориентированные на специалистов, связанных с определенной предметной областью. В этом случае, на наш взгляд, процесс реферирования может быть описан в терминах теории *скопоса*, поскольку здесь именно *цель* (точнее, «целевая аудитория») определяет и отбор, и ком-

бинирование элементов содержания. Меняется, соответственно, и критерий адекватности: текст реферата должен быть адекватен не тексту-источнику, а той коммуникативной ситуации, в которой ему предстоит функционировать.

С целью исследования механизмов свертывания информации нами было выполнено экспериментальное реферирование трех англоязычных публикаций, представляющих собой тексты докладов, прочитанных на российской нефтегазовой технической конференции и выставке SPE. Задача заключалась в выполнении реферативного перевода одного и того же текста для разных групп специалистов: РНГМ (разработка нефтяных и газовых месторождений), ГНГ (геология нефти и газа), МОН (машины и оборудование нефтяных и газовых промыслов) и НГП (нефте- и газопроводы). Исходя из концепции понимания текста как перехода от его внешней (языковой) формы к внутренней, мы представили содержание реферируемых текстов в виде денотатных графов, что и позволило выделить как общую информацию (на ее основе были написаны информативные рефераты, отражающие содержание исходного текста в целом), так и специальную информацию, представляющую интерес для вышеназванных групп специалистов. Эта информация и была отражена в созданных нами соответствующих специализированных рефератах. Все написанные рефераты были предложены специалистам, которые оценили их с точки зрения информативности и соответствия скопосу.

Выполненное нами экспериментальное реферирование еще раз доказывает, что исследования школы Н. И. Жинкина и А. И. Новикова в области семантики текста и его понимания могут быть методологической основой как для исследования механизмов понимания и смыслового свертывания текста, так и для разработки методики реферирования, включая и *речевое сжатие*, необходимое для порождения соответствующего текста реферата, имеющего фиксированную языковую форму. Суть речевого сжатия состоит в членении денотатной структуры на крупные фрагменты и в их лексическом и синтаксическом «означивании».

Таким образом, установка на реферирование предполагает обязательное осмысление всего исходного текста в целом, поскольку моделирование предметной ситуации позволит сохранить целостность денотатной структуры при существенном сжатии речевой формы. Это означает, что исходный текст (T_1) и текст реферата (T_2) представляют собой семантически адекватные речевые произведения, характеризующиеся разными логико-композиционными структурами и

созданные на основе языковых единиц с различными объемами содержания. Их адекватность определяется на содержательном уровне, что еще раз подтверждает, что для достижения подобной семантической равноценности необходим высокий уровень понимания исходного текста.

Анализ имеющихся исследований в области аналитико-синтетической обработки исходного текста дает основание считать, что развитие методов реферирования осуществлялось в основном благодаря совершенствованию «ручного» реферирования. Однако в современных условиях мощных информационных потоков невозможно обойтись без автоматизированных технологий.

В настоящее время формализация реферирования стала одной из важнейших задач компьютерной обработки текстов. Современные технологии позволяют уже намного больше, чем в то время, когда создавались методики реферирования, о которых речь шла выше. Исследователи, занимающиеся вопросом автоматизированного реферирования, предлагают несколько способов обработки информации с помощью компьютера, каждый из которых характеризуется своим методом выделения информации: это так называемое квазиреферирование и краткое изложение содержания первичного текста [WebGround].

Квазиреферирование заключается в выделении наиболее информативных фраз и формировании квазирефератов на их основе. Сегодня различают следующие методы отбора информации:

1. Статистические методы, основанные на оценке уровня информативности элементов первичного текста по частоте их появления. Именно частота появления служит ключевым критерием информативности слов или целых фрагментов.

2. Позиционные методы, основной идеей которых можно считать предположение о том, что информативность слова (или предложения) в тексте зависит от его позиции.

3. Индикаторные методы, позволяющие выявить наличие специальных слов в документе, так называемых маркеров важности, которые характеризуют содержательную значимость документа.

Другой способ автоматизированного реферирования – это краткое изложение содержания, базисом которого является выделение наиболее важной информации с помощью автоматизированного качественного контент-анализа. Такой анализ проходит в три стадии. На первой стадии происходит сведение исходного текста к заданному числу фрагментов, т.е. единиц значения. На второй стадии подключается процесс поиска свя-

зей между выбранными единицами значения, после чего начинается третья, последняя, стадия – формирование выводов и обобщений, которые ложатся в основу текста реферата в виде связанных смысловых единиц.

