

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

На правах рукописи

Худякова Мария Викторовна

**МЕТОДЫ АНАЛИЗА И ОЦЕНКИ УСТНОГО ДИСКУРСА
ЛЮДЕЙ С РЕЧЕВЫМИ НАРУШЕНИЯМИ**

Резюме

диссертации на соискание ученой степени
кандидата филологических наук

Научный руководитель:
Доктор филологических наук
Драгой Ольга Викторовна

Москва 2023

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский университет «Высшая школа экономики».

Публикации

На защиту выносятся четыре публикации, в которых описывается разработка методов анализа и оценки устного дискурса людей с речевыми нарушениями. Автор диссертации является первым автором статьи (Khudyakova et al., 2023), которая опубликована в журнале, индексируемом в базе данных Scopus, и единственным автором в статье (Худякова, 2020), опубликованной в журнале, входящем в перечень журналов высокого уровня, подготовленный НИУ ВШЭ. В выносимой на защиту статье (Linnik et al., 2021) опубликованной в журнале, входящем в 1 квартиль в базах данных WoS и Scopus, автор диссертации является четвертым автором, и в статье (Ivanova et al., 2021), опубликованной в журнале, входящем в 1 квартиль в базе данных Scopus, – восьмым автором.

1. Ivanova, M. V., Akinina, Y. S., Soloukhina, O. A., Iskra, E. V., Buivolova, O. V., Chrabaszcz, A. V., Stupina, E. A., Khudyakova, M. V., Akhutina, T. V., & Dragoy, O. (2021). The Russian Aphasia Test: The first comprehensive, quantitative, standardized, and computerized aphasia language battery in Russian. PLOS ONE, 16(11), e0258946. <https://doi.org/10.1371/JOURNAL.PONE.0258946> (Русский афазнологический тест: первая комплексная, квантитативная, стандартизированная и компьютеризированная батарея языковых тестов на русском языке)
2. Linnik, A., Bastiaanse, R., Stede, M., & Khudyakova, M. (2021). Linguistic mechanisms of coherence in aphasic and non-aphasic discourse. *Aphasiology*, 36(2), 123-146. <https://doi.org/10.1080/02687038.2020.1852527> (Лингвистические механизмы связности в афатическом дискурсе)
3. Khudyakova, M., Antonova, N., Nelubina, M., Surova, A., Vorobyova, A., Minnigulova, A., Gronskaya, N., Yashin, K., Medyanik, I., Shishkovskaya, T., Ryazanskaya, G., Zuev, A., & Dragoy, O. (2023). Discourse Diversity Database (3D) for Clinical Linguistics Research: Design, Development, and Analysis. *Bakhtiniana: Revista de Estudos Do Discurso*. <https://doi.org/10.1590/2176->

[4573e55885](#) (База дискурсивного разнообразия (3D) для клинических лингвистических исследований: дизайн, развитие и анализ)

4. Худякова, М. В. (2020). Влияние степени усталости говорящего на характеристики устного дискурса. *Российский Журнал Когнитивной Науки*, 7(3), 78–88.

По результатам диссертационного исследования было сделано 5 устных и постерных докладов на международных конференциях:

1. VII Международная научная конференция «Общетеоретические и типологические проблемы языкознания» (Россия, онлайн, 2022). Устный доклад: Корпусный анализ устной речи пациентов до и после резекции опухолей мозга.
2. The International Online Workshop on Language in Healthy and Pathological Aging (Испания, онлайн, 2021). Устный доклад: Effect of age on lexical diversity in three discourse genres.
3. Linguistic Forum 2020: Language and Artificial Intelligence (Россия, онлайн, 2020). Устный доклад: Speech characteristics depend on discourse elicitation task.
4. The Fifth Saint Petersburg Winter Workshop on Experimental Studies of Speech and Language (Россия, 2019). Постерный доклад: Effect of speaker's fatigue and discourse genre on speech characteristics: a pilot study.
5. XI International interdisciplinary congress: Neuroscience for medicine and Psychology (Россия, 2015). Устный доклад: Developing a standardized test for language assessment in aphasia.

Введение

В диссертацию включены статьи, посвященные разработке методов оценки порождения и понимания устного дискурса, а также созданию корпуса устной речи пациентов с различными неврологическими и психиатрическими нарушениями. Одна из статей посвящена разработке и стандартизации субтестов «Понимание дискурса» и «Порождение дискурса» Русского афазиологического теста (РАТ). Во второй статье представлена новая шкала для оценки четырех аспектов связности дискурса, а также представлены результаты исследования вклада микро-и макролингвистических параметров в аспекты связности. Третья статья посвящена описанию корпуса 3D (База

дискурсивного разнообразия) и его разделов. В четвертой главе описано исследование на базе одного из разделов 3D, посвященное влиянию усталости говорящего на различные параметры речи.

