

Лингвистически сбалансированный инструментарий для оценки навыков чтения у взрослых носителей русского языка

Светлана Дорофеева, Яна Кокорева

Статья поступила
в редакцию
в апреле 2024 г.

Дорофеева Светлана Валентиновна — кандидат филологических наук, научный сотрудник Центра языка и мозга, Национальный исследовательский университет «Высшая школа экономики». Адрес: 101000 Москва, Кривоколенный переулок, 3. E-mail: sdorofeeva@hse.ru. ORCID: <https://orcid.org/0000-0002-4299-3069> (контактное лицо для переписки)

Кокорева Яна Александровна — стажер-исследователь Центра языка и мозга, Национальный исследовательский университет «Высшая школа экономики». E-mail: yana_kokoreva28@mail.ru. ORCID: <https://orcid.org/0009-0003-8809-9104>

Аннотация

Нарушения чтения значительно снижают качество жизни, поэтому их необходимо своевременно выявлять, корректировать или создавать особые условия для обучения и работы, если коррекция неэффективна. Оценить наличие нарушений чтения и результаты коррекционного воздействия невозможно без надежных и валидных тестов. При том что русскоязычные тесты для оценки навыков чтения у детей систематически создаются и стандартизируются, инструменты, позволяющие оценить навык чтения у взрослых носителей русского языка, до сих пор отсутствовали.

Проведено исследование с целью создания лингвистически сбалансированного инструментария для оценки навыков чтения текста у взрослых. Разработан тест, включающий две параллельные версии текстов, которые составлены таким образом, что между ними нет значимых различий по ряду психолингвистических параметров. К каждому тексту предлагается набор вопросов, что позволяет оценить как технику чтения, так и успешность понимания прочитанного. Проведена апробация данного инструментария в группе 111 молодых взрослых носителей русского языка. В ходе апробации подтверждена параллельность версий: между текстами не обнаружено значимых различий ни в скорости чтения, ни в числе верных ответов на вопросы. Поскольку для оценки внешней валидности теста не было возможности использовать другой русскоязычный стандартизированный тест на чтение для взрослых ввиду отсутствия такового, проведено сравнение с широко используемой стандартизированной русскоязычной методикой для оценки навыков чтения у детей, при этом использован наиболее сложный вариант текста из данной методики.

В статье приведены описание нового теста, результаты сравнения параллельных версий, результаты оценки внешней валидности, а также рассчитанные контрольные уровни для диагностики нарушений чтения у молодых русскоговорящих взрослых.

Ключевые слова

чтение, тест на чтение, чтение текстов, скорость чтения, понимание прочитанного, взрослые, исследование чтения

Для цитирования Дорофеева С.В., Кокорева Я.А. (2025) Лингвистически сбалансированный инструментарий для оценки навыков чтения у взрослых носителей русского языка. *Вопросы образования / Educational Studies Moscow*, № 2, сс. ??? <https://doi.org/10.17323/vo-2025-21026>

Linguistically Balanced Tool for Reading Skills Assessment in Russian-Speaking Adults

Svetlana Dorofeeva, Jana Kokoreva

Svetlana V. Dorofeeva — PhD in Philology and Linguistics, Research Fellow at the Center for Language and Brain, HSE University. Address: 3 Krivokolenny lane, 101000 Moscow, Russian Federation. E-mail: sdorofeeva@hse.ru. ORCID: <https://orcid.org/0000-0002-4299-3069> (corresponding author)

Jana A. Kokoreva — Research Assistant at the Center for Language and Brain, HSE University. E-mail: yana_kokoreva28@mail.ru. ORCID: <https://orcid.org/0009-0003-8809-9104>

Abstract Reading impairments substantially reduce quality of life, that is why it is necessary to detect them timely, provide intervention programs, or provide special learning or working environment for people with reading impairments in case the intervention was not effective. The diagnostics of reading impairments and evaluation of intervention efficiency is impossible without reliable and valid assessment tools. Whereas Russian-language tools for reading skills assessment in children are being developed and standardized systematically, there still was a lack of tools for reading skills assessment in Russian-speaking adults.

We conducted a study aiming to construct the linguistically balanced tool for reading skills assessment in adults. We developed a test including two parallel versions of texts with no significant differences in several psycholinguistics parameters. Each text is accompanied with a set of questions that allows to assess both reading skills and written text comprehension skills. The test passes an approbation in a group of 111 Russian-speaking young adults. During the approbation, two versions were proved to be parallel: no significant differences were revealed neither in reading speed nor in number of correct answers to the comprehension questions. Since it was impossible to evaluate the external validity of an instrument due to the absence of another standardized reading test for adults, the comparison with the widely used standardized reading test for children was made, using the most difficult variant of that test.

The paper provides the description of a new test, the results of comparison of parallel versions, the results of external validity evaluation, and the computed cut-off levels for diagnosing reading impairments in Russian-speaking young-adults.

Keywords reading, reading assessment test, text reading, reading fluency, reading comprehension, adults, reading research

For citing Dorofeeva S.V., Kokoreva J.A. (2025) Linguistically Balanced Tool for Reading Skills Assessment in Russian-Speaking Adults. *Voprosy obrazovaniya / Educational Studies Moscow*, no 2, pp. ??? (In Russian). <https://doi.org/10.17323/vo-2025-21026>

Современный человек ежедневно сталкивается с необходимостью использовать навык чтения. Люди, у которых этот навык нарушен, могут страдать от дискриминации на работе [Cavalli et al.,

2017], испытывать трудности с учебой в школе или в институте [Копиц, 2002]. Поскольку нарушения чтения значительно снижают качество жизни, людям с такими нарушениями требуется помощь: коррекционная работа или создание особых условий, если коррекционная работа оказалась недостаточно успешной. Ни то ни другое невозможно без точной диагностики, для проведения которой необходимы надежные и валидные стандартизированные инструменты, позволяющие оценить уровень развития навыка чтения у конкретного человека.

Ввиду важности задачи своевременной диагностики нарушений чтения работа по созданию и стандартизации русскоязычных тестов на чтение для детей ведется систематически. Для оценки навыка чтения на русском языке в настоящее время стандартизованы три методики, предназначенные для выявления дислексии: Стандартизированная методика исследования навыка чтения (СМИНЧ) [Корнев, Ишимова, 2010], тест для оценки навыков чтения у младших школьников *LexiMetr* [Здорова и др., 2025], батарея тестов на чтение слов и псевдослов [Дорофеева и др., 2021].

СМИНЧ ориентирована на учащихся 1–6-х классов и предназначена для оценки скорости чтения и уровня понимания прочитанного с целью выявления нарушений чтения. Методика представлена в двух параллельных версиях, и каждая из них включает два текста, различающихся по уровню сложности. Безусловным достоинством СМИНЧ является проведенная стандартизация методики, наличие опубликованных нормативных показателей, в том числе контрольных уровней для диагностики дислексии, а также данных о ее диагностической валидности [Корнев, Ишимова, 2010; Dorofeeva et al., 2019]. Ограничение данной методики состоит в том, что при составлении текстов не учитывались психолингвистические параметры материала для чтения.

Методика *LexiMetr*, разработанная недавно в Центре языка и мозга НИУ ВШЭ, представляет собой более современный инструмент для оценки навыков чтения у детей. Она опирается на те же принципы, на которых основана СМИНЧ, и также дает возможность оценить скорость чтения текстов и уровень понимания прочитанного, однако учитывает психолингвистические параметры при определении уровня сложности текстов и реализована в виде удобного в применении планшетного приложения, позволяющего делать аудиозапись голоса ребенка во время чтения и проводить разметку.

