

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

На правах рукописи

Здорова Нина Станиславовна

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЧТЕНИЯ В РАЗЛИЧНЫХ  
ПОПУЛЯЦИЯХ

Резюме диссертации  
на соискание ученой степени  
кандидата филологических наук

Научный руководитель:  
доктор филологических наук  
Драгой Ольга Викторовна

Москва 2024

Диссертация была подготовлена в Национальном исследовательском университете «Высшая школа экономики», Москва, Россия.

## **Публикации**

На защиту выносятся три публикации, описывающие исследования чтения в различных популяциях, где изучаются факторы, влияющие на процесс чтения:

1. Parshina O., Lopukhina A., Sofya Goldina, Ekaterina Iskra, Serebryakova M., Staroverova V., **Zdorova N.**, Dragoy O. Global reading processes in children with high risk of dyslexia: a scanpath analysis. // *Annals of Dyslexia*. 2022. Vol. 72. No. 3. P. 403-425. <https://doi.org/10.1007/s11881-021-00251-z>
2. **Zdorova N.**, Malyutina S., Laurinavichyute A., Kaprielova A., Anastasia Zuibanova, Lopukhina A. Do we rely on good-enough processing in reading under auditory and visual noise? // *Plos One*. 2023. Vol. 18. No. 1. Article e0277429. <https://doi.org/10.1371/journal.pone.0277429>
3. **Zdorova N.**, Parshina O., Ogly B., Bagirokova I., Krasikova E., Ziubanova A., Unarokova Sh., Makerova S. and Dragoy O. (2023) Eye movement corpora in Adyghe and Russian: an eye-tracking study of sentence reading in bilinguals. // *Front. Psychol.* 14:1212701. <https://doi.org/10.3389/fpsyg.2023.1212701>

Нина Станиславовна Здорова является первым автором двух из перечисленных статей, опубликованных в журналах второго квартиля (Q2) и входящих в список А высокорейтинговых журналов ВШЭ. В статье Parshina et al. (2022), опубликованной в журнале первого квартиля (Q1), вклад Нины Станиславовны Здоровой включает сбор данных, участие в разработке исследования и обсуждении результатов.

## **Доклады на конференциях**

Предварительные и итоговые результаты исследования были представлены на научных конференциях, приведённых ниже.

1. AMLaP 2023: Architectures and Mechanisms of Language Processing (San-Sebastian, Spain, 2023). Poster presentation: Lexical and morphological effects on eye movements while reading a sentence corpus in a polysynthetic language.
2. Когнитивная наука в Москве 2023: новые исследования (Москва). Доклад: Паттерны движений глаз при чтении предложений у русскоязычных школьников и взрослых.
3. Когнитивная наука в Москве 2023: новые исследования (Москва). Доклад: Базовые характеристики движения глаз при чтении на полисинтетическом языке среди адыгейско-русских билингвов.
4. The 36th Annual Conference on Human Sentence Processing (Pittsburgh, USA, online, 2023). Poster presentation: Global reading patterns in Russian-speaking children and adults: a scanpath analysis.
5. Проблемы онтолингвистики-2022 (Санкт-Петербург, онлайн, 2022). Доклад: Формирование чтения у русскоязычных школьников 1-6 классов по данным видеоокулографии.
6. 9th Summer Neurolinguistics School 2022 on Experimental Studies of Russia's Ethnic Languages (Moscow, Russia, online, 2022). Oral presentation: Reading in morphologically distinct languages: an eye-tracking study in Russian-Adyghe bilingual children and adults.
7. SRCO 2021 Biennial Meeting (USA, online, 2021). Poster presentation: The Impact of Phonological and Orthographic Processing on Reading Development in Russian Children.

## **Введение**

Чтение – это сложный когнитивный процесс, который включает в себя декодирование графической информации, её соотнесение с фонетической формой, извлечение нужной единицы из ментального лексикона, синтаксическую интеграцию и, наконец, концептуализацию письменной информации, которая приводит к пониманию. Фундаментальные исследования чтения направлены на изучение механизмов развития навыков чтения, моделирование процесса чтения, когда этот навык уже сформирован, и выявление психолингвистических факторов, влияющих на чтение.

Развитие навыков чтения проходит через несколько этапов от буквенного и слогового чтения к целостному чтению слов и предложений, что в конечном итоге должно приводить как к беглости чтения, так и к пониманию прочитанного. Годы психолингвистических исследований, использовавших поведенческие методы и видеоокулографию (метод удалённой регистрации движений глаз), показывают менее беглое, более долгое и менее точное чтение у детей с нарушениями чтения по сравнению с типично читающими детьми (см. Barrington, 2019 для обзора). Однако вопрос о том, обязательно ли атипичное чтение подразумевает качественно разные этапы развития навыков чтения, остается дискуссионным. Иной взгляд на проблему говорит о том, что атипичное чтение отличается от типичного лишь количественной задержкой.

Уже развитый навык чтения у взрослых ставит перед психолингвистикой иные вопросы — о модели языковой обработки при чтении. Так, согласно теории коммуникации, канал связи неизбежно зашумлён (Shannon, 1948), где шум понимается широко. При наличии шума наша языковая обработка склонна опираться на семантику больше, чем на действительные синтаксические связи (Gibson et al., 2013; Levy 2008, 2011). Бóльшая опора на семантику и соответствие описанного условиям реальности постулируется и в модели поверхностной обработки независимо от шума, которая преимущественно апробируется на семантически неправдоподобных предложениях (Ferreira 2003; Ferreira et al., 2002). Однако гипотеза о

поверхностной обработке практически не изучалась на других языках, отличных от английского, и не была проверена в более натуралистических условиях языковой обработки с зашумленным каналом.