Для выполнения всех этих преобразований используются манипуляции логическими предположениями и выделяются шаблоны в имеющейся базе данных. В результате рождается концептуальная структура текста, т.е. концептуальные «выжимки» из текста [WebGround].

Безусловно, автоматизированное реферирование дает возможность ускорить процесс и исключить субъективность при изложении содержания оригинала. Однако необходимо признать, что методы автоматизированного реферирования, которые применяются сегодня, не удовлетворяют всем требованиям реферирования, и качество вторичных текстов пока остается достаточно низким. Это связано с тем, что при выделении текстовых блоков не учитываются связи и отношения между ними, в результате порождаются тексты, лишенные внешней связности и внутренней целостности.

Неудачи в области формализации реферирования, несомненно, связаны с тем, что данная проблема относится к классу задач, эффективное решение которых, как писал А. И. Новиков, невозможно без формализации содержательной стороны текста. Однако мы считаем, что при определенных ограничениях, прежде всего связанных с предметно-тематической областью текста и его определенной стандартизованностью, методика денотативного анализа позволит формализовать процесс реферирования, что и было мечтой Анатолия Ивановича Новикова, ученого, чья книга «Семантика текста и ее формализация», увидевшая свет в 1983 г., значительно опередила свое время.

Список литературы

Блюменау Д. И. Проблемы свертывания научной информации. Л.: Наука, 1982. 166 с.

Вейзе А. А. Реферирование текста. Минск: Изд-во Белорус. гос. ун-та, 1978. 128 с.

ГОСТ 7.9-95. Реферат и аннотация. Общие требования. URL: <http://www.tehnorma.ru/> (дата обращения: 05.09.2013).

Жанры информационной литературы: Обзор. Реферат / А. А. Гречихин, И. Г. Здоров, В. И. Соловьев. М.: Книга, 1983. 320 с.

Жинкин Н. И. Сенсорная абстракция // Проблемы общей, возрастной и педагогической психологии. М., 1978. С. 38–59.

Жолкова А. И. Фасетный метод реферирования как проблема библиографирования техниче-

ской литературы: автореф. дис. ... канд. пед. наук. Л.: ЛГИК им. Н.К. Крупской, 1985. 15 с.

Жолкова А. И. К вопросу о реферировании научно-технической периодики в реферативном журнале ВИНТИ // НТИ. Сер. 1. 1970. № 5. С. 24–25.

Колесникова Н. И. От конспекта к диссертации: учеб. пособие по развитию навыков письменной речи. М.: Флинта: Наука, 2008. 288 с.

Нестерова Н. М. Реферативный перевод: проблема смыслового свертывания и семантической адекватности // Вестник ЧелГУ. Сер. Филология. Искусствоведение. 2011. Вып. 58, №25(240). С. 112–118.

Новиков А. И., Нестерова Н. М. Реферативный перевод научно-технических текстов. М.: Ин-т языкозн. АН СССР, 1991. 148 с.

Новиков А.И. Текст и его смысловые доминанты. М.: «Азбуковник», 2007. 224 с.

Соловьев В.И., Гречихин А.А., Здоров И.Г. Составление и редактирование рефератов: Вопросы теории и практики // Жанры информационной литературы. Обзор. Реферат. М., 1983. С.199–320.

Соловьев В.И. Поаспектный метод реферирования // НТИ. Сер. 2. 1971. №2. С.14–17.

WebGround. URL: http://www.webground.ru/services.php?param=book&part=chapter%203_2_3.htm (дата обращения: 03.09.2013).

ABSTRACTING AS A WAY TO EXTRACT AND PRESENT THE MAIN INFORMATION OF THE TEXT

Natalia M. Nesterova

**Professor of Department of Foreign Languages, Linguistics and Intercultural Communication
Perm National Research Polytechnic University**

Natalia A. Gerte

**Post-graduate Student of Department of Foreign Languages, Linguistics and Intercultural Communication
Perm National Research Polytechnic University**

The paper describes the process of abstracting as one of the most difficult types of verbal and cognitive activity. Particular attention is paid to the research of the main abstracting methods and the mechanisms of content compression and its representation in a new text (an abstract). The paper presents the denotative analysis approach aimed to extract from the source text its denotative structure which presents its main content. The paper also touches upon the problem of automated abstracting as one of the emerging computer text processing approaches in the age of highly-informed society. The authors suggest the presented method of abstracting, based on the denotative analysis, can be used for formalization of abstracting.

Key words: abstracting; abstract; methods of abstracting; semantic adequacy; text denotative structure; automated abstracting.