Батарея тестов на чтение слов и псевдослов предназначена для оценки скорости и точности чтения на уровне слова и разработана с особым вниманием к лингвистической составляющей: параллельные версии листов для чтения слов не имеют значимых различий по частотности, длине, слоговой структуре и частеречной принадлежности слов, стимулы на листах для чтения псевдослов сопоставлены по длине и слоговой структуре со стимулами на ли-

стах для чтения слов. Проведенное исследование подтвердило отсутствие значимых различий между параллельными версиями тестов на чтение слов и псевдослов из данной батареи, а также показало, что результаты этих тестов имеют высокий коэффициент корреляции с результатами СМИНЧ [Дорофеева и др., 2021]. Данная батарея тестов пока не может использоваться для диагностики нарушений чтения у взрослых, поскольку нормативные данные собраны только для детей младшего школьного возраста [Там же] и для подростков 14–17 лет [Бурдына и др., 2023]. Однако, в отличие от СМИНЧ и теста *LexiMetr*, материалы тестов на чтение слов и псевдослов подобраны таким образом, что не имеют возрастных ограничений и после сбора нормативных данных могут быть использованы для оценки навыков чтения на уровне слова у взрослых.

В 2023 г. опубликованы нормативные данные для нового теста, предназначенного для оценки навыка чтения с использованием текста у подростков 14–17 лет [Лезина и др., 2023]. Данный инструмент разработан с опорой на опыт использования СМИНЧ для оценки навыков чтения у детей, но, по утверждению авторов, включает лингвистически более сложные тексты.

Все рассмотренные тесты предназначены для детей или подростков, инструментов для диагностики нарушений чтения у взрослых носителей русского языка до сих пор не было. При этом оценка навыков чтения на уровне текста у взрослых является актуальной исследовательской задачей, поскольку механизмы чтения на уровне текста, включающие влияние морфосинтаксической обработки и предсказуемости контекста, отличаются от механизмов чтения на уровне слова. Инструменты оценки чтения необходимы как для исследований механизмов чтения у взрослых, в частности для отбора в исследования взрослых участников с типичным или нарушенным чтением, а также для исследования влияния различных факторов на скорость или точность чтения, так и для диагностики нарушений чтения. Лингвистически сбалансированный, стандартизированный инструмент был бы полезен как исследователям, так и специалистам-практикам — логопедам, нейропсихологам, дефектологам.

Получить возможность подтвердить или опровергнуть наличие дислексии может быть особенно важно для молодых взрослых — студентов высших учебных заведений, для которых чтение является необходимым инструментом обучения. У некоторых студентов трудности с чтением, даже если они не помешали поступить в вуз, в ситуации длительной повышенной когнитивной нагрузки становятся ограничением, препятствующим успешной учебе в вузе. Ряд исследователей отмечают, что численность студентов высших учебных заведений, которые сталкиваются с трудностями в чтении, неуклонно растет [Tops et al., 2022; Ottosen et al., 2022]. Все чаще требуется адаптировать учебные програм-

мы и экзамены под нужды людей с трудностями чтения [Konur, 2002; Tops et al., 2022; Desrochers, Glickman, 2009; Phillips, Kelly, Symes, 2013]. Для выявления студентов с трудностями чтения и создания им особых условий учебные заведения нуждаются в валидных и надежных инструментах для оценки этого навыка.

Первая цель настоящего исследования состоит в разработке лингвистически сбалансированного инструментария для оценки навыков чтения текстов и понимания прочитанного у взрослых носителей русского языка и его апробации, вторая — в сборе нормативных данных в группе молодых взрослых носителей русского языка и расчете контрольных диагностических уровней для этой группы. Учитывая ограниченность ресурсов, мы приняли решение в первую очередь собрать нормативные данные для группы молодых взрослых, поскольку для ее представителей, студентов, отсутствие возможности своевременно диагностировать дислексию может иметь особенно тяжелые последствия, например снижение успеваемости вплоть до невозможности продолжить обучение, что окажет негативное влияние на всю последующую жизнь. Безусловно, сбор нормативных данных для оценки навыков чтения текста в других возрастных группах взрослых носителей русского языка также важен и может стать задачей дальнейших исследований.

1. Теоретический обзор

1.1. Определение навыка чтения и выделение параметров, подлежащих оценке

Чтение — один из наиболее сложных когнитивных процессов, которые осваивает человек [Nicolson, Fawcett, 2008]. Чтение можно определить как комбинацию процессов декодирования и понимания письменного текста [Cline, Johnstone, King, 2006]. Декодирование подразумевает соотнесение письменной формы слова с его фонетическим обликом и далее с его значением. Понимание текста включает интерпретацию семантики, морфосинтаксическую обработку словоформ, интерпретацию синтаксических структур предложений, а также уяснение особенностей дискурса [Hoover, Gough, 1990; Driscoll, 2012].

Инструменты для оценки могут фокусироваться на разных аспектах навыка: на чтении отдельных слов/псевдослов или на чтении более длинных единиц — предложений и текстов; на понимании прочитанного текста или только на скорости и/или точности чтения; на оценке способности читать вслух или про себя.

Много инструментов создано для оценки умения быстро распознавать отдельные слова или декодировать псевдослова [Keenan, Betjemann, Olson, 2008], так как успешность чтения напрямую зависит от способности человека соотносить письменную форму слова с его фонетической формой и значением [Catts, Kamhi, 1999]. Этой цели служат тесты на чтение слов и псевдослов, которые позволяют оценить такие показатели, как скорость и точность чтения, и косвенно — фонологическую осознанность, т.е. уме-

ние воспринимать звучащее слово как последовательность звуков речи [Spector, 1992; Catts, Kamhi, 1999]. Однако подобные тесты не позволяют оценить навык чтения текста [Share, 2021]: они проверяют чтение на уровне слов вне контекста, тогда как для чтения текста важно умение оценивать контекст [Catts, Kamhi, 1999]. Кроме того, для чтения текста необходимо не только умение читать отдельные слова или быстро декодировать текст, но и умение проводить морфологическую и синтаксическую обработку, понимать содержание текста. Чтение более длинных, чем отдельные слова, единиц часто оказывается более показательным для оценки скорости чтения [Trauzettel-Klosinski, Dietz, 2012; Altpeter et al., 2015], особенно у взрослых.

В данном исследовании мы будем исходить из определения чтения как комбинации процессов декодирования и понимания текста, поэтому при создании инструментария для оценки навыка чтения мы сфокусируемся на разработке лингвистически сбалансированных текстов и вопросов к ним, чтобы иметь возможность оценивать такие параметры, как скорость и точность чтения текста, а также уровень понимания прочитанного.

1.2. Требования к структуре инструментария для тестирования навыков чтения

1.2.1. Соответствие материала задачам тестирования

1.2.1.1. Выбор типа текста

Тип текста может влиять на скорость чтения. Эмпирически установлено, что школьники читали повествовательный текст значительно быстрее, чем информационный [Uysal, Bilge, 2019]. Кроме того, тип текста может влиять и на его понимание: так, ученики начальной школы в Турции понимали повествовательный текст, содержащий описания событий, более успешно, чем информационный, целью которого является передача и актуализация информации [Şahin, 2013]. У взрослых читателей уровень понимания текста также различается в зависимости от жанра: понимание повествовательных текстов происходит успешнее по сравнению с описательными текстами и текстами-инструкциями [Harris, Rogers, Qualls, 1998]. При этом молодым взрослым легче воспринимать динамические тексты, чем статические [Доброго, Петрова, 2016].

1.2.1.2. Адекватность возрасту по сложности

Текст, предложенный для оценки навыка чтения, должен соответствовать возрасту читающего по сложности. Одна из метрик, часто используемых, для того чтобы оценить, насколько текст соответствует возрасту читающего, — возраст усвоения входящих в текст слов. Однако она применима в основном для оценки текстов, предназначенных для детей либо для людей, изучающих второй язык. Другая метрика, формула Флеша — Кинкейда [Kincaid et al., 1975], позволяет оценить уровень образования, необходимый для успешного прочтения и понимания текста. Формула Флеша — Кинкейда

является модифицированной версией формулы Флеша [Flesch, 1948]. Обе эти формулы основаны на таких параметрах текста, как средняя длина предложений в словах и средняя длина слов в символах. По формуле Флеша сложность текста оценивается в баллах от 0 до 100: чем ниже оценка, тем сложнее текст для восприятия. Формула Флеша — Кинкейда оценивает количество лет формального образования, которое необходимо, чтобы понять текст.

