

ОТЗЫВ

об автореферате диссертации **И. А. Меньшениной**
“Автоматическая смысловая компрессия публикаций научного
медицинского дискурса (на материале английского языка)”,
представленной на соискание ученой степени кандидата
филологических наук по специальности
10.02.21 – прикладная и математическая лингвистика

Активный рост научных исследований, который подтверждается экспоненциальным нарастанием объёмов профессиональной информации, отражаемой в научных журналах, требует автоматизации обработки текстов, форсирующей их освоение и использование. При переходе к доказательной парадигме в самых различных областях знаний, медицине в том числе, обеспечение возможности корректной компрессии информации, отражаемой в публикациях, позволяет интенсифицировать и сам обмен информацией и ее извлечение для поддержания исследований в актуальном состоянии. В рамках многолетних исследований процесса компрессии информации, осуществляемого при аннотировании и реферировании научных текстов, накоплен важный опыт, на основе которого была осуществлена разработка множества моделей и систем автоматизации компрессии информации. Однако следует согласиться с автором АР, они не обеспечивают достаточную точность, столь востребованную в медицинской сфере. Соответственно, разработка новых подходов к решению задачи смысловой компрессии научных публикаций в области медицины, при которой должен создаваться предельно точный вторичный текст, в котором сохраняется полнота и точность профессионально значимой информации, безусловно **актуально**.

Теоретически при решении задачи автоматизации создания аннотации возможны два подхода: формальный, основанный на компрессии научного текста за счет удаления из него синсемантических предложений и целых абзацев, опоры на реляторы и коннекторы, которые также позволяют выделить несамостоятельные предложения, и количественного анализа оставшихся предложений за счет учета используемых терминологических единиц. Второй подход является по сути фреймовым, основанным на порождении нового вторичного текста по заданным заранее слотам. Сложность такого подхода связана с необходимостью детального семантического анализа текста статьи и генерации нового текста.

В работе И. А. Меньшениной выбран второй подход, при этом достаточно взвешено и обосновано представлены теоретические основания для применения субъектно-ориентированного подхода, позволяющего проводить когнитивный анализ внутренней структуры научной статьи с учетом заинтересованной включенности разных участников медицинского дискурса, что позволило автору обнаружить и представить многослойную структуру организации смысла научной медицинской публикации. Это и определяет возможность точного выделения семантически наиболее важной информации для осуществления формализации процесса компрессии текста. Соответственно, целью диссертационного исследования является разработка лингвистического обеспечения и системы автоматической компрессии научных медицинских публикаций на английском языке, направленной на существенное уменьшение их объема с сохранением основного авторского замысла.

Теоретическая основа исследования позволила сформулировать и экспериментально проверить такую важную особенность медицинского дискурса как обязательность суперструктуры, определяющей глубинный логический каркас содержания публикации, выявить и классифицировать 4 категории дискурсивных маркеров, формирующих микроструктуру текста. На основе проведенного лингвистического исследования представительного массива текстов разработана и апробирована прикладная компьютерная программа MedCompress (Medical Compression) с шестью базовыми модулями, каждый из которых отвечает за выполнение конкретной задачи. Связь между этими модулями в соответствии с кибернетическим подходом осуществляется на основе передачи параметров созданного автотром лингвистического ресурса – базы знаний программы MedCompress.

Научная новизна исследования бесспорна и определяется как выявлением структурно-смысловой организации научной медицинской публикации, так и разработкой программы MedCompress, позволяющей выполнять компрессию текста публикации с сохранением основного содержания. Достоверность результатов исследования обеспечивается репрезентативностью материала, подтвержденной статистическими методами, а также корректностью использованных методов анализа. Теоретическая значимость и практическая ценность исследования несомненны. Чрезвычайно важно, что предложенные методы позволяют обеспечить оперативное ознакомление специалистов в области медицины с основными результатами, направлениями и тенденциями в этой активно развивающейся науке.

Таким образом, можно утверждать, что, судя по автореферату, диссертация **И. А. Меньшениной** представляет собой завершенное исследование, имеющее как практическую ценность, так и новизну, соответствующие заявленной специальности.

Работа **И. А. Меньшениной** безусловно соответствует требованиям, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения ученой степени по специальности 10.02.21 – прикладная и математическая лингвистика.

Выражаю свое согласие на размещение моего отзыва об автореферате диссертации И. А. Меньшениной “Автоматическая смысловая компрессия публикаций научного медицинского дискурса (на материале английского языка)” на сайте учреждения образования «Минский государственный лингвистический университет».

Профессор кафедры образовательных технологий в филологии
филологического факультета РГПУ им. А. И. Герцена,
доктор филологических наук, профессор
заслуженный деятель науки РФ

Л.Н. Беляева

20 марта 2025 г.

РГПУ им. А.И. ГЕРЦЕНА.

подпись Л.Н. Беляева

удостоверяю «14» 03 2025 г.

Отдел кадров управления по работе с кадрами
и организационно-контрольному обеспечению

Ведущий документовед
отдела кадров

Ю.В. Пасечник