В клинической лингвистике анализ устной речи – важная составляющая оценки языковых навыков у людей с различными неврологическими расстройствами. Такой анализ позволяет выявить нарушения как на микролингвистических (фонетическом, лексическом, семантическом), так и макролингвистических (дискурсивный, прагматический) уровнях (Bryant et al., 2017; Prins & Bastiaanse, 2004). Поскольку понимание и порождение связного дискурса – основа человеческой коммуникации (Mar, 2004; Schank, 1995), при работе с речевыми нарушениями крайне важно оценивать дискурсивный уровень. Именно поэтому субтесты, оценивающие понимание и порождение дискурса включены во многие стандартизированные батареи для оценки речи, например, CAT (Swinburn et al., 2004); BDAE (Goodglass et al., 2001); QAB (Wilson et al., 2018).

Разработка субтестов для оценки порождения и понимания дискурса сопряжена с определенными сложностями. Во-первых, понимание связного текста требует не только обработки фонологической, лексико-семантической и синтаксической информации, но и построения логических выводов и понимания связей между элементами дискурса, а также создает нагрузку на рабочую память (см. Carpenter et al., 1995; Linda E Nicholas & Brookshire, 1995). Поэтому в субтесте для оценки понимания дискурса размер текста и количество вопросов должны быть ограничены, чтобы пациент был способен выполнить задание, но при этом достаточны для того, чтобы различить разные уровни тяжести нарушений. Чтобы оценить порождение речи, необходимо разработать такой метод оценки (схему транскрибирования или оценочную шкалу), который бы был достаточно экономным с точки зрения затраченного времени и усилий, но при этом соблюдался бы достаточный уровень тест-ретестовой и межэкспертной надежности. В первой главе описаны субтесты «Понимание дискурса» и «Порождение дискурса», разработанные автором диссертации для РАТ, а также результаты стандартизации теста.

Оценка с помощью шкал – стандартный метод для анализа устного дискурса, и особенно макролингвистических параметров дискурса, применяемый как в клинической практике, так и в фундаментальных исследованиях. Однако до сих пор не до конца изучена связь между тем, как слушающий интерпретирует и оценивает дискурс как некоторое «единое целое», и количественными параметрами дискурса (например, количеством ошибок, лексическим разнообразием и т. п.). Во второй главе описано исследование, в

фокусе которого – связь различных аспектов дискурсивной связности, измеренных с помощью разработанной авторами оценочной шкалы, и количественными характеристиками дискурса, извлеченными из аннотированных аудиозаписей людей с афазией и контрольной группы.

Для многосторонней оценки устной речи при различных нарушениях создают аннотированные мультимедийные корпуса. Корпусный анализ позволяет исследовать различные источники вариативности характеристик речи в клинических популяциях и норме. В клинической корпусной лингвистике установлены определенные стандарты сбора данных и аннотации, а также созданы многие корпуса речи людей с различными неврологическими и психиатрическими расстройствами, например коллекция TalkBank (MacWhinney, 2007), кембриджский корпус рассказов «Воровство печени» (Williams et al., 2010), Греческий корпус афатического дискурса (Varlokosta, 2016) и другие. Однако на данный момент нет большого разнообразия подобных корпусов для русского языка. Третья глава посвящена описанию Базы дискурсивного разнообразия (3D), протокола сбора данных и описанию разделов корпуса. Также приведены результаты исследования влияния усталости на параметры речи, основанные на одном из разделов 3D.

Цель диссертации – представить набор методов для оценки и анализа порождения и понимания дискурса для применения в клинической практике и научных исследованиях. Статьи, включенные в диссертацию, используют различные методы оценки дискурса, позволяющие оценить как отдельные аспекты речи, так и дискурсивные способности в целом. **Актуальность** работы определяется тем, что в настоящее время не существует достаточного количества современных и разработанных для русского языка шкал оценки дискурса, а также больших корпусов речи людей с неврологическими нарушениями.

Объект исследования – порождение и понимание дискурса при различных неврологических и психиатрических нарушениях, а также в различных функциональных состояниях. **Предмет исследования** – методы оценки порождения и понимания дискурса с помощью разработанных шкал и анализ микро- и макролингвистических параметров речи на основе корпусов речи в различных клинических популяциях.

Новизна работы заключается в следующем:

- РАТ – первый комплексный стандартизированный тест на русском языке с субтестами для оценки понимания и порождения дискурса, разработанный с учетом современных психолингвистических теорий.

- Новая шкала для оценки связности дискурса позволила провести исследование связи различных аспектов связности и микро- и макролингвистических параметров речи.
- Корпус 3D – первый большой русскоязычный корпус, в который входят образцы устного дискурса людей с неврологическими и психиатрическими расстройствами, а также здоровых людей различных возрастных групп и в различных функциональных состояниях.

Теоретическая значимость диссертации заключается в следующем:

- Анализ связности образцов дискурса показал, что при афазии связность нарушена, а также что нарушения различных аспектов связности связаны с разными микро- и макролингвистическими параметрами дискурса.
- Уровень усталости говорящего влияет на темповые характеристики речи.

Практическая значимость исследования заключается в следующем:

- Были разработаны и стандартизированы субтесты «Понимание дискурса» и «Порождение дискурса» Русского афазиологического теста.
- Были разработаны шкалы для оценки четырех аспектов связности дискурса.
- Был создан корпус 3D, а также описан протокол сбора данных и создана схема разметки.