Наконец, механизмы чтения у взрослых на разных языках, согласно видеоокулографическим исследованиям, имеют как общие универсальные черты (напр., значимое влияние частотности и длины слов), так и демонстрируют различия в базовых характеристиках движений глаз (ср. Kliegl et al., 2004 на материале немецкого и Laurinavichyute et al., 2019 на материале русского). Однако исследования, посвященные универсальности или языкоспецифичности механизмов чтения, проводятся преимущественно на материале широко изученных языков. Важно отметить, что на сегодня до сих пор нет экспериментальных данных о движениях глаз при чтении на полисинтетических языках, что обуславливает необходимость проверки универсальности многих лексических и морфологических факторов на материале этих языков в том числе.

Настоящая диссертация представляет три видеоокулографических исследования чтения, описанных в разделах 1-3 соответственно. Раздел 1 описывает исследование развития навыков чтения у детей (Parshina et al., 2022). Раздел 2 описывает тестирование модели поверхностной обработки при чтении у монолингвальных взрослых (Zdorova et al., 2021). Раздел 3 излагает исследование базовых характеристик движений глаз при чтении у билингвальных взрослых, читающих на двух типологически различных языках (Zdorova et al., 2023). Одно исследование применяет классический экспериментальный дизайн 2x2, а два других исследования используют корпусный подход, который всё больше распространён среди кросс-лингвистических исследований движения глаз при чтении.

**Целью** диссертационного исследования является выявление того, как возраст и психолингвистические особенности модулируют механизмы чтения. Отсутствие кросс-лингвистических исследований в данном направлении психолингвистики с применением как экспериментальных, так и корпусных

методов определяет **актуальность** настоящей работы, которая расширяет фундаментальные исследования чтения в сторону малоизученных типологически различных языков. **Объектом** исследования является чтение как сложный когнитивный процесс. Возраст и психолингвистические факторы, влияющие на чтение, составляют **предмет** исследования. В качестве основного метода используется регистрация движений глаз (видеоокулография), а также применяются статистические методы для анализа данных в среде R (R Core Team, 2020).

**Теоретическая значимость** исследования заключается в расширении фундаментальных исследований с использованием видеоокулографии в сторону малоизученных и типологически различных языков. Экспериментальные данные на этих языках позволяют проверить языкоспецифичность моделей чтения и психолингвистические факторы, которые предположительно влияют на чтение универсально во всех языках. Более того, исследование развития навыков чтения применяет передовые методы анализа данных движений глаз (анализ рисунков чтения или анализ сканпафов, от англ. scanpaths analysis), что дает нам представление о языковой обработке на глобальном уровне предложения, а не только на уровне слова.

**Практическая значимость** исследования заключается в следующем. Во-первых, рисунки чтения (сканпафы), описанные у типично читающих детей и детей с высоким риском дислексии, предоставляют специалистам в области образования и логопедии более глубокое понимание механизмов чтения при дислексии. Это важный этап в диагностике дислексии, открывающий поле для будущего развития автоматизированных инструментов диагностики дислексии на основе данных о движениях глаз при чтении. Во-вторых, результаты экспериментального исследования языковой обработки при чтении и наличии внешнего шума могут быть применены в образовании, маркетинге и дизайне, где влияние внешнего шума (аудиального и визуального) должно учитываться для более эффективного восприятия целевой информации клиентом. Наконец, Адыгейский корпус предложений

(ASC, от англ. Adyghe Sentence Corpus), созданный в рамках третьего исследования и выложенный в свободном доступе на платформе [osf](#), может быть использован исследователями, педагогами и студентами в научных и учебных целях.

**Ключевые результаты** исследования и **положения**, выносимые на защиту:

1. Развитие навыков чтения проходит качественно схожие этапы как у типично читающих детей, так и у детей с высоким риском дислексии, при этом у последних наблюдается количественное 3-летнее отставание в чтении.
2. Взрослые носители русского языка полагаются на поверхностную обработку при чтении, однако их опора на неё не модулируется внешним языковым шумом (аудиальным или визуальным).
3. Чтение на полисинтетическом языке (адыгейском) зависит от психолингвистических характеристик, которые считаются универсально значимыми, однако вклад данных характеристик в чтение на синтетическом (русском) и полисинтетическом (адыгейском) языке для адыго-русских билингвов может быть различным.

## **1. Паттерны развития навыков чтения у типично читающих детей и детей с риском дислексии на материале данных о движениях глаз**

Статья, выносимая на защиту: Parshina et al. (2022)

Данное исследование направлено на понимание особенностей чтения у типично читающих детей и детей с дислексией. В работе исследуется, отличаются ли (и если да, то как) базовые характеристики движений глаз при чтении и глобальные рисунки чтения предложений в этих группах детей с 1 по 5 классы. Кроме того, исследование отвечает на вопрос о критическом возрасте в развитии чтения, в котором эти различия между группами проявляются.

Многочисленные исследования движений глаз при чтении показывают, что дети с трудностями в чтении (включая дислексию) делают более длительные фиксации на словах, больше регрессий к предыдущим словам и более короткие саккады вправо (Barrington, 2019 for review, De Luca et al., 2002; Hatzidaki et al., 2010; Hawelka & Wimmer, 2005; Lefton et al., 1979). Более того, они склонны пропускать мало слов и чаще перечитывать их по сравнению с типично читающими ровесниками (De Luca et al., 1999, 2002; Hawelka et al., 2010; Hutzler & Wimmer, 2004).