Поскольку наш тест предназначен для диагностики дислексии, мы не ограничивались задачей разработать тексты, соответствующие целевой аудитории по возрасту. Мы стремились к тому, чтобы тексты, с одной стороны, субъективно воспринимались как достаточно простые и не делали процедуру тестирования мучительной для участников с дислексией, а с другой — давали достаточную когнитивную нагрузку, повышающую вероятность выявления трудностей в чтении у взрослых при их наличии. Мы выбрали повествовательные тексты (по данным предыдущих исследований, они легче других для восприятия) в формате новостной статьи, хорошо знакомом взрослым читателям. Для создания достаточной когнитивной нагрузки мы включили в текст предложения с относительно редко встречающимися синтаксическими конструкциями с обратным порядком слов. Этот тип предложений допустим в русском языке, но не является преобладающим [Slioussar, Makarchuk, 2022]. Эмпирически показано, что слова в предложениях с непревалирующим порядком слов обрабатываются типичными носителями русского языка почти так же успешно, как и слова в предложениях с обычным порядком слов (с эффектом непревалирующего порядка слов небольшого размера или с отсутствием эффекта в зависимости от условий), но при этом непревалирующий порядок слов вызывает больше трудностей при обработке предложения у людей, осваивающих русский язык как второй [Slioussar, Harchevnik, 2024]. Мы ожидаем, что на основании подобного эффекта окажется возможным различать носителей русского языка с успешно сформированным навыком чтения и с нарушениями чтения. У людей с дислексией во время чтения синтаксически сложных предложений трудности обработки могут вызвать паузы, что приведет к снижению скорости чтения, или дополнительные ошибки, что отразится на показателях точности чтения, или усложнит понимание прочитанного, что скажется на количестве верных ответов на вопросы. Проверить эту гипотезу мы сможем только в будущем, после проведения исследования с участием взрослых с дислексией.

1.2.1.3. Наличие
вопросов
для оценки
понимания

Понимание прочитанного оценивают с помощью вопросов по содержанию прочитанного текста, пересказа текста или наборов ключевых слов [Мурзин, Штерн, 1991; Magliano et al., 2007; McArthur et al., 2013]. Оценка понимания прочитанного с помощью клю-

чевых слов более трудозатратна для проводящего тестирование, чем основанная на вопросах по содержанию текста. Если тест создается для использования широким кругом специалистов, целесообразно предусмотреть в нем оценивание понимания прочитанного именно при помощи вопросов. Послетекстовые вопросы как способ тестирования имеют свои недостатки: они отражают не процесс понимания, как он идет в действительности непосредственно во время чтения, а сохранность понятой во время чтения текста информации после прочтения всего фрагмента [Magliano et al., 2007]. Однако ответы на послетекстовые вопросы в большей степени, чем пересказ текста, поддаются объективному оцениванию [Корнев, 1997; Корнев, Ишимова, 2010]. Кроме того, оценка понимания при помощи пересказа занимает больше времени, а также ставит под вопрос адекватность оценки навыка чтения у людей, испытывающих сложности с экспрессивной речью [Klingner, 2004]. Поэтому в большинстве как отечественных, так и зарубежных тестов для оценки навыка чтения [Корнев, 1997; Корнев, Ишимова, 2010; Лезина, 2023; Wiederholt, Bryant, 2001] понимание прочитанного проверяется именно с помощью послетекстовых вопросов. Таким образом, чтобы обеспечить возможность сопоставления результатов наших тестов с показателями, полученными с применением других методик, мы решили оценивать понимание прочитанного текста посредством послетекстовых вопросов.

1.2.2. Наличие параллельных версий

Для исследования развития навыков чтения в динамике, и особенно для проверки успешности коррекционной работы в случаях нарушений чтения, необходимо проводить повторные тестирования одного и того же участника. При использовании одной и той же версии теста более высокий результат при повторном тестировании может быть обусловлен тем, что тестируемый уже знаком с предложенным тестом. Следовательно, тестовые тексты должны иметь несколько параллельных версий — как минимум две.

Параллельность текстов во многом обеспечивается их лингвистической составляющей. Например, в *Test of Everyday Reading Comprehension* [McArthur et al., 2013] две версии составлены таким образом, чтобы входящие в них слова как можно лучше соответствовали друг другу по частотности. В результате такого подбора слов для разных версий теста удалось добиться того, что количество верных ответов участников на вопросы к этим тестам значимо не различалось. Авторы *International Reading Speed Texts* [Trauzettel-Klosinski, Dietz, 2012] преимущественное внимание при создании параллельных версий текста уделяли длине слов. В этом случае версии теста читались с одинаковой скоростью. Однако и на скорость чтения, и на его понимание могут влиять и другие параметры, помимо частотности и длины слов.

В частности, важной оказывается синтаксическая структура предложений [Poulsen, Gravgaard, 2016; Magliano et al., 2007; Лапошина и др., 2022]. Например, показатели понимания предложений сложной структуры, таких как предложения в пассивном залоге, у школьников оказались ниже, чем показатели понимания предложений синтаксически более простых [Poulsen, Gravgaard, 2016].

**1.2.3. Наличие
норматив-
ных данных
и определение
контрольных
диагностических
уровней**

Существует несколько способов рассчитывать контрольные диагностические уровни. Один из наиболее распространенных в клинической практике способов — 95%-ный доверительный интервал от среднего. При таком подходе принимается, что результаты, выходящие за рамки интервала «среднее $\pm$ 2 стандартных отклонения» при распределении, соответствующем нормальному, будут встречаться менее чем в 5% случаев. Таким образом, уровень «среднее минус 2 стандартных отклонения» может считаться срезовым диагностическим уровнем [Sharma, Jain, 2014]. В случаях, когда дисперсия велика, в некоторых тестах используют в качестве отсечки уровень «среднее минус 1,5 стандартного отклонения». Расчет контрольных диагностических уровней на основе среднего и стандартных отклонений используется во многих диагностических инструментах, в частности в СМИНЧ [Корнев, Ишимова, 2010], а также в недавно разработанных и апробированных инструментах [Бурдына и др., 2023; Лезина и др., 2023].

Еще один достаточно распространенный способ расчета контрольных диагностических уровней — определение их на основе процентилей. В таком случае за показатели, соответствующие уровню тяжелого нарушения, уровню нарушения и уровню риска развития нарушений, могут принимаются цифры, ниже которых оказались результаты по данному тесту соответственно у 5, 7 и 10% участников нормативной выборки. Проценти́ли для расчета контрольных уровней используются, например, в Русском афазиологическом тесте [Ivanova et al., 2021] и в *Hasbrouck-Tindal Oral Reading Fluency Chart* [Hasbrouck, Tindal, 1992; 2006].

Выбор предпочтительного варианта расчета контрольных уровней может определяться уже используемым в отрасли подходом (например, [Бурдына и др., 2023]), но зависит и от распределения данных в выборке нормирования: при нормальном распределении данных уместно использовать подход с опорой на средние значения и стандартные отклонения, при распределении данных, отличающемся от нормального, предпочтительным вариантом может быть подход на основе процентилей.

Что касается размера выборки для нормирования, он также зависит от типа теста. Если тестирование проводится в группе и оценка также ведется на групповом уровне, нормативная выборка должна быть не менее 1 тыс. человек. Однако если тест проводится

индивидуально, часто достаточно бывает выборки в 100–200 человек [Weiner, Graham, Naglieri, 2012]. При этом в современной практике диагностические уровни рассчитываются и на меньшей выборке: в недавних работах [Бурдына и др., 2023; Лезина и др., 2023] приводятся уровни, рассчитанные на выборках по 103 человека.

Наша работа по созданию теста для оценки навыков чтения у взрослых состояла из нескольких этапов. Во-первых, мы составили два текста, эквивалентных друг другу по психолингвистическим параметрам входящих в них слов и по наборам используемых синтаксических структур, после чего разработали наборы вопросов на понимание к каждому из текстов, также параллельные по структуре. Эквивалентность сложности текстов позже была подтверждена анализом с использованием инструмента «Текстометр» [Лапошина, Лебедева, 2021]. Во-вторых, мы провели эксперимент с участием 111 молодых взрослых носителей русского языка, в котором проверили, действительно ли благодаря высокой лингвистической сбалансированности версий теста между двумя составленными текстами нет различий ни в скорости чтения, ни в точности чтения, ни в правильности ответов на вопросы. В-третьих, мы использовали собранные в ходе эксперимента нормативные данные для расчета контрольных диагностических уровней в группе молодых взрослых, что позволит использовать разработанный инструментарий в диагностических и исследовательских целях.