Основные положения, выносимые на защиту:

- 1) Субтесты «Понимание дискурса» и «Порождение дискурса» Русского афазиологического теста были созданы с учетом современной психолингвистической теории и стандартизированы в группе людей с афазией и контрольной группе. РАТ отвечает современным психометрическим стандартам и позволяет выявить нарушения речи при афазии.
- 2) Новая шкала для оценки четырех аспектов связности дискурса позволила выявить нарушения связности дискурса при афазии. Разные аспекты связности наиболее связаны с разными комбинациями микро- и макролингвистических параметров.
- 3) Корпус 3D включает в себя образцы устной речи людей с различными неврологическими и психиатрическими нарушениями. Для элиситации трех типов дискурса были подобраны три типа заданий. Схема разметки позволяет анализировать как микро-, так и макролингвистические параметры дискурса.

- 4) Результаты пилотного исследования выявили вариативность фонетических и темповых параметров дискурса в зависимости от дискурсивного типа и состояния говорящего.

1. Создание и стандартизация субтестов «Понимание дискурса» и «Порождение дискурса» Русского афазиологического теста

Статья, выносимая на защиту: (Ivanova et al., 2021)

Русский афазиологический тест (РАТ) – стандартизированный тест для оценки речи, позволяющий оценить понимание и порождение устной речи. В отличие от скрининговых инструментов, таких как Быстрый тест на афазия (Vuivolova et al., 2020), Токен тест (Bastiaanse et al., 2015), Афазия: быстрая верификация (Буйволова и др., 2020), РАТ представляет собой комплексный инструмент для оценки речевого тяжести речевого дефицита на различных языковых уровнях. РАТ состоит из 13 субтестов, оценивающих понимание речи (Восприятие псевдослов, Лексическое решение, Понимание существительных, Понимание глаголов, Понимание предложений, Понимание дискурса), повторение (Повторение псевдослов, Повторение слов, Повторение предложений) и порождение речи (Называние объектов, Называние действий, Порождение предложений, Порождение дискурса). Материалы каждого субтеста созданы с учетом современной психолингвистической теории.

В настоящем исследовании мы представляем опыт создания и апробации субтестов для оценки понимания и порождения дискурса. Понимание и порождение дискурса – одни из наиболее экологически валидных заданий для оценки речи, так как истории, их рассказывание и понимание, являются одной из основополагающих частей человеческой коммуникации (Mar, 2004). Понимание связного текста требует фонологической, лексико-семантической и синтаксической обработки отдельных элементов дискурса, а также понимания связи между этими единицами (Carpenter et al., 1995; Nicholas & Brookshire, 1995) и создает нагрузку на рабочую память. По образцу порождённой связной речи можно провести оценку речевых дефицитов на фонологическом, лексико-семантическом, синтаксическом и дискурсивном уровнях (Bryant et al., 2016).

Субтест «Понимание дискурса» представляет собой короткую историю «Про кота» и 16 утверждений, которые соответствуют или не соответствуют тексту. Текст предзаписан в исполнении диктора и предъявляется на слух. Далее по одному предъявляются утверждения, и тестируемый отвечает, соответствует ли это утверждение истории. История «Про кота» создана специально для РАТ и не повторяет известные сюжеты из

кино, литературы, народной культуры и т.п., чтобы тестируемые не могли ориентироваться на эвристики и общие знания о мире во время ответов на вопросы (Caplan & Evans, 1990; Ferstl et al., 2005; Kintsch & Van Dijk, 1978; Van Dijk & Kintsch, 1983).

Текст содержит 156 слов (средняя частотность слов значимых частей речи - 261.32 ipm; медианная = 94.7 ipm, (Ляшевская & Шаров, 2009), и состоит из 31 клаузы (средняя длина клаузы в словах – 5.03, от 2 до 10 слов) с базовым порядком слов, 4 клаузы зависимые. Утверждения после текста расположены в псевдослучайном порядке и организованы в пары. Каждая пара состоит из двух утверждений, относящихся к одному факту из истории, одно из которых верно, а второе ложно. Каждая пара утверждений относится либо к основной сюжетной линии, либо к детали, а также либо к эксплицитной, либо имплицитной информации; подобное разделение вопросов есть в Тесте на понимание дискурса (DCT, Nicholas & Brookshire, 1993)). Балл дается за каждую пару ответов на утверждения в том случае, если оба ответа верные; таким образом, максимальный балл за субтест – 8 баллов.

В субтесте «Порождение дискурса» в качестве стимула для порождения речи используется специально созданный для теста рисунок «Велосипед». Как и сюжет текста «Про кота», сюжет рисунка не напоминает какой-либо известный в культуре сюжет. Инструкция сформулирована таким образом, чтобы побудить к порождению рассказа, то есть дискурсивной структуры с упорядоченными во времени элементами сюжета, а не описания отдельных элементов рисунка (Olness, 2006). Рассказ по рисунку оценивают по четырем шкалам (информативность, беглость, грамматическая сложность, парафазии) от 1 (сильные нарушения) до 5 (нет нарушений). Использование оценки по нескольким шкалам является оптимальным подходом оценки дискурса в клинической практике, так как позволяет оценить речь на нескольких языковых уровнях, но при этом не требует значительных временных затрат (Prins & Bastiaanse, 2004). Итоговый балл за субтест представляет собой сумму баллов по каждой шкале. Ноль баллов ставится, если тестируемый не произнес как минимум пять релевантных значимых слов в течение одной минуты.