Однако другие исследования (Hyönä & Olson, 1995; Rayner, 1985a, b, 1998) показывают, что различия в движениях глаз при чтении между данными группами детей исчезают, если сопоставить читающих на основании их скорости чтения, а не на основании их возраста. Rayner (1998) утверждает, что наблюдаемые различия в движениях глаз являются следствием не качественной, а количественной задержки в развитии навыков чтения, происходящей в связи с дефицитами языковой обработки (cf. Hutzler et al., 2006).

Настоящее исследование чтения было проведено в двух группах детей с 1 по 5 класс: в группе детей с высоким риском дислексии ( $N = 72$ , средний возраст = 9.47,  $SD = 1.24$ ) и в контрольной группе типично читающих детей ( $N$

= 72, средний возраст = 9.35, SD = 1.14). Отнесённость ребёнка к группе определялась на основании его результатов чтения вслух из Стандартизированной методики исследования навыков чтения на русском языке (СМИНЧ, Корнев, 1997). Участники читали 30 предложений детской версии (Korneev et al., 2017) Русского корпуса предложений (Laurinavichyute et al., 2019) и отвечали на вопросы по ним. Во время чтения предложений с экрана компьютера движения их глаз записывались с помощью видеоокулографа EyeLink 1000+ или EyeLink Portable Duo (SR Research) с частотой дискретизации 1000 Гц.

Анализ данных в R (R Core Team, 2020) включал в себя сравнение базовых характеристик движений глаз между двумя группами. К базовым характеристикам относятся следующие глазодвигательные меры: (FFD, от англ. First Fixation Duration — длительность первой фиксации на слове; SFD, от англ. Single Fixation Duration — длительность единственной фиксации на слове; GD, от англ. Gaze Duration — сумма длительности фиксаций при первом прочтении слова; TT, от англ. Total Time reading — общее время прочтения слова; пропуск слова и вероятность регрессии на слово). Для сравнения чтения между группами была использована серия (обобщенных) линейных моделей, где одна из перечисленных глазодвигательных мер была зависимой переменной, а принадлежность к группе, класс учащегося, частотность слова, длина слова, четыре взаимодействия 2\*2 (группа X длина слова, группа X частотность слова, класс X длина, класс X группа), а также относительное положение слова в предложении являлись предикторами. Случайными эффектами послужили номер участника, номер предложения и слово.

Результаты анализа показали статистически значимые различия между группами в мерах общего времени прочтения слова (TT) и сумме длительности фиксаций при первом прочтении (GD): дети с риском дислексии дольше фиксировались на словах, чем контрольная группа (все  $ps < 0.004$ ). Однако отличий не было в длительности первой фиксации. Класс учащихся оказался значимым предиктором ряда глазодвигательных мер (FFD, GD, TT) для обеих

групп: длительность фиксаций снижалась от класса к классу (все  $p_s < 0.001$ ). Однако длительность единственной фиксации (SFD) сокращалась в классах старше лишь в контрольной группе.

Мы обнаружили значимое взаимодействие между группой и классом, влияющее на вероятность пропуска слова. А именно, в то время как с увеличением возраста детей в контрольной группе вероятность пропуска слов увеличивалась, в группе с высоким риском дислексии она стремилась к нулю ( $p = 0.003$ ). Кроме того, дети с риском дислексии не пропускали короткие слова, как это делали типично читающие дети ( $p < 0.001$ ).

После сравнения базовых характеристик движений глаз мы проанализировали глобальные паттерны чтения, используя анализ рисунков чтения всего предложения, или анализ сканпафов (от англ. scanpath analysis). Он включает в себя создание рисунков чтения каждого предложения на каждого участника и применение кластерного анализа для выявления кластеров схожих паттернов (von der Malsburg & Vasishth, 2011). В результате данного вида анализа мы выявили пять паттернов чтения.

Первый паттерн чтения – паттерн беглого чтения – напоминает стратегию чтения взрослых, у которых навык чтения сформирован и развит: короткие фиксации (GD = 426 мсек), пропуск некоторых слов (23%) и малое количество регрессий (14%). Два следующих паттерна продвинутого чтения похожи на беглый паттерн с основным отличием в длительности фиксаций (GD = 585 мсек и GD = 716 мсек соответственно). Промежуточный паттерн характеризуется ещё более длительными фиксациями (GD = 808 мсек) и бóльшим количеством коротких регрессий (22%). Паттерн «начинающего» характеризуется глобальными трудностями в интеграции информации и общем понимании, что выражается в множественных перекидываниях не только отдельных слов, но и всего предложения целиком (вероятность регрессии = 26%, в среднем каждое слово перечитывается по 3 раза, GD = 1004 мсек, общее время прочтения предложения = 20 сек).

Анализ пяти паттернов чтения показал качественное сходство в этапах овладения чтением между группами. Дети из контрольной группы быстро прогрессировали в чтении и уже к 4-му классу демонстрировали паттерн беглого чтения. Дети с риском дислексии начинали овладение чтением с того же начинающего паттерна, как и их сверстники-первоклассники из контрольной группы. Однако далее переход от одного паттерна чтения к другому происходил медленнее, чем у типично читающих детей. Со 2-го по 4-ый классы они демонстрировали преимущественно промежуточные паттерны и достигали продвинутого паттерна лишь к 5-му классу, который дети из контрольной группы достигали уже во 2ом классе.

Итак, в данной работе мы обнаружили, что главное различие в механизмах чтения между типично читающими детьми и детьми с риском дислексии заключается в длительности фиксаций на словах. Анализ сканпафов показал, что несмотря на количественные межгрупповые различия в глазодвигательных мерах, качественно дети в группе с риском дислексии проходили те же этапы овладения чтением, как и типично читающие сверстники, но с задержкой в 3 года.