2. Разработка лингвистически сбалансированного инструмента для оценки навыков чтения у взрослых носителей русского языка

Разработанный нами тест на чтение текста состоит из двух параллельных версий. Параллельность версий обеспечивалась тщательным подбором материала по ряду психолингвистических параметров.

Каждая версия теста включает текст и набор из 11 открытых вопросов. Чтобы тексты были привычными по формату для взрослых читателей, мы не стали использовать художественные тексты и остановили свой выбор на формате новостной статьи. Оба текста являются повествовательными [Валгина, 2003; Хамаганова, 2002], так как содержат рассказ о последовательности событий, и публицистическими [Филиппов, 2003]. Первый текст («Пейзаж») посвящен украденной из музея картине, второй («Корабль») — терпящему бедствие кораблю.

2.1. Сбалансированность материалов теста на уровне лексических характеристик

Многочисленными исследованиями показано, что на скорость чтения могут влиять различные психолингвистические параметры, в частности длина слов и их частотность [Kliegl et al., 2004; Laurinavichyute et al., 2019; Лапошина, Лебедева, Берлин Хенис, 2022]. Учитывая эти данные, мы составили два параллельных текста для на-

шего теста таким образом, чтобы между ними не было значимых различий ни в длине слов, оцениваемой в буквах, ни в частотности слов, ни в числе слогов в полнозначительных словах.

Под полнозначительными словами здесь мы подразумеваем все слова, кроме союзов, предлогов и частиц. Поскольку союзы, предлоги и частицы — это слова преимущественно относительно короткие и высокочастотные, взрослые здоровые люди обычно не фиксируются на них при чтении и, следовательно, эти слова почти не влияют на скорость чтения [Laurinavichyute et al., 2019]. При этом союзы, предлоги и частицы имеют довольно большой разброс частотности: например, частотность предлога «об» составляет 649,61 IPM¹, а предлога «на» — 15 406,88 IPM. Таким образом, не давая почти никакой информации о сложности текста для чтения, союзы, предлоги и частицы могут значительно исказить результаты сравнения частотности слов в целом. Поэтому далее при сравнении параметров текстов мы в расчетах используем только полнозначительные слова. В тексте «Пейзаж» встретились 153 полнозначительных слова, а в тексте «Корабль» — 148.

Распределения значений по всем психолингвистическим параметрам использованных в текстах полнозначительных слов отличаются от нормального: при применении теста Шапиро — Уилка $p < 0,05$ ($W = 0,41$, $p < 0,001$ и $W = 0,44$, $p < 0,001$ для распределения частотности слов в текстах «Пейзаж» и «Корабль» соответственно; $W = 0,94$, $p < 0,001$ и $W = 0,95$, $p < 0,001$ для длины слов в буквах; $W = 0,89$, $p < 0,001$ и $W = 0,91$, $p < 0,001$ для длины слов в слогах). Поэтому попарное сравнение текстов проводилось при помощи критерия Манна — Уитни. Анализ показал отсутствие значимых различий между характеристиками текстов (табл. 1, рис. 1). Кроме того, при сравнении характеристик всех слов текстов, включая служебные части речи, такие как союзы, частицы и предлоги, при помощи критерия Манна — Уитни также не обнаружено значимых различий между текстами (рис. 2).

Таблица 1. Психолингвистические характеристики текстов (среднее $\pm$ стандартное отклонение) и результаты сравнения текстов с использованием критерия Манна — Уитни

	«Пейзаж»	«Корабль»	W	p -value
Количество полнозначительных слов в текстах	153	148	—	—
Частотность слов (IPM)	83 $\pm$ 207	91 $\pm$ 221	12 258	0,215
Длина слов в буквах	7 $\pm$ 3	7 $\pm$ 2	11 785	0,536
Длина слов в слогах	3 $\pm$ 1	3 $\pm$ 1	12 006	0,344

Примечание. Психолингвистические параметры слов определены при помощи базы данных *StimulStat* (дата обращения 04.02.2023) [Alexeeva, Slioussar, Chernova, 2018].

¹ Instances per million — количество вхождений на миллион словоформ.

Рис. 1. Сравнение полнозначительных слов, входящих в тексты, по частотности, длине в буквах и числу слогов при помощи критерия Манна — Уитни

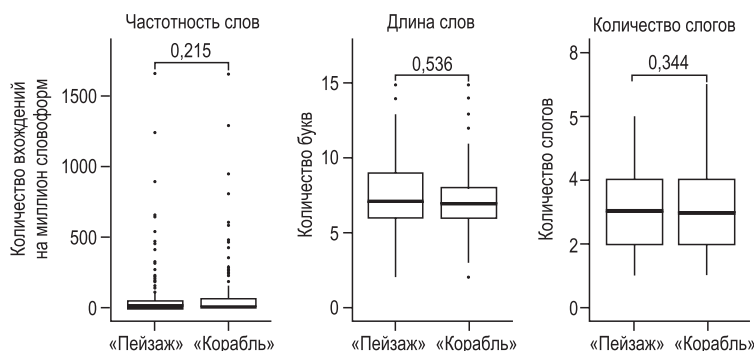
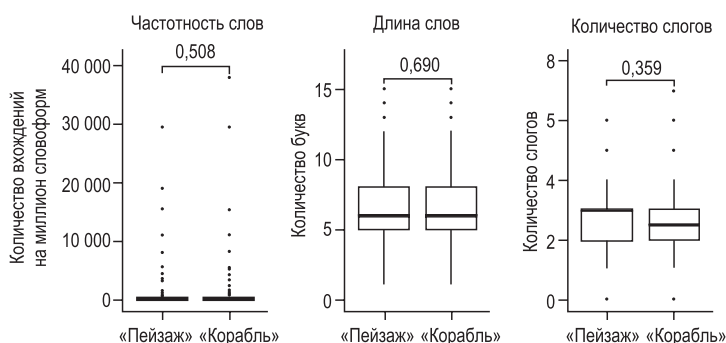


Рис. 2. Сравнение всех слов, входящих в тексты, по частотности, длине в буквах и числу слогов при помощи критерия Манна — Уитни



2.2. Сбалансированность материалов теста на уровне синтаксических структур

Каждый текст состоит из 18 предложений. Шестнадцать первых предложений каждого текста разделены на четыре группы по четыре предложения в зависимости от их синтаксической структуры. Два последних предложения добавлены с целью увеличить длину текстов, чтобы даже быстро читающие участники читали более одной минуты. При этом учитывались ключевые члены предложения: подлежащее, сказуемое и прямое дополнение, а также деепричастные обороты и порядок клауз сложного предложения. Наполненность предложений словами, играющими роли других второстепенных членов, могла варьировать. Предложения разного типа распределены по текстам в псевдослучайном порядке так, чтобы тексты воспринимались как естественные. Порядок предложений с определенной структурой в текстах «Пейзаж» и «Корабль» совпадает. В тексты включены предложения следующей структуры:

- простые предложения;
- сложноподчиненные предложения, начинающиеся с независимой клаузы;

- сложноподчиненные предложения, начинающиеся с зависимой клаузы;
- предложения с деепричастными оборотами.

Каждая группа предложений также разделена на две подгруппы по два предложения в каждой в зависимости от синтаксической сложности предложений. Например, группа простых предложений состоит из предложений с порядком слов «подлежащее — сказуемое — дополнение» (более простая подгруппа) и предложений с порядком слов «подлежащее — дополнение — сказуемое» (более сложная подгруппа). В табл. 2 представлены структуры составленных предложений и приведены примеры, а в табл. 3 приведено количество полнозначенных слов в предложениях каждой группы в каждом из текстов.