Для оценки психометрических характеристик РАТ было проведено нормирование, валидация и тестирование надежности теста. В исследовании приняли участие 106 неврологически здоровых носителей русского языка (77 женщин; от 19 до 86 лет; средний возраст – 49.9, SD = 18.4) и 85 человек с афазией (26 женщин; от 25 до 80 лет; средний возраст – 57.6, SD = 12.1). Для оценки тест-ретестовой надежности «Понимание дискурса» и «Порождение дискурса» были дополнительно протестированы 20 человек (11 женщин,

от 39 до 82 лет; средний возраст – 58.8, SD = 12.9). Межэкспертная надежность была подсчитана на материале оценок 20 рассказов двумя логопедами. Внутренняя надежность была подсчитана только для субтеста «Понимание дискурса». Результаты оценки надежности субтестов представлены в Таблице 1.

Таблица 1. Психометрические характеристики субтестов «Понимание дискурса» и «Порождение дискурса» РАТ

		Понимание дискурса	Порождение дискурса
Внутренняя надежность	Cronbach alpha	0,83	-
	95% CI	[0,776, 0,885]	-
	ICC	-	0,833
Межэкспертная надежность	95% CI	-	[0,609-0,932]
	Эксперт 1 (Mean)	-	57,0%
	Эксперт 2 (Mean)	-	62,5%
	ICC	0,784	0,71
Тест-ретестовая надежность	95% CI	[0,53, 0,909]	[0,394-0,875]
	Тест 1 (Mean)	70,0%	75,5%
	Тест 2 (Mean)	76,9%	75,2%

Были проанализированы баллы за субтесты в двух возрастных группах: молодые взрослые (18-59 лет) и пожилые взрослые (60+ лет), см. Таблицу 2. В группе здоровых участников не было значимых различий в баллах за субтест «Порождение дискурса», в то время как в субтесте «Понимание дискурса» пожилые участники получили значимо более низкие баллы (критерий Уэлча для двух независимых выборок, $p_{adj} < .05/13 = .0038$). В обеих возрастных когортах участники с афазией получили более низкие баллы по обоим субтестам (критерий Уэлча для независимых выборок с $p_{adj} < .05/26 = .0019$).

Таблица 2. Баллы за субтесты «Понимание дискурса» и «Порождение дискурса»

Субтест	Возраст	Группа	N	От-до	Среднее (SD)	Медиана (IQR)
Понимание дискурса	Молодой	Контроль	69	62.5 - 100	94.36 (8.74)	100 (12.5)
		Афазия	40	0 - 100	63.44 (31.82)	75 (37.5)
	Пожилкой	Контроль	37	50 - 100	86.15 (14.06)	87.5 (25)
		Афазия	38	0 - 100	56.58 (32.33)	62.5 (50)
Порождение дискурса	Молодой	Контроль	69	75 - 100	90.65 (6.3)	90 (5)
		Афазия	38	0 - 95	46.97 (27.67)	50 (33.75)
	Пожилкой	Контроль	35	75 - 100	91.14 (5.83)	90 (5)
		Афазия	38	0 - 80	45 (30.78)	57.5 (70)

2. Лингвистические механизмы связности дискурса при афазии

Статья, выносимая на защиту: (Linnik et al., 2021)

Элементы дискурса связаны между собой определенным образом и организованны в соответствии с определенной иерархией и законами. Одним из важных качеств дискурса является связность – то, насколько текст может быть понятен читателю или слушателю. Это ментальный феномен, который включает в себя необходимость установления контакта между говорящим и адресатом (или репрезентацией адресата) и определения общего контекста, референции и тематической структуры (Jucker, 1997).

Механизмы нарушения связности при афазии и их связь с нарушениями на других языковых уровнях пока недостаточно изучены. Во многом это связано с тем, что различные исследователи дают разные определения связности и применяют разные методы ее измерения, а также тем, что большую роль в оценке связности играет интерпретация адресата (слушающего). В настоящем исследовании мы разработали шкалы оценки четырех аспектов связности, чтобы ответить на два исследовательских вопроса: 1) нарушена ли связность дискурса у рассказчиков с афазией? 2) какие микро- и макролингвистические параметры влияют на субъективно воспринимаемую связность?

Исследование было проведено на материале пересказов фильма «О грушах» (Chafe, 1980), которые позже вошли в состав корпуса Russian CliPS (Khudyakova et al., 2016). В исследовании приняли участие 10 человек с афазией (4 женщины; от 40 до 73 лет, средний возраст – 56,4) и 10 здоровых носителей языка (5 женщин; от 42 до 84 лет, средний возраст – 58,7). Аудиозаписи были оценены тремя экспертами по специально разработанной шкале. Каждый из четырех аспектов связности: информативность, ясность, понятность, связанность – оценивался по четырехбалльной шкале от 1 (полностью несвязный дискурс) до 4 (полностью связный дискурс). Аудиозаписи были размечены в формате СНАТ (Codes for the Human Analysis of Transcripts; MacWhinney, 2010) с сегментацией на элементарные дискурсивные единицы по критериям, принятым в работе (Carlson & Marcu, 2001). Транскрипты были размечены в соответствии с Теорией риторических структур (Mann & Thompson, 1987). Для каждого пересказа была проведена оценка количества основных событий в соответствии с критериями (Wright et al., 2010).