## **2. Модель поверхностной языковой обработки при чтении: влияние внешнего шума**

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

Данное исследование тестирует модель поверхностной языковой обработки при чтении в условиях шума, который, согласно теории коммуникации, считается неизбежной особенностью коммуникативного потока (Shannon, 1948). В широком смысле шум — это любое нарушение канала связи или любой дополнительный сигнал, который мешает целевому сигналу. Согласно модели «зашумленного» канала (Levy, 2008; 2011; Gibson et al., 2013), в условиях шума человек воспринимает информацию поверхностно и опирается на семантику, а не на синтаксические отношения между словами.

Вместе с тем, согласно модели поверхностной обработки (Ferreira 2003; Ferreira et al., 2002; Ferreira & Patson, 2007), независимо от наличия или отсутствия шума, основным фактором для понимания языковой информации является соответствие описываемой ситуации привычным для жизни сценариям. Когда человек сталкивается с предложением, у него одновременно запускаются два механизма обработки: алгоритмический, основанный на синтаксисе, и семантический, основанный на значении слов. Семантическая обработка может быть завершена быстрее, если возникшее понимание соответствует реальным знаниям о мире.

Тестирование модели поверхностной обработки, как правило, проводится на материале семантически неправдоподобных предложений, напр. *Собака была укушена человеком* (перевод примера с английского из Ferreira 2003, Ferreira & Stacey, 2000). Треть участников (32%) ошибочно определяют агенса в данном предложении (Ferreira & Stacey, 2000), склоняясь к семантически более правдоподобной интерпретации, а не к описанной в предложении.

Так, целью двух видеоокулографических экспериментов было проверить опору на семантику при чтении предложений на русском языке (в соответствии с моделью поверхностной обработки) и изменение опоры на семантику при сопутствующем шуме (аудиальном и визуальном) в группе монолингвальных носителей русского языка (в соответствии с моделью зашумленного канала). Эксперименты 1 и 2 основаны на идентичном материале предложений и процедуре исследования с двумя чередующимися сессиями (с шумом и без), но были проведены с разными видами шума (аудиальным и визуальным) и в разных группах участников. Мы ожидаем, что бóльшая опора на семантику будет проявляться в большем количестве ошибок в вопросах после семантически неправдоподобных предложений (независимо от наличия шума).

Эффект шума, в свою очередь, может быть двояким. С одной стороны, участники могут отдавать предпочтение скорости, а не пониманию. Тогда быстрое и поверхностное чтение приведёт к бóльшему количеству ошибок в понимании как в семантически правдоподобных, так и в неправдоподобных предложениях. Это может привести к увеличению времени прочтения из-за шума и ухудшению понимания. С другой стороны, участники могут приоритизировать правильность ответов и читать предложения медленнее, чтобы компенсировать повышенную когнитивную нагрузку. Замедление чтения может помочь им в правильной интерпретации предложений (такой паттерн был отмечен ранее в работах Cauchard et al., 2012; Hyönä, & Eklholm, 2016; Kemper et al., 2008; Vasilev et al., 2019; Yan et al., 2018).

Материалами чтения были 56 русских предложений с причастным оборотом, где мы манипулировали падежным согласованием и семантическим правдоподобием описываемой ситуации. После каждого экспериментального предложения следовал вопрос на понимание. Пример экспериментального блока приведен ниже. Предложения (1) и (2) являются семантическими правдоподобными, а предложения (3) и (4) – семантическими неправдоподобными. Причастие в предложениях (1) и (3) согласуется с

вершиной именной группы, а в предложениях (2) и (4) – со вторым элементом этой группы. Однако стоит отметить, что предпочтение в отнесении причастия к вершине или второму элементу именной группы не было в фокусе нашего анализа, так как экспериментальное исследование Chernova et al. (2016) показало, что в русском языке носителям предпочтительно раннее закрытие, а исследование Lopukhina et al. (2022a) показало, что семантическое правдоподобие влияет на процесс обработки обоих видов предложений, но не меняет предпочтения к верхнему закрытию.

- (1) *Дима работал с доктором президента, лечащим* маленьких детей.  
Dima worked with **the doctor (Instr, masc)** of the President (Gen, masc),  
**\*who treat-PART\* (Instr, masc)** small children.  
*Кто лечил маленьких детей? — Доктор / Президент*  
Who treated small children? — **Doctor / President**
- (2) *Дима работал с доктором президента, управляющего* целой страной.  
Dima worked with the doctor (Instr, masc) of **the President (Gen, masc)**,  
**\*who run-PART\* (Gen, masc)** an entire country.  
*Кто управлял целой страной? — Доктор / Президент*  
Who ran an entire country? — **Doctor / President**
- (3) *Дима работал с доктором президента, управляющим* целой страной.  
Dima worked with the **doctor (Instr, masc)** of the President (Gen, masc),  
**\*who run-PART\* (Instr, masc)** an entire country.  
*Кто управлял целой страной? — Доктор / Президент*  
Who ran an entire country? — **Doctor / President**
- (4) *Дима работал с доктором президента, лечащего* маленьких детей.  
Dima worked with the doctor (Instr, masc) of the **President (Gen, masc)**,  
**\*who treat-PART\* (Gen, masc)** small children.  
*Кто лечил маленьких детей? — Доктор / Президент*  
Who treated small children? — **Doctor / President**

Кроме того, стимулы включали 128 филлеров, отличающихся от экспериментальных предложений по структуре. Стимулы были равномерно

распределены на 2 набора, которые предъявлялись в сессии с шумом и без. Каждый набор стимулов состоял из 28 экспериментальных и 64 филлерных предложений, что в совокупности дало 4 экспериментальных листа, составленных по латинскому квадрату.