Таблица 2. Синтаксические структуры составленных предложений и примеры предложений с этими структурами

Группа	«Простой» вариант (a)	«Сложный» вариант (b)
1	Простое предложение с порядком слов SVO <i>Директор сразу же вызвал полицию на место преступления</i>	Простое предложение с порядком слов SOV <i>Доблестная полиция пропажу искала целых четыре недели</i>
2	Сложное предложение структуры SV, что SVO <i>Поздно вечером дежурные по кораблю обнаружили, что судно получило большую пробоину в металлической обшивке днища ниже уровня погружения</i>	Сложное предложение структуры SV, что SOV <i>Созвонившись с экипажем судна спасателям капитан заявил, что команда прибытия их вертолета давно ожидает</i>
3	Сложное предложение структуры подчинительный союз SVO, SVO <i>Когда капитан заметил вертолет, команда эвакуировала последнего пассажира с борта корабля</i>	Сложное предложение структуры подчинительный союз SOV, SVO <i>Как только моряки свой корабль оставили, судно опустилось на дно моря</i>
4	Простое предложение, начинающееся с деепричастного оборота <i>Тщательно осмотрев залы музея, полицейские не нашли ничего странного</i>	Сложное предложение, начинающееся с деепричастного оборота <i>Отвечая следователю, сторож сообщил, что ему об украденной картине не известно</i>

Примечание. S — подлежащее, V — сказуемое, O — дополнение.

Тексты имеют одинаковую синтаксическую структуру, и одинаковые группы предложений в текстах не различаются между собой по частотности и длине слов в символах и слогах, что также подтверждается результатом сравнения при помощи критерия Манна — Уитни. Размах частотности слов в некоторых группах предложений кажется большим, но в действительности он довольно скромный, а набор использованных слов относительно однороден. Слова с очень высокой частотностью, такие как

Таблица 3. Количество полнозначительных слов в группах предложений

Группа предложений	«Пейзаж»	«Корабль»
1a	11	13
1b	13	13
2a	29	26
2b	25	23
3a	19	18
3b	22	18
4a	15	14
4b	19	23

«весь» с частотностью 1105,14 IPM, встречаются на протяжении текста единично. Более подробно сравнение текстов отражено в табл. 4, расчеты выполнены только для полнозначительных слов (об использовании в расчетах только полнозначительных слов см. комментарий в разделе 2.1).

Таблица 4. Психолингвистические характеристики текстов (минимум — максимум (медиана; среднее $\pm$ стандартное отклонение)) и результаты сравнения текстов с использованием критерия Манна — Уитни

		«Пейзаж» ($n = 153$)	«Корабль» ($n = 148$)	W	p -value
Частотность слов (вхождений на миллион словоформ)	Все предложения	0,17 – 1653,89 (20,6; 83,43 $\pm$ 207,06)	0,17 – 1653,89 (11,1; 91,01 $\pm$ 220,58)	12 258	0,215
	1a	0,51 – 477,47 (25,5; 105,32 $\pm$ 157,85)	0,17 – 477,47 (23,6; 87,61 $\pm$ 138,42)	76,5	0,794
	1b	0,51 – 179,62 (10; 42,63 $\pm$ 60,71)	0,17 – 242,56 (7,8; 49,50 $\pm$ 76,11)	90,5	0,778
	2a	0,17 – 632,94 (25,3; 67,92 $\pm$ 124,15)	0,17 – 1283,75 (7,5; 94,95 $\pm$ 270,41)	449,5	0,225
	2b	0,17 – 1242,93 (10,9; 80,72 $\pm$ 251,27)	0,17 – 945,25 (12,2; 86,39 $\pm$ 204,19)	283,5	0,942
	3a	0,85 – 455,7 (25; 81,36 $\pm$ 138,71)	0,17 – 809,34 (21,7; 79,44 $\pm$ 186,86)	182	0,753
	3b	1,02 – 535,65 (15,3; 56,07 $\pm$ 117,69)	1,02 – 425,76 (9,3; 58,65 $\pm$ 122,50)	222,5	0,514
	4a	2,72 – 894,05 (28,2; 123,36 $\pm$ 268,26)	1,36 – 351,59 (11,5; 80,98 $\pm$ 123,58)	111	0,81
	4b	0,85 – 1653,89 (13,8; 128,16 $\pm$ 374,09)	0,34 – 1653,89 (10,2; 157,05 $\pm$ 363,94)	248,5	0,456
Длина слов в буквах	Все предложения	2 – 15 (7; 7,37 $\pm$ 2,55)	2 – 15 (7; 7,09 $\pm$ 2,38)	11 785	0,536
	1a	5 – 12 (6; 6,91 $\pm$ 2,07)	4 – 11 (6; 6,85 $\pm$ 1,95)	70,5	0,976
	1b	5 – 14 (6; 7,85 $\pm$ 3,16)	4 – 15 (7; 8,46 $\pm$ 3,62)	78	0,755
	2a	4 – 14 (7; 7,48 $\pm$ 2,37)	3 – 13 (7; 6,92 $\pm$ 2,12)	421,5	0,452
	2b	4 – 15 (8; 8,08 $\pm$ 2,81)	2 – 13 (7; 7,22 $\pm$ 2,45)	335,5	0,322

Окончание табл. 4

		«Пейзаж» (n = 153)	«Корабль» (n = 148)	W	p-value
	3a	2 – 12 (8; 7,63 ± 2,75)	4 – 12 (7; 7,06 ± 2,18)	200	0,382
	3b	4 – 12 (6; 6,82 ± 2,52)	3 – 13 (7; 6,94 ± 2,58)	184,5	0,721
	4a	4 – 11 (7; 6,73 ± 2,02)	4 – 12 (7; 7,43 ± 2,21)	86,5	0,427
	4b	3 – 14 (7; 7,05 ± 2,57)	2 – 11 (7; 6,43 ± 2,02)	240,5	0,577
Длина слов в слогах	Все предложения	1 – 6 (3; 3,07 ± 1,04)	1 – 7 (3; 2,95 ± 1,13)	12 006	0,344
	1a	2 – 5 (3; 3,09 ± 1,04)	1 – 5 (2; 2,92 ± 1,32)	79	0,67
	1b	2 – 5 (3; 3,23 ± 1,09)	1 – 6 (3; 3,46 ± 1,56)	78,5	0,772
	2a	2 – 6 (3; 3,14 ± 1,03)	1 – 5 (3; 3,08 ± 1,06)	381	0,95
	2b	2 – 5 (3; 3,32 ± 0,95)	1 – 5 (3; 2,96 ± 1,02)	342	0,245
	3a	1 – 5 (3; 3,21 ± 1,23)	2 – 7 (3; 3,06 ± 1,26)	196,5	0,431
	3b	2 – 5 (2; 2,82 ± 1,10)	1 – 5 (3; 2,78 ± 1,11)	195	0,943
	4a	2 – 5 (3; 2,80 ± 0,77)	1 – 5 (3; 3,00 ± 1,04)	89	0,46
	4b	1 – 5 (3; 2,84 ± 1,07)	1 – 4 (3; 2,57 ± 0,84)	244	0,503

Примечание. Психолингвистические параметры слов определены при помощи базы данных *StimulStat* (дата обращения 04.02.2023) [Alexeeva, Slioussar, Chernova, 2018].

Использование параллельных синтаксических структур преследовало две цели. Во-первых, поскольку разница в синтаксической структуре двух текстов может влиять на скорость и точность чтения, параллельный синтаксис необходим для уменьшения разницы в этих метриках. Во-вторых, разные синтаксические структуры могут по-разному влиять на понимание текста, и, таким образом, параллельность необходима, чтобы снизить разницу в качестве ответов на вопросы после первого и второго текстов [Magliano et al., 2007].

Каждый текст сопровождается набором из 11 вопросов, позволяющих оценить понимание прочитанного. Вопросы составлены так, чтобы между первым и вторым набором было как можно меньше различий: ответами на них были слова, являющиеся одним и тем же членом предложения в исходном тексте. Например, если к первому тексту составлен вопрос, относящийся к подлежащему предложения структуры SVO, то аналогичный вопрос составляли к подлежащему предложения второго текста, обладающего такой же структурой. Вопросы одного набора различались между собой по тому, к какому типу предложения и к какой части предложения они заданы. Примеры вопросов приведены в табл. 5.