Для оценки вклада микро-и макролингвистических параметров в аспекты связности дискурса был сформирован набор из девяти метрик, извлеченных из разметки: пять микролингвистических параметров (количество информационных единиц в минуту,

количество ошибок на уровне слов, количество филлеров, количество неграмматических ЭДЕ, лексическое разнообразие) и четыре макролингвистических параметра (количество основных событий, количество структурных нарушений, набор риторических отношений, количество мета-комментариев). Для классификационного анализа оценки по четырехбалльной шкале были переведены в оценки по двухбалльной шкале (0 – дискурс несвязный, 1 – дискурс связный). Классификация была проведена с использованием метода случайного леса (random forests, Breiman, 2001) для всех 20 образцов дискурса.

Межгрупповое сравнение было проведено с помощью односторонних пермутационных тестов и выявило значимо более низкие оценки по всем аспектам связности в группе людей с афазией по сравнению с неврологически здоровыми рассказчиками. Результаты межгруппового сравнения для каждого аспекта связности, а также точность классификации и наиболее значимые предикторы представлены в Таблице 3.

Таблица 3. Дискурсивная связность: результаты межгруппового сравнения и классификационного анализа

Межгрупповое сравнение				
	p_{adj}	Медианная разница в оценках		
Информативность	0.04	1		
Ясность	0.04	1		
Понятность	0.002	1.5		
Связанность	0.002	2		

Классификационный анализ				
	Точность классификации	Карра	ООВ error	Наиболее значимые предикторы
Информативность	0.84	0.355	15%	- количество основных событий
				- набор риторических отношений
	0.8	0.57	20%	- количество мета-комментариев
				- количество неграмматических ЭДЕ
				- количество ошибок на уровне слов
Ясность				- количество неграмматических ЭДЕ
				- количество информационных единиц в минуту
				- лексическое

Понятность	0.8	0.34	20%	<ul style="list-style-type: none"> - разнообразие - количество ошибок на уровне слов - количество мета-комментариев - количество структурных нарушений - лексическое разнообразие - количество неграмматичных ЭДЕ
Связанность	0.95	0.89	5%	<ul style="list-style-type: none"> - Лексическое разнообразие - количество ошибок на уровне слов

В этом исследовании связность рассматривалась не столько как внутреннее неизменное качество текста, а как то, что создается не только говорящим, но и адресатом дискурса. Результаты исследования показали, что для различных аспектов связности наиболее важны различные наборы микро- и макролингвистических параметров дискурса.

3. Корпус 3D (Discourse Diversity Database)

В корпус 3D вошли образцы дискурса трех типов: рассказы по серии рисунков, личные истории, инструкции по серии рисунков. В качестве стимулов для рассказов по серии рисунков были отобраны три комикса Херлуфа Бидструпа с единственным главным персонажем (истории «Зимний день», «Приключение» и «Спортсмен»); для элиситации личной истории были отобраны три вопроса про самые лучшие или запоминающиеся подарок, путешествие и праздник; в качестве стимулов для элиситации инструкций по серии рисунков были отобраны инструкции по сборке мебели фирмы ИКЕА (стол, скамейка и табурет), каждая в виде серии из восьми картинок без сопровождающего текста.

3.1 Корпус 3D для исследований в области клинической лингвистики

Статья, выносимая на защиту: (Khudyakova et al., 2023)

Корпус 3D включает в себя несколько разделов с образцами устной речи людей с неврологическими и психиатрическими расстройствами, а также без нарушений. В корпусе представлены записи дискурса двух клинических групп: людей до и после удаления опухолей мозга и людей с психиатрическими нарушениями. Также в 3D есть три нормативных раздела: записи дискурса людей разных возрастных групп, записи дискурса

и результатов психиатрических опросников, а также раздел с записями дискурса людей с различной степенью усталости. Сбор данных корпуса в настоящий момент продолжается. Основные демографические данные, а также сведения о дополнительных мета-данных для каждого из разделов приведены в Таблице 4.