В эксперименте 1 с аудиальным шумом приняли участие 71 человек (из них 38 женщин, средний возраст участников — 22 года,  $SD = 4.9$ ; в диапазоне 20–40 лет; среднее количество лет образования = 14, в диапазоне 11–20 лет). Участникам эксперимента необходимо было читать про себя предложения с экрана компьютера и отвечать на вопросы по их содержанию. При чтении предложений движения глаз участников регистрировались с помощью видеоокулографа EyeLink1000 Plus. В сессии с шумом во время чтения предложений в наушниках у участников играл аудиальный шум – наложенные друг на друга звуковые дорожки подкастов.

В эксперименте 2 с визуальным шумом приняли участие 70 человек (из них 30 женщин, средний возраст участников — 23 года,  $SD = 5.5$ , в диапазоне 20–40 лет; среднее количество лет образования = 14.5, в диапазоне 11–22). Процедура эксперимента была идентична эксперименту 1, однако в сессии с шумом использовался визуальный шум— в виде идиом и устойчивых словосочетаний длиной от двух до пяти слов, появлявшихся вокруг целевого предложения на 300–400 мсек. Вокруг одного стимульного предложения в любой точке экрана могли появиться случайно выбранные 3-4 идиомы.

Анализ данных в R (R Core Team, 2020) включал подсчет правильных ответов на вопросы и анализ трех базовых глазодвигательных мер (FFD, GD, TT). Эти меры анализировались в двух регионах интереса: на причастии и предшествующем ему существительном. Структура модели с правильностью ответов как зависимой переменной была следующей  $accuracy \sim plausibility * noise + (1 + plausibility || unique.item) + (1 + plausibility + noise || subject.id)$ . Структура модели для глазодвигательных мер в качестве зависимой переменной была следующей:  $\log(eyetrackingmeasure\_AOI) \sim plausibility *$

noise + plausibility:accuracy + length.centered + frequency.centered + (1 + plausibility + noise || unique.item) + (1 + plausibility + noise || subject.id).

Результаты анализа, проведенного в среде R (R Core Team, 2020), для эксперимента 1 показали, что понимание семантически неправдоподобных предложений было статистически ниже, чем правдоподобных ( $p < 0.001$ ), однако правдоподобие не влияло на длительность фиксации в двух регионах интереса, на причастии и предшествующем ему существительном. Понимание предложений при аудиальном шуме оставалось сохранным, но аудиальный шум значимо влиял на время прочтения предложений: увеличивал длительность первой фиксации ( $p < 0.001$  в обоих регионах интереса) снижал GD ( $p < 0.001$  в обоих регионах интереса). Вероятно, таким образом при более длительной первой фиксации испытуемые компенсировали вызванную шумом когнитивную нагрузку, но затем справлялись с ней и не нуждались в длительных последующих фиксациях.

Мы не обнаружили значимого взаимодействия между семантическим правдоподобием и шумом, но обнаружили взаимодействие между семантическим правдоподобием и уровнем понимания в общем времени прочтения (ТТ) регионов интереса:  $p = 0.008$  на причастии и  $p < 0.001$  для предшествующего существительного. Сохранное понимание семантически неправдоподобных продолжений требовало от участников более длительного общего времени прочтения причастий.

Результаты эксперимента 2 с визуальным шумом тоже показали ухудшение понимания семантически неправдоподобных предложений в сравнении с правдоподобными ( $p < 0.001$ ), хотя на длительность фиксации правдоподобие не влияло. Мы обнаружили изменение общей скорости чтения причастий в условиях визуального шума: длительность фиксации была меньше как при первом прочтении причастия, так и в общем времени прочтения (оба  $ps = 0.01$ ). Вероятно, такое ускорение могло быть вызвано желанием участников выполнить задание по чтению как можно быстрее, так как визуальный шум делает чтение неудобным. В то же время не было

обнаружено значимого влияния шума на точность понимания. Испытуемым удавалось прочитывать предложения быстрее, сохранив при этом высокий уровень понимания. Как и в эксперименте 1, мы не обнаружили значимого взаимодействия между сематическим правдоподобием и шумом, но обнаружили взаимодействие между сематическим правдоподобием и уровнем понимания в общем времени прочтения (ТТ) причастия:  $p = 0.004$ .

Результаты исследования отчасти подтверждают предыдущие работы по поверхностной языковой обработке (Ferreira et al., 2002; Ferreira & Patson, 2007): при чтении на русском языке участники опираются на семантику в первую очередь, но ни аудиальный, ни визуальный шум не увеличивали опору читателей на семантику — чтение не становилось более поверхностным. Важно отметить, что мы обнаружили различное влияние шума на чтение: аудиальный шум отвлекал читателей на первых этапах языковой обработки, но позволял ускориться на поздних этапах, в то время как при визуальном шуме участники в целом ускорялись в чтении. Однако мы должны признать, что аудиальный и визуальный шум отличались по своей природе, что не позволяет нам напрямую определить, различаются ли эффекты аудиального и визуального шума ввиду их различной модальности или ввиду других факторов (выраженности, статическому/динамическому характеру).

### **3. Базовые характеристики движений глаз билингвов при чтении на полисинтетическом и синтетическом языках**

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

Настоящее исследование расширяет кросс-лингвистические исследования движений глаз при чтении, изучая малый язык (адыгейский) и типологически различные языки (полисинтетический адыгейский и синтетический русский) с кириллической письменностью. Годы психолингвистических исследований чтения, использующих метод видеоокулографии, показывают значимое влияние лексических характеристик слов на движения глаз при чтении. Так, частотность, длина и предсказуемость слова влияют как минимум на длительность фиксаций и на вероятность пропуска слова (Inhoff & Rayner, 1986; Rayner, 1998; Staub & Rayner, 2007). Кроме того, морфологические и морфосинтаксические характеристики тоже влияют на движения глаз и, следовательно, языковую обработку.