Таблица 5. Примеры вопросов на понимание

Структура предложения	Предложение текста «Пейзаж»	Вопрос к тексту «Пейзаж»	Предложение текста «Корабль»	Вопрос к тексту «Корабль»
Простое предложение с порядком слов SVO	Директор сразу же вызвал полицию на место преступления	Кто вызвал полицию на место преступления?	Капитан сразу же собрал всех членов экипажа на совещание	Кто собрал членов экипажа на совещание?

Окончание табл. 5

Структура предложения	Предложение текста «Пейзаж»	Вопрос к тексту «Пейзаж»	Предложение текста «Корабль»	Вопрос к тексту «Корабль»
Простое предложение с порядком слов SOV	Доблестная полиция пропажу искала целых четыре недели	Сколько недель полиция искала пропажу?	Прибывшие спасатели экипаж доставили на противоположный берег лишь через пять часов	Через сколько часов спасатели доставили команду на берег?
Сложное предложение структуры «подчинительный союз SVO, SVO»	Чтобы раскрыть преступление, сотрудники полиции стали спрашивать работников музея о картине	Кто стал спрашивать работников музея о картине?	Ради того, чтобы помочь пассажирам, члены экипажа начали спускать лодки на воду	Кто стал спускать лодки на воду?
Сложное предложение структуры «подчинительный союз SOV, SVO»	Когда ранним утром охранник про происшествие узнал, прочие сотрудники музея вовсе искали ценный пейзаж	Про что охранник узнал утром?	Когда ответственные матросы эту странную просьбу отклонили, сердитый пассажир громко выругался	Что отклонили матросы?

2.3. Оценка сложности разработанных текстов внешним инструментом

Для того чтобы проверить, имеют ли тексты одинаковый уровень сложности, мы сравнили их между собой при помощи инструмента «Текстомер» [Лапошина, Лебедева, 2021]. Анализ показал, что тексты не различаются по сложности. Кроме того, наши тексты по уровню сложности действительно подходят для взрослых людей: формула Флеша — Кинкейда, показывающая количество лет школьного образования, которое необходимо для успешного понимания текста, соответствует полному школьному образованию — 11 годам обучения (табл. 6).

Таблица 6. Сложность текста, оцененная с помощью инструмента «Текстомер» (дата обращения — 04.02.2023) [Лапошина, Лебедева, 2021]

Сложность текста	«Пейзаж»	«Корабль»
Общий балл (из 100)	54	54
Структурная сложность (из 10)	7	7
Лексическая сложность (из 10)	5	5
Динамичность текста (из 10)	6	6
Описательность текста (из 10)	4	3
Формула Флеша (из 100)	38	40
Формула Флеша — Кинкейда (из 11)	11	11
Лексическая плотность (из 10)	8	8
Лексическое разнообразие	0,69	0,71

3. Апробация и сбор нормативных данных

Одна из важных задач исследования состояла в сборе нормативных данных. Однако инструмента, который бы позволил надежно оценить уровень навыка чтения у взрослых носителей русского языка, чтобы исключить из нашей базы нормативных данных показатели тех участников, чей уровень мог оказаться значительно

3.1. Участники

Уровень невербального интеллекта всех протестированных участников находится в пределах возрастной нормы для теста Равена. В контрольном тесте на чтение с использованием текста «Неблагодарная ель» из СМИНЧ [Корнев, Ишимова, 2010] все участники показали результат намного выше уровня 6-го класса (нормы для более старшего возраста отсутствуют, нижняя граница нормы для 6-го класса составляет 84 слова в минуту, минимальная скорость чтения этого текста в нашей выборке составляла 108 слов в минуту). Поскольку как основной статистический анализ, так и расчет контрольных уровней проведен на данных выборки из 96 человек (о причинах исключения 15 участников см. раздел 3.2), в табл. 7 представлены демографические данные и показатели предварительного тестирования не всех 111 участников, а только 96 из них, включенных в основную выборку.

Возраст (лет)	Число участников			Количество лет формального образования (минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))	Уровень невербального интеллекта (минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))	Скорость чтения текста «Неблагодарная ель» (количество верно прочитанных слов за первую минуту чтения; минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))
	Девушки	Юноши	Всего			
18	11	7	18	11 – 13 (12; 11,61 $\pm$ 0,61)	110 – 141 (127; 126,76 $\pm$ 11,08)	125 – 208 (152; 154,50 $\pm$ 22,50)
19	21	18	39	11 – 13 (12; 12,05 $\pm$ 0,60)	92 – 141 (125; 122,66 $\pm$ 14,85)	108 – 222 (160; 160,21 $\pm$ 25,11)
20	15	5	20	12 – 14 (13; 12,90 $\pm$ 0,55)	103 – 141 (128; 125,88 $\pm$ 13,32)	121 – 220 (165,5; 164,80 $\pm$ 28,87)

Вопросы образования/Educational Studies Moscow. 2025. № 2

Возраст (лет)	Число участников			Количество лет формального образования (минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))	Уровень невербального интеллекта (минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))	Скорость чтения текста «Неблагодарная ель» (количество верно прочитанных слов за первую минуту чтения; минимум – максимум (медиана; среднее $\pm$ стандартное отклонение))
	Девушки	Юноши	Всего			
21	8	2	10	13 – 15 (14; 14,00 $\pm$ 0,47)	108 – 141 (127,5; 125,00 $\pm$ 15,11)	142 – 219 (163,5; 174,80 $\pm$ 27,23)
22	6	3	9	12 – 15 (15; 14,22 $\pm$ 1,20)	108 – 130 (121; 121,71 $\pm$ 7,72)	146 – 211 (173; 179,00 $\pm$ 22,19)
18–22	61	35	96	11 – 15 (12; 12,55 $\pm$ 1,09)	92 – 141 (125; 124,35 $\pm$ 13,17)	108 – 222 (160; 163,38 $\pm$ 26,03)

3.2. Процедура и преобработка данных

Апробация разработанного теста проходила в рамках более крупного проекта, посвященного исследованию чтения у молодых взрослых, поэтому процедура тестирования включала не только тесты, необходимые для данного исследования, но и ряд дополнительных заданий. В частности, участники также читали списки слов и псевдослов [Дорофеева и др., 2021]; эти материалы будут проанализированы и опубликованы позже.

Предварительно участники проходили тест для оценки уровня невербального интеллекта. Во время основного тестирования сначала мы просили участников прочесть текст «Неблагодарная ель» [Корнев, Ишимова, 2010] и ответить на вопросы к нему. Затем участнику предлагалось прочесть по одному листу слов и псевдослов [Дорофеева и др., 2021]. Все участники с четными номерами читали листы А и В, участники с нечетными номерами — листы С и D. После этого участники читали один из новых текстов и должны были ответить на вопросы к нему (участники с четными номерами читали текст «Пейзаж», участники с нечетными номерами — текст «Корабль»). Потом мы предлагали участникам прочесть оставшиеся листы слов и псевдослов и, наконец, второй из составленных нами текстов. Мы предупреждали участников, что после чтения текста им нужно будет ответить на вопросы, и при этом просили их читать вслух так быстро, как они могут читать без ошибок.

Для 15 участников процедура оказалась непреднамеренно нарушена: первым текстом им был предъявлен не текст «Неблагодарная ель», а текст «Пейзаж». Поэтому в дальнейшем все результаты статистического анализа приводятся на выборке из 96 человек, которые прошли процедуру исследования в точном соответствии с разработанной программой (средний возраст 19,5 года, SD = 1,2; 61 девушка). Данные 15 участников, проходивших тестирование с нарушенной процедурой, проанализированы отдельно. Результаты этого анализа описаны в разделе 4.2.1.