Таблица 4. Разделы корпуса 3D

	Разделы				
	Клинические группы		Нормативные группы		
	Нейрохирургия	Психиатрия	Разные возрастные группы	Психиатрические опросники	Усталость
Количество участников	87	107	86	76	10
Количество временных точек	3	1	1	1	2
Возраст	M=49.7, SD=14.6	M=28.8, SD=4.3	18-29 y.o. (M = 21.2, SD = 2.6); 30-49 y.o. (M = 38.1, SD = 6.6); 50-64 y.o. (M = 57, SD = 3.8); 65+ y.o. (M = 72, SD = 7.0)	M=23.9, SD=4.3	M=28.80, SD=2.86
Диагнозы	Опухоли мозга	Расстройства шизофренического спектра, аффективные расстройства	нет	возможно	нет
Метаданные	MPT; PAT; ИКЭф	ICD-10; PANSS; HDRS		SCL-90-R; AMSR; QIDS-16SR	CAH

Разметка аудиозаписей была проведена в программе ELAN Annotation Tool (Wittenburg et al., 2006). Аудио были затранскрибированы вручную с определением незаполненных и заполненных пауз. На отдельных слоях были размечены леммы и части речи на основании стандартов, применяемого в разметке Национального корпуса русского

языка (<https://ruscorpora.ru/>). Все аудиозаписи были разделены на элементарные дискурсивные единицы (ЭДЕ), примерно равные клаузе (то есть имеющие один предикат, или пропущенный предикат); ЭДЕ с синтаксически зависимыми от них ЭДЕ объединялись в высказывания (подробнее см. Bergelson & Khudyakova, 2020). На отдельных слоях были размечены ошибки, фальстарты, семантически пустые слова и автоматизированные выражения. Каждая ЭДЕ была размечена как относящаяся к основной теме (Основная линия, Фон или Комментарий) или не относящаяся (Метакомментарий, Регулятор).

3.1 Влияние степени усталости говорящего на характеристики устного дискурса

Статья, выносимая на защиту: (Худякова, 2020)

В современной логопедической практике и психо- и нейролингвистических исследованиях анализ различных характеристик устного дискурса используется для оценки тяжести речевых нарушений. Однако известно, что и у неврологически здоровых носителей языка такие параметры, как темп речи и различные акустические характеристики голоса могут изменяться под воздействием усталости, стресса и когнитивной нагрузки (Quatieri et al., 2015; Rao et al., 2020; Sloboda et al., 2018). Таким образом, чтобы оценить тяжесть нарушений при различных расстройствах, необходимо знать границы и вариативность речевой нормы. Для оценки этого вклада было проведено пилотное исследование речи людей с разной степенью усталости и собраны образцы дискурса, вошедшие в корпус 3D.

В исследовании приняли участие десять здоровых носителей русского языка (8 женщин и 2 мужчин; возраст от 23 до 33 лет, средний возраст – 28,8). Исследование проходило онлайн. Каждый рассказчик принимал участие в исследовании два раза: один раз – в активном, бодром состоянии, второй раз – в усталом состоянии. Помимо субъективного определения состояния, рассказчики заполняли протокол стандартизированного Теста дифференцированной самооценки функционального состояния САН (Доскин et al., 1973), предназначенный для оценки самочувствия, активности и настроения. Для каждого участника различие между баллами за активность по САН между двумя сессиями составляло не менее 1,5 баллов по семибалльной шкале. Во время каждой сессии участники исследования записывали один рассказ по серии рисунков, одну личную историю и одну инструкцию.

Анализ данных с помощью регрессионных моделей с фиксированными эффектами состояния, типа дискурса и их взаимодействия и случайным эффектом участника не выявил значимых эффектов взаимодействия типа дискурса и состояния говорящего.

Значимый эффект состояния говорящего был обнаружен для полного темпа речи ($p_{adj} = 7.20e-06$). Значимый эффект типа дискурса был обнаружен для следующих параметров: длительность в секундах ($p_{adj} = 3,57e-04$), количество слов ($p_{adj} = 0,038$), полный темп речи ($p_{adj} = 5e-07$), артикуляторный темп речи ($p_{adj} = 1,23e-07$), среднее количество слов в ЭДЕ ($p_{adj} = 0,006$), лексическое разнообразие ($p_{adj} = 0,042$) и пропорция значимых частей речи ($p_{adj} = 0,014$). Не был выявлен значимый эффект типа дискурса на длину пауз и среднее количество ЭДЕ в высказывании.

Результаты пилотного исследования позволили выявить вариативность фонетических и темповых характеристик речи в зависимости от усталости говорящего, а также значимые отличия между дискурсом разных жанров в характеристиках на разных языковых уровнях.

4. Заключение

Статьи, включенные в диссертацию, описывают разработанные новые методы и инструменты для исследования порождения и понимания дискурса у людей с неврологическими и психиатрическими расстройствами. Эти методы могут применяться как в клинической практике, так и в фундаментальных исследованиях.

В первой главе описаны субтесты для оценки понимания и порождения дискурса Русского афазиологического теста. Результаты стандартизации РАТ показали, что эти субтесты обладают необходимым уровнем надежности, а также результаты этих субтестов значимо различаются в группе людей с афазией и контрольной группе.

Во второй главе было описано исследование лингвистических механизмов дискурсивно й связности. Мы представили новую шкалу для оценки четырех аспектов дискурсивной связности и показали, что все эти аспекты нарушены при афазии. Также мы исследовали вклад различных микро- и макролингвистических параметров в каждый из аспектов связности. Классификационный анализ показал, что для каждого из аспектов связности наиболее значимыми являются различные комбинации параметров.

В третьей главе мы представили описание корпуса 3D, трех заданий для элиситации дискурса, включенных в протокол, параметрах, включенных в схему разметки. Было представлено описание текущего состояния всех разделов корпуса. Также были представлены результаты пилотного исследования, проведенного на материале одного из разделов 3D. Исследование показало, что на отдельные характеристики речи влияет степень усталости говорящего и тип дискурса.