Описанные знания о влиянии психолингвистических характеристик слова на чтения накоплены преимущественно на материале романо-германских языков или крупных представителей иных языковых групп, таких как китайский, финский, турецкий. В этой связи, ограниченные кросс-лингвистические ресурсы на сегодня требуют продолжения исследований на различных языках для проверки универсальности и языкоспецифичности моделей языковой обработки и психолингвистических факторов, влияющих на чтение. Для этого используется корпусный подход: см. Мультиязычный корпус движений глаз при чтении текстов (MECO, Siegelman et al., 2022); Потсдамский корпус предложений (Kliegl et al., 2004); Гентский билингвальный корпус текстов (Sui et al., 2022); Русский корпус предложений (RSC, Laurinavichyute et al., 2019); Детская версия Русского корпуса предложений (ChiRSC, Korneev et al., 2017; Lopukhina et al., 2022b), Русский билингвальный корпус предложений (BiRSC, Parshina et al., 2021) и пр.

Продолжая линию кросс-лингвистических исследований, но смещая её фокус на малые и типологически различные языки, мы собрали данные

движений глаз при чтении предложений в группе 50 взрослых билингвов (44 женщины), носителей русского и адыгейского языков в возрасте от 18 до 60 лет, средний возраст составил 32.7 лет,  $SD = 14.1$ . Средний уровень образования участников составил 15.1 лет,  $SD = 2.1$ , в диапазоне 11–20 лет. Согласно заполненным анкетам LEAP-Q (Marian et al., 2007), участники используют адыгейский язык 58,6% времени в течение дня, а русский – 41,4% времени. На шкале 1-5 участники оценили свои навыки чтения на адыгейском на 4 балла, а на русском – на 4,8.

Материалом для чтения послужили два корпуса (набора) предложений: Русский корпус предложений (RSC, Laurinavichyute et al., 2019) из 144 вхождений и Адыгейский корпус предложений (ASC, от англ. Adyghe Sentence Corpus) из 100 вхождений, составленный по аналогии с русским корпусом для данного исследования. Оба корпуса содержат лингвистическую разметку по частотности слов (по Адыгейскому корпусу текстов от Arkhangelskiy et al., 2018), их длине, частеречной принадлежности и морфемной структуре. Для возможности проведения экспериментального дизайна исследования, Адыгейский корпус включал также целевые слова в 8 условиях по параметрам части речи (существительные и глаголы), длина слова (короткие и длинные), и частотность слова (низкая и высокая). Каждое условие было представлено 8 словами в центре предложения, что в сумме дает 64 целевых слова.

Участники читали два набора предложений с экрана ноутбука, во время чего велась регистрация движений их глаз с помощью видеоокулографа EyeLink Portable Duo (SR Research, Canada), с частотой дискретизации 1000 Hz. Порядок предъявления корпусов чередовался между участниками, между чтением корпусов был перерыв. После части предложений в корпусах следовал вопрос на понимание прочитанного предложения, на который участники отвечали с помощью мыши, выбирая правильный ответ из 2 или трех вариантов ответа.

Первая часть анализа данных в среде R (R Core Team, 2020) включала построение регрессионных моделей (линейных моделей для континуальных

мер и логистических – для вероятностных), в которых изучалось влияние (психо)лингвистических факторов и навыка чтения на глазодвигательные меры при чтении на адыгейском языке. Всего были проанализированы 9 зависимых переменных (длительность первой фиксации FFD, длительность единственной фиксации SFD, сумма длительности фиксаций при первом прочтении GD, общее время прочтения слова TT, вероятность пропуска слова P0, вероятность единственной фиксации на слове P1, вероятность более двух фиксаций P2+, вероятность регрессии от слова RO и на слово RG) и ряд факторов в качестве независимых переменных (длина и частотность слова, частеречная принадлежность, длина и частотность предыдущего и последующего слова, относительная позиция слова в предложении, количество лексических аффиксов, место приземления первой фиксации на слове и уровень навыков чтения на адыгейском). Такие модели были построены для двух наборов данных: для всех слов в Адыгейском корпусе предложений и для целевых слов.

Результаты анализа на всех словах Адыгейского корпуса предложений показали значимое влияние длины и частотности слова на ряд глазодвигательных мер (GD, TT, P1, P2+, RG). Однако значимый эффект частотности слова в анализе на целевых словах сохраняется лишь в общем времени прочтения (TT). Вероятно, для полисинтетического языка более важно смотреть на частотность морфемы или биграммы, а не словоформы, что ранее обсуждалось лишь в теоретических работах по лингвистике (Haspelmath, 2018) о проблеме выделения границ слова в полисинтетических языках, но не было изучено экспериментально.

Более длинные предыдущие слова ускоряли общее время чтения (TT) текущего слова на адыгейском языке, в то время как более длинные последующие слова не показали значимого влияния. Глаголы на адыгейском языке читались значительно медленнее, чем существительные (TT), в то время как другие части речи не отличались по скорости чтения от глагола. Наличие лексических аффиксов в слове увеличивало общее время прочтения слова.

Вторая часть анализа данных в среде R была направлена на изучение чтения адыго-русских билингвов на двух языках (на русском и адыгейском). Абсолютные значения глазодвигательных мер при чтении у адыго-русских билингвов характеризуются более длительными фиксациями и меньшим пропуском слов при чтении на адыгейском, чем на русском. При чтении на адыгейском участники приземляются ближе к началу слова (на первые 31% букв слова), при чтении на русском – ближе к центру слова (на первые 48% букв слова).