Чтение всех материалов фиксировалось на диктофон. При обработке результатов мы прослушивали записи каждого из заданий на чтение и считали количество верно прочитанных слов за первую минуту чтения. При этом неверное ударение и оговорки при условии самокоррекции, т.е. при наличии в записи правильно произнесенного целевого слова, не засчитывались за ошибочное прочтение. Для оценки техники чтения мы использовали три метрики: скорость чтения (количество верно прочитанных слов за одну минуту чтения), точность чтения (доля верно прочитанных слов от всех прочитанных за одну минуту), взвешенные баллы за ошибки, рассчитанные по методике, описанной в [Dorofeeva et al., 2019]. Кроме того, мы подсчитывали количество верных ответов на вопросы на понимание прочитанного.

3.3. План анализа

Основная задача настоящего исследования заключалась в разработке параллельных версий теста. Как было показано выше, тексты «Пейзаж» и «Корабль» составлены так, что их лингвистические характеристики не имеют значимых различий. Наша гипотеза состояла в том, что благодаря лингвистической сбалансированности версий теста в результатах участников по метрикам чтения двух версий теста также не обнаружится значимых различий. Для проверки гипотезы о параллельности версий мы используем три статистических критерия: равенство средних значений в выборке, равенство дисперсий и наличие корреляции [Henning, 1987].

Кроме того, мы ставили задачу проверить внешнюю валидность разработанного теста. В качестве внешнего критерия для сравнения использовались результаты тестирования чтения при помощи СМИНЧ. Данная методика предназначена для выявления дислексии у детей, и мы выбрали наиболее сложный из включенных в нее текстов. СМИНЧ не предназначен для тестирования взрослых, но на момент проведения исследования это был единственный стандартизированный тест для оценки скорости чтения текста на русском языке, поэтому мы сочли возможным его использовать. Для оценки внешней валидности проведен корреляционный анализ.

Разработанный лингвистически сбалансированный тест предполагается применять не только для исследовательских, но и для диагностических целей, т.е. для выявления нарушений чтения у взрослых. Поэтому мы рассчитали контрольные диагностические уровни в группе молодых взрослых для оценки следующих параметров: скорости чтения, точности чтения (с использованием двух вариантов метрик) и количества правильных ответов на вопросы.

В анализе использованы тест Шапиро — Уилка на нормальность распределения, тесты Уилкоксона и Манна — Уитни для сравнения средних значений в выборках, тест Левена для срав-

нения дисперсий выборок и критерий корреляции Спирмена для корреляционного анализа. Кроме того, в случае множественных сравнений применялась поправка Бонферрони, для того чтобы избежать ложноположительных результатов. При довольно большом числе сравнений она может оказаться слишком строгой, но ее тем не менее можно использовать при числе пар сравнений, не превышающем восемь [Мастицкий, Шитиков, 2015]. Кроме того, строгость поправки при определении параллельности текстов в некоторой степени полезна, так как позволяет избежать ложноположительных результатов и исключает признание параллельными тестов, не являющихся таковыми.

4. Результаты

4.1. Описательная статистика

Данные описательной статистики по скорости и точности чтения двух текстов, а также по точности ответов на вопросы приведены в табл. 8.

Таблица 8. Описательная статистика по метрикам чтения: минимум — максимум (медиана; среднее $\pm$ стандартное отклонение), $n = 96$

Текст	Возраст	Скорость чтения	Точность чтения	Количество взведенных баллов за ошибки	Количество верных ответов на вопросы
«Пейзаж»	18	100 – 164 (136,5; 135,67 $\pm$ 19,17)	0,981 – 1 (1; 0,99 $\pm$ 0,01)	0,5 – 8 (3,2; 3,69 $\pm$ 2,35)	4 – 10 (7,5; 7,39 $\pm$ 1,82)
	19	103 – 180 (148; 147,69 $\pm$ 19,50)	0,936 – 1 (1; 0,99 $\pm$ 0,01)	0 – 21 (4; 5,27 $\pm$ 5,09)	2 – 10 (7; 6,64 $\pm$ 2,27)
	20	96 – 196 (140,5; 146,75 $\pm$ 25,32)	0,938 – 1 (1; 0,99 $\pm$ 0,01)	0 – 13,5 (4,2; 4,45 $\pm$ 3,59)	1 – 11 (6; 6,15 $\pm$ 2,54)
	21	122 – 196 (156; 156,10 $\pm$ 22,82)	0,982 – 1 (1; 0,99 $\pm$ 0,01)	1 – 9,5 (3,8; 4,65 $\pm$ 3,29)	5 – 9 (7; 7,10 $\pm$ 1,10)
	22	136 – 187 (165; 162,22 $\pm$ 20,55)	0,993 – 1 (1; 1,00 $\pm$ 0,00)	0 – 9 (2,5; 3,67 $\pm$ 3,00)	3 – 10 (7; 7,44 $\pm$ 2,19)
«Корабль»	18–22	96 – 196 (145,5; 147,48 $\pm$ 22,04)	0,936 – 1 (1; 0,99 $\pm$ 0,01)	0 – 21 (4; 4,59 $\pm$ 4,02)	1 – 11 (7; 6,80 $\pm$ 2,16)
	18	110 – 175 (135; 137,72 $\pm$ 17,05)	0,973 – 1 (1; 0,99 $\pm$ 0,01)	1 – 8,5 (4,8; 4,39 $\pm$ 2,20)	3 – 10 (7,5; 7,17 $\pm$ 2,12)
	19	106 – 185 (145; 144,82 $\pm$ 19,77)	0,952 – 1 (1; 0,99 $\pm$ 0,01)	0 – 19,5 (4,5; 5,17 $\pm$ 4,69)	1 – 11 (8; 7,00 $\pm$ 2,51)
	20	107 – 191 (143; 148,45 $\pm$ 22,39)	0,979 – 1 (1; 0,99 $\pm$ 0,01)	0,5 – 8,5 (3,2; 3,85 $\pm$ 2,49)	4 – 10 (7; 7,35 $\pm$ 1,76)
	21	119 – 192 (145; 152,70 $\pm$ 26,28)	0,98 – 1 (1; 1,00 $\pm$ 0,01)	1 – 7 (3,8; 3,75 $\pm$ 1,69)	4 – 9 (7,5; 7,10 $\pm$ 1,66)
	22	129 – 193 (161; 158,78 $\pm$ 24,23)	0,978 – 1 (1; 0,99 $\pm$ 0,01)	2 – 16,5 (6,5; 7,61 $\pm$ 4,25)	3 – 9 (7; 6,78 $\pm$ 1,64)
	18–22	106 – 193 (143,5; 146,38 $\pm$ 21,41)	0,952 – 1 (1; 0,99 $\pm$ 0,01)	0 – 19,5 (4; 4,83 $\pm$ 3,72)	1 – 11 (7; 7,09 $\pm$ 2,11)