Таким образом, материалы, входящие в настоящее диссертационное исследование, описывает новые инструменты для оценки и анализа дискурса на русском языке. В настоящее время проводится работа над расширением корпуса 3D и исследованием вариативности характеристик дискурса в зависимости от состояния говорящего и типа дискурса, а также над корреляцией оценки дискурса слушающим и объективными дискурсивными характеристиками. Работа, проведенная автором диссертации, носит междисциплинарный характер и объединяет традиции клинической, корпусной и экспериментальной лингвистики.

Литература

- Bastiaanse, R., Raaijmakers, S., Satoer, D., & Visch-Brink, E. (2015). The Multilingual Token Test. *Http://Dx.Doi.Org/10.1080/02687038.2015.1121710*, 30(4), 508–508.
<https://doi.org/10.1080/02687038.2015.1121710>
- Bergelson, M. B., & Khudyakova, M. V. (2020). Narrative discourse segmentation in clinical linguistics. In *In Search of Basic Units of Spoken Language. A Corpus-driven Approach*. (pp. 258–284). <https://doi.org/10.1075/scl.94.08ber>
- Breiman, L. (2001). Random Forests. *Machine Learning 2001 45:1*, 45(1), 5–32.
<https://doi.org/10.1023/A:1010933404324>
- Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, 30(7), 489–518.
<https://doi.org/10.3109/02699206.2016.1145740>
- Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology*, 31(10), 1105–1126.
<https://doi.org/10.1080/02687038.2016.1239013>
- Buivolova, O., Vinter, O., Bastiaanse, R., & Dragoy, O. (2020). The Aphasia Rapid Test: adaptation and standardisation for Russian.
Https://Doi.Org/10.1080/02687038.2020.1727836, 35(5), 730–744.
<https://doi.org/10.1080/02687038.2020.1727836>

- Caplan, D., & Evans, K. L. (1990). The effects of syntactic structure on discourse comprehension in patients with parsing impairments. *Brain and Language*, 39(2), 206–234. [https://doi.org/10.1016/0093-934X\(90\)90012-6](https://doi.org/10.1016/0093-934X(90)90012-6)
- Carlson, L., & Marcu, D. (2001). Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, 2, 1–87. <ftp://128.9.176.20/isi-pubs/tr-545.pdf>
- Carpenter, P. A., Miyake, A., & Just, M. A. (1995). Language comprehension: sentence and discourse processing. *Annual Review of Psychology*, 46(1), 91–120. <https://doi.org/10.1146/annurev.psych.46.1.91>
- Chafe, W. (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production* (W. Chafe (ed.)). Ablex.
- Ferstl, E. C., Walther, K., Guthke, T., & von Cramon, D. Y. (2005). Assessment of Story Comprehension Deficits After Brain Damage. *Journal of Clinical and Experimental Neuropsychology*, 27(3), 367–384. <https://doi.org/10.1080/13803390490515784>
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *BDAE-3: Boston Diagnostic Aphasia Examination* (Third Edit). Lippincott Williams & Wilkins.
- Ivanova, M. V., Akinina, Y. S., Soloukhina, O. A., Iskra, E. V., Buivolova, O. V., Chrabaszcz, A. V., Stupina, E. A., Khudyakova, M. V., Akhutina, T. V., & Dragoy, O. (2021). The Russian Aphasia Test: The first comprehensive, quantitative, standardized, and computerized aphasia language battery in Russian. *PLOS ONE*, 16(11), e0258946. <https://doi.org/10.1371/JOURNAL.PONE.0258946>
- Jucker, A. H. (1997). Coherence in spontaneous text. *Journal of Pragmatics*, 27(3), 400–402. [https://doi.org/10.1016/S0378-2166\(97\)83633-8](https://doi.org/10.1016/S0378-2166(97)83633-8)
- Khudyakova, M., Antonova, N., Nelubina, M., Surova, A., Vorobyova, A., Minnigulova, A., Gronskaya, N., Yashin, K., Medyanik, I., Shishkovskaya, T., Ryazanskaya, G., Zuev, A., & Dragoy, O. (2023). Discourse Diversity Database (3D) for Clinical Linguistics Research: Design, Development, and Analysis. *Bakhtiniana: Revista de Estudos Do Discurso*. <https://doi.org/10.1590/2176-4573e55885>
- Khudyakova, M. V., Bergelson, M. B., Akinina, Y. S., Iskra, E. V., Toldova, S., & Dragoy, O. V. (2016). Russian CliPS: a Corpus of Narratives by Brain-Damaged Individuals. *LREC Proceedings*.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production.

Psychological Review, 85(5), 363–394. <https://doi.org/10.1037/0033-295X.85.5.363>

Linnik, A., Bastiaanse, R., Stede, M., & Khudyakova, M. (2021). Linguistic mechanisms of coherence in aphasic and non-aphasic discourse. *Aphasiology*, 36(2), 123–146.