Статистический анализ о вкладе различных параметров в чтение на обоих языках и взаимодействии данных факторов с языком чтения включал построение моделей, аналогичных первой части анализа, однако модели были построены на объединённом наборе данных чтения всех слов адыгейского и русского корпусов, где язык чтения был заложен как фактор (и был значимым для 8 из 9 мер). Значимые основные эффекты длины и частотности слова наблюдались в ряде глазодвигательных мер (FFD, SFD, GD, TT, P0, P2+), однако значимое взаимодействие данных параметров с языком было обнаружено лишь в P2+ и RO для частотности и в GD, TT, P1, P2+ для длины. Значимое взаимодействие языка и характеристик соседних слов (частотность и длина) наблюдалось лишь для длины предыдущего слова (TT). Это даёт основание полагать, что адыго-русские билингвы по-разному полагаются на информацию о соседних словах при чтении на русском и адыгейском языках.

Настоящее исследование подтверждает эффекты частотности и длины слова на движения глаз при чтении на полисинтетическом языке, хотя и ставит под сомнение необходимость опоры на частотность слова, а не на частотность биграммы или морфемы. Мы также подтвердили морфологические эффекты при чтении на адыгейском языке (влияние частеречной принадлежности и количества лексических аффиксов), которые ранее были показаны на материале морфологически богатых языков. Кроме того, исследование показало сходства и различия в чтении на синтетическом русском и полисинтетическом адыгейском языке в группе адыго-русских билингвов.

## **Заключение и выводы**

Исследования, представленные в данной диссертации, отвечают на фундаментальные вопросы, связанные со сложным когнитивным процессом чтения. С помощью серии видеоокулографических исследований в различных популяциях (у детей, монолингвальных и билингвальных взрослых) и на материале различных языков исследование расширяет понимание механизмов чтения в менее изученных языках. Полученные результаты демонстрируют значимый вклад возраста и психолингвистических факторов в процесс чтения.

Исследование, описанное в разделе 1, показало, что развитие навыков чтения проходит качественно аналогичные стадии как у типично развивающихся детей, так и у детей с высоким риском дислексии. Однако сохраняется количественные межгрупповые различия в развитии этих навыков. Так, типично читающие дети демонстрируют быстрый переход от начального паттерна чтения к беглому, в то время как детям с риском дислексии для такого перехода требуется на 3 года больше.

Исследование в разделе 2 предоставило экспериментальное подтверждение модели поверхностной языковой обработки и показало, что носители русского языка больше полагаются на семантику, чем на синтаксис при чтении предложений. Тем не менее, наши результаты не подтверждают предположения о том, что внешний шум (аудиальный или визуальный) приведет к большей опоре на семантику. Следует отметить, что воздействие аудиального и визуального шума на чтение различалось: при аудиальном шуме читающие замедляются на первых этапах обработки, но ускоряются впоследствии, в то время как визуальный шум в целом ускорял чтение.

Исследование в разделе 3 описало базовые характеристики движений глаз при чтении на полисинтетическом языке (адыгейском) и описало схожие и различающиеся механизмы чтения у билингвальных носителей, читающих на двух языках (адыгейском и русском). Мы подтвердили широко известные лексические и морфологические эффекты на движения глаз у адыго-русских билингвов для обоих языков. Более того, мы показали, что вклад некоторых

психолингвистических характеристик в чтение отличается в зависимости от языка чтения (русского или адыгейского), что представляет интерес для дальнейшего изучения.

Результаты настоящего диссертационного исследования вносят вклад в понимание универсальных и языкоспецифичных характеристик чтения и демонстрируют важность продолжения исследований в данном направлении на материале менее изученных языков. Теоретическая значимость исследования заключается в расширении фундаментального знания на материале типологически различных языков с менее изученными письменностями и орфографиями. Практическая значимость исследования заключается в возможности применения полученных результатов в логопедии для диагностики дислексии, в образовании, маркетинге и дизайне, а также в возможностях свободно использовать Адыгейский корпус предложений для исследовательских и учебных целей.

## Список литературы

1. Корнев А.Н. Нарушения чтения и письма у детей: учебно-метод. пособие. СПб.: МиМ, 1997.
2. Чернова, Д. А., Слюсарь, Н. А., Прокопеня, В. К., Петрова, Т. Е., & Черниговская, Т. В. (2016). Экспериментальные исследования грамматики: синтаксический анализ неоднозначных предложений. *Вопросы языкознания*, 6, 36-50.
3. Arkhangel'skiy, T., Bagirokova, I., Lander, Y., & Lander, A. (2018). Adyghe corpus. Available at: [http://adyghe.web-corpora.net/index\\_en.html](http://adyghe.web-corpora.net/index_en.html) [Accessed April 26, 2023].
4. Barrington, R., 2019. Understanding dyslexia by measuring eye-movements during reading. [Doctoral Thesis, Bournemouth University].
5. Cauchard, F., Cane, J.E., Weger, U.W. (2012). Influence of Background Speech and Music in Interrupted Reading: An Eye-Tracking Study. *Applied Cognitive Psychology*, 26: 381–390.
6. De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language*, 80(3), 617–626. <https://doi.org/10.1006/brln.2001.2637>
7. De Luca, M., Di Pace, E., Judica, A., Spinelli, D., & Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia*, 37(12), 1407–1420. [https://doi.org/10.1016/s0028-3932\(99\)00038-x](https://doi.org/10.1016/s0028-3932(99)00038-x)
8. Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive psychology*, 47(2), 164-203.
9. Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current directions in psychological science*, 11(1), 11-15.
10. Ferreira, F., & Patson, N. D. (2007). The ‘good enough’ approach to language comprehension. *Language and linguistics compass*, 1(1-2), 71-83.