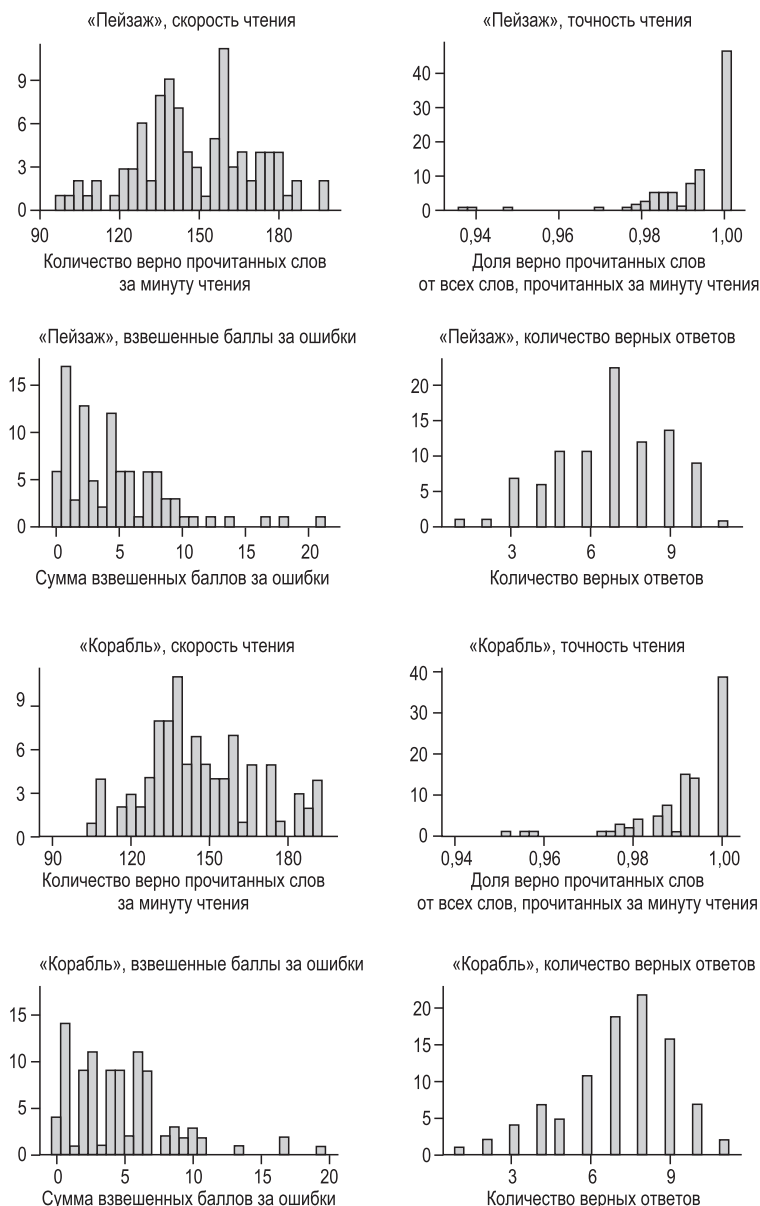
Распределение результатов в выборке мы оценивали при помощи теста Шапиро — Уилка (табл. 9). Распределение соответствовало нормальному только для скорости чтения текста «Пейзаж» ($W = 0,99, p = 0,496$). Все остальные метрики нарушали закон нормального распределения. Гистограммы, позволяющие визуально оценить распределения, представлены на рис. 3. Можно заметить, что нарушения нормального распределения в случае с точностью чтения и взвешенными баллами за ошибки (смещением выборки к максимуму или минимуму) связаны с эффектом потолка или пола соответственно. Поскольку в исследовании принимали участие люди, не испытывающие трудностей с чтением, они допускали очень мало ошибок при чтении. Таким образом, точность чтения часто близка к максимально возможному показателю — к 1, а количество взвешенных баллов за ошибки сравнительно мало. Поскольку в рассматриваемой выборке ни для одной пары сравнения по целевым метрикам не оказалось такого случая, когда оба распределения данных соответствуют критериям нормального, сравнение средних мы производили при помощи непараметрических тестов: теста Уилкоксона в случае зависимых выборок и теста Манна — Уитни в случае независимых выборок, а сравнение дисперсий — при помощи теста Левена. Корреляционный анализ осуществлялся при помощи критерия Спирмена.

Таблица 9. Результаты теста Шапиро — Уилка на нормальность распределения для целевых метрик ($n = 96$)

		<i>W</i>	<i>p-value</i>
«Пейзаж»	Скорость чтения	0,99	0,496
	Точность чтения	0,66	< 0,001
	Количество взвешенных баллов за ошибки	0,87	< 0,001
	Количество верных ответов на вопросы	0,96	0,005
«Корабль»	Скорость чтения	0,97	0,022
	Точность чтения	0,77	< 0,001
	Количество взвешенных баллов за ошибки	0,89	< 0,001
	Количество верных ответов на вопросы	0,94	< 0,001

Мы проверили наличие корреляции результатов чтения с возрастом участников, количеством лет образования и уровнем невербального интеллекта. Скорость чтения положительно коррелировала с возрастом в случае текста «Пейзаж» (тест Спирмена: $r_s = 0,28, p = 0,01$). Значимой корреляции скорости чтения с количеством лет образования и уровнем невербального интеллекта, а также с возрастом в случае теста «Корабль» не обнаружено. Кроме того, ни точность чтения, ни успешность ответов на вопросы оказались не связаны ни с одной из проанализированных характеристик (табл. 10 и 11).

Рис. 3. Гистограммы распределения данных (на оси Y на всех графиках отражено число участников, показавших указанный на оси X результат)



При сравнении результатов тестов на чтение в группах юношей и девушек на основании критерия Манна — Уитни обнаружены значимые различия в числе взвешенных баллов за ошибки в случае с текстом «Корабль». Во всех остальных метриках различий не выявлено (табл. 12).

Таблица 10. Коэффициент корреляции Спирмена (в скобках указан уровень значимости) между результатами тестов на чтение (текст «Пейзаж») и характеристиками участников теста

	Возраст	Количество лет образования	Уровень невербального интеллекта
Скорость	0,28 (0,006)*	0,16 (0,112)	−0,08 (0,473)
Точность	0 (0,966)	0,04 (0,693)	0,05 (0,65)
Сумма взвешенных баллов за ошибки	0,01 (0,887)	−0,06 (0,532)	−0,07 (0,543)
Количество верных ответов на вопросы	−0,05 (0,624)	0,07 (0,509)	0,08 (0,49)

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони (цифры в таблице указаны до применения поправки). Результаты, сохранившие свою значимость при данном числе сравнений (три пары в каждой группе сравнений; $p < 0,017$), отмечены *.

Для всех пар сравнений число участников равнялось 96, число степеней свободы во всех случаях равно 94.

Таблица 11. Коэффициент корреляции Спирмена (в скобках указан уровень значимости) между результатами тестов на чтение (текст «Корабль») и характеристиками участников теста

	Возраст	Количество лет образования	Уровень невербального интеллекта
Скорость	0,23 (0,023)	0,13 (0,211)	−0,04 (0,721)
Точность	0,05 (0,649)	0,03 (0,740)	0,02 (0,894)
Сумма взвешенных баллов за ошибки	0,06 (0,568)	0,09 (0,407)	−0,15 (0,176)
Количество верных ответов на вопросы	−0,06 (0,56)	0,05 (0,598)	0,06 (0,569)

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном числе сравнений (три пары в каждой группе сравнений; $p < 0,017$), отсутствуют.

Для всех пар сравнений число участников равнялось 96, число степеней свободы во всех случаях равно 94.

Таблица 12. Сравнение результатов чтения в группах юношей и девушек: минимум — максимум (медиана; среднее $\pm$ стандартное отклонение)

		Девушки ($n = 61$)	Юноши ($n = 35$)	W	P
Пейзаж	Скорость чтения	96 – 196 (143; 146,15 $\pm$ 21,74)	103 – 196 (148; 149,80 $\pm$ 22,69)	962	0,424
	Точность чтения	0,938 – 1 (1; 0,99 $\pm$ 0,01)	0,936 – 1 (1; 0,99 $\pm$ 0,01)	1169	0,408
	Сумма взвешенных баллов за ошибки	0 – 13,5 (4; 4,04 $\pm$ 3,10)	0 – 21 (4,5; 5,54 $\pm$ 5,16)	97	0,321
	Количество верных ответов на вопросы	1 – 11 (7; 6,87 $\pm$ 2,09)	3 – 10 (7; 6,69 $\pm$ 2,30)	1114	0,723
Корабль	Скорость чтения	106 – 192 (144; 145,97 $\pm$ 19,69)	107 – 193 (143; 147,09 $\pm$ 24,40)	1070	0,988

Окончание табл. 12

		Девушки ($n = 61$)	Юноши ($n = 35$)	W	P
	Точность чтения	0,977 – 1 (1; 0,99 ± 0,01)	0,952 – 1 (1; 0,99 ± 0,01)	1272	0,106
	Сумма взвешенных баллов за ошибки	0 – 10,5 (3,5; 3,89 ± 2,70)	0 – 19,5 (6; 6,47 ± 4,63)	693	0,004
	Количество верных ответов на вопросы	1 – 11 (7; 6,97 ± 2,03)	3 – 11 (8; 7,31 ± 2,26)	901	0,200

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном числе сравнений (четыре пары сравнения для каждого текста; $p < 0,013$), отмечены *.

4.2. Проверка параллельности версий

Для проверки параллельности версий проведен сравнительный и корреляционный анализ скорости чтения, точности чтения, взвешенных баллов за ошибки и количества верных ответов на вопросы.