<https://doi.org/10.1080/02687038.2020.1852527>

MacWhinney, B. (2007). The Talkbank Project. *Creating and Digitizing Language Corpora*, 163–180. https://doi.org/10.1057/9780230223936_7

MacWhinney, B. (2010). Part 1: The CHAT Transcription Format. In *The CHILDES Project: Tools for Analyzing Talk*. <https://doi.org/10.1111/1460-6984.12101/abstract>

Mann, W. C., & Thompson, S. A. (1987). *Rhetorical structure theory: Description and construction of text structures*. Springer Netherlands.

Mar, R. A. (2004). The neuropsychology of narrative: story comprehension , story production and their interrelation. *Neuropsychologia*, 42(10), 1414–1434.

<https://doi.org/10.1016/j.neuropsychologia.2003.12.016>

Nicholas, L E, & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research*, 36(2), 338–350.

Nicholas, Linda E, & Brookshire, R. H. (1995). Comprehension of Spoken Narrative Discourse by Adults With Aphasia, Right-Hemisphere Brain Damage, or Traumatic Brain Injury.

American Journal of SpeechLanguage Pathology, 4(3), 69–81.

<http://ajslp.asha.org/cgi/content/abstract/4/3/69>

Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, 20(2/3/4), 175–187. <https://doi.org/10.1080/02687030500472710>

Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers.

Aphasiology, 18(12), 1075–1091. <https://doi.org/10.1080/02687030444000534>

Quatieri, T. F., Williamson, J. R., Smalt, C. J., Patel, T., Perricone, J., Mehta, D. D., Helfer, B. S., Ciccarelli, G., Ricke, D., Malyska, N., Palmer, J., Heaton, K., Eddy, M., & Moran, J. (2015). *Vocal biomarkers to discriminate cognitive load in a working memory task*.

Rao, H. M., Smalt, C. J., Rodriguez, A., Wright, H. M., Mehta, D. D., Brattain, L. J., Edwards, H. M., Lammert, A., Heaton, K. J., & Quatieri, T. F. (2020). Predicting Cognitive Load and Operational Performance in a Simulated Marksmanship Task. *Frontiers in Human Neuroscience*, 14, 222. <https://doi.org/10.3389/fnhum.2020.00222>

- Schank, R. C. (1995). *Tell me a story: Narrative and intelligence*. Northwestern University Press.
- Sloboda, J., Lammert, A., Williamson, J., Smalt, C., Mehta, D. D., Curry, C. I., Heaton, K., Palmer, J., & Quatieri, T. (2018). Vocal Biomarkers for Cognitive Performance Estimation in a Working Memory Task. *Interspeech 2018, 2018-Septe*, 1756–1760.
<https://doi.org/10.21437/Interspeech.2018-2418>
- Swinburn, K., Porter, G., & Howard, D. (2004). *CAT: comprehensive aphasia test*.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*.
- Varlokosta, S. (2016). A Greek Corpus of Aphasic Discourse: Collection, Transcription, and Annotation Specifications. *LREC 2016 Workshop Resources and Processing of Linguistic and Extra-Linguistic Data from People with Various Forms of Cognitive/Psychiatric Impairments (RaPID-2016)*, May, 14–21.
- Williams, C., Thwaites, A., Buttery, P., Geertzen, J., Randall, B., Shafto, M., Devereux, B., & Tyler, L. K. (2010). The Cambridge Cookie-Theft corpus: A corpus of directed and spontaneous speech of brain-damaged patients and healthy individuals. *LREC 2010. Proceedings of the 7th International Conference on Language Resources and Evaluation, November 2015*, 2824–2830. <http://www.lrec-conf.org/proceedings/lrec2010/summaries/327.html>
- Wilson, S. M., Eriksson, D. K., Schneck, S. M., & Lucanie, J. M. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PLOS ONE*, 13(2), e0192773. <https://doi.org/10.1371/journal.pone.0192773>
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*, 1556–1559.
<https://tla.mpi.nl/tools/tla-tools/elan/>
- Wright, H. H., Capilouto, G. J., Wagovich, S. A., Cranfill, T. B., & Davis, J. E. (2010). Development and reliability of a quantitative measure of adults' narratives. <Http://Dx.Doi.Org/10.1080/02687030444000732>, 19(3–5), 263–273.
<https://doi.org/10.1080/02687030444000732>
- Буйволова, О. В., Бастиансе, И. Р. М., Драгой, О. В., Винтер, О. Б., Позднякова, В. А., Самоукина, А. А., Шляхова, А. А., & Виш-Бринк, Э. (2020). Adaptation of the Aphasia Bedside Check for Russian. *Российский Журнал Когнитивной Науки*, 7(3), 45–67.

- Доскин, В. А., Лаврентьева, Н. А. Мирошников, Н. П., & Шарай, В. Б. (1973). Тест дифференцированной самооценки функционального состояния. *Вопросы Психологии*, 6, 141–145.
- Ляшевская, О. Н., & Шаров, С. А. (2009). *Частотный словарь современного русского языка (на материалах Национального корпуса русского языка)*. Азбуковник.
- Худякова, М. В. (2020). Влияние степени усталости говорящего на характеристики устного дискурса. *Российский Журнал Когнитивной Науки*, 7(3), 78–88.