11. Ferreira, F., & Stacey, J. (2000, April 10). The misinterpretation of passive sentences.  
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2c5a94a3935b4008566f27b15e761579c9dba772> [Preprint]
12. Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, 110(20), 8051-8056.
13. Hatzidaki, A., Gianneli, M., Petrakis, E., Makaronas, N., & Aslanides, I. M. (2010). Reading and visual processing in Greek dyslexic children: An eye-movement study. *Dyslexia*, 17(1), 85–104. <https://doi.org/10.1002/dys.416>
14. Haspelmath, M. (2018). The last word on polysynthesis: A review article. *Linguistic Typology*, 22(2), 307-326. <https://doi.org/10.1515/lingty-2018-0011>
15. Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, 115(3), 367–379. <https://doi.org/10.1016/j.cognition.2009.11.004>
16. Hawelka, S., & Wimmer, H. (2005). Impaired visual processing of multi-element arrays is associated with increased number of eye movements in dyslexic reading. *Vision Research*, 45(7), 855–863. <https://doi.org/10.1016/j.visres.2004.10.007>
17. Hutzler, F., Kronbichler, M., Jacobs, A. M., & Wimmer, H. (2006). Perhaps correlational but not causal: No effect of dyslexic readers' magnocellular system on their eye movements during reading. *Neuropsychologia*, 44, 637–648. <https://doi.org/10.1016/j.neuropsychologia.2005.06.006>
18. Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language*, 89(1), 235–242. [https://doi.org/10.1016/S0093-934X\(03\)00401-2](https://doi.org/10.1016/S0093-934X(03)00401-2)
19. Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of*

- Experimental Psychology: Learning, Memory, and Cognition, 21(6), 1430–1440. <https://doi.org/10.1037//0278-7393.21.6.1430>
20. Hyönä, J., & Eklholm, M. (2016). Background Speech Effects on Sentence Processing during Reading: An Eye Movement Study. *PLOS ONE*, 11: e0152133.
  21. Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & psychophysics*, 40(6), 431-439.
  22. Kemper, S., McDowd, J., Metcalf, K., Liu, C-J. (2008). Young and Older Adults' Reading of Distracters. *Educational Gerontology*, 34: 489–502.
  23. Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European journal of cognitive psychology*, 16(1-2), 262-284.
  24. Korneev, A., Akhutina, T., & Matveeva, E. (2017). Silent reading in Russian junior schoolchildren: An eye tracking study. *Psychology: Journal of Higher School of Economics*, 14(2), 219–235.
  25. Laurinavichyute, A. K., Sekerina, I. A., Alexeeva, S., Bagdasaryan, K., & Kliegl, R. (2019). Russian Sentence Corpus: Benchmark measures of eye movements in reading in Russian. *Behavior Research Methods*, 51(3), 1161–1178. <https://doi.org/10.3758/s13428-018-1051-6>
  26. Lefton, L. A., Nagle, R. J., Johnson, G., & Fisher, D. F. (1979). Eye movement dynamics of good and poor readers: Then and now. *Journal of Reading Behavior*, 11(4), 319–328. <https://doi.org/10.1080/10862967909547338>
  27. Levy, R. (2008, October). A noisy-channel model of human sentence comprehension under uncertain input. In *Proceedings of the 2008 conference on empirical methods in natural language processing* (pp. 234-243), Honolulu, Hawaii. Association for Computational Linguistics.
  28. Levy, R. (2011, June). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results.

- In *Proceedings of the 49th annual meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 1055-1065), Portland, Oregon, USA. Association for Computational Linguistics.
29. Lopukhina, A., Laurinavichyute, A., Malyutina, S., Ryazanskaya, G., Savinova, E., Simdianova, A., ... & Korkina, I. (2022a). Reliance on semantic and structural heuristics in sentence comprehension across the lifespan. *Quarterly Journal of Experimental Psychology*, 75(7), 1367-1381.
  30. Lopukhina, A., Zdorova, N., Staroverova, V., Ladinskaya, N., Kaprielova, A., Goldina, S., ... Dragoy, O. (2022b, August 10). Benchmark measures of eye movements during reading in Russian children. <https://doi.org/10.31234/osf.io/2x5pk> [Preprint]
  31. Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of speech, language, and hearing research: JSLHR*, 50(4), 940–967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
  32. Parshina, O., Laurinavichyute, A. K., & Sekerina, I. A. (2021). Eye-movement benchmarks in Heritage Language reading. *Bilingualism: Language and Cognition*, 24(1), 69-82.
  33. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>
  34. Rayner, K. (1985a). The role of eye movements in learning to read and reading disability. *Remedial and Special Education*, 6, 53–60.
  35. Rayner, K. (1985b). Do faulty eye movements cause dyslexia? *Developmental Neuropsychology*, 1(1), 3–15. <https://doi.org/10.1080/87565648509540294>
  36. Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>

37. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
38. Siegelman, N., Schroeder, S., Acartürk, C., Ahn, H. D., Alexeeva, S., Amenta, S., ... & Kuperman, V. (2022). Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). *Behavior research methods*, 54(6), 2843-2863.
39. Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In: G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 327-342). Oxford, UK: Oxford University Press.
40. Sui, L., Dirix, N., Woumans, E., & Duyck, W. (2023). GECO-CN: Ghent eye-tracking corpus of sentence reading for Chinese-English bilinguals. *Behavior Research Methods*, 55(6), 2743-2763.
41. Vasilev, M.R., Liversedge, S.P., Rowan, D., Kirkby, J.A., Angele, B. (2019). Reading is disrupted by intelligible background speech: Evidence from eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance*, 45: 1484.
42. Yan, G., Meng, Z., Liu, N., He, L., Paterson, K.B. (2018). Effects of irrelevant background speech on eye movements during reading. *Journal of Experimental Psychology*, 71: 1270–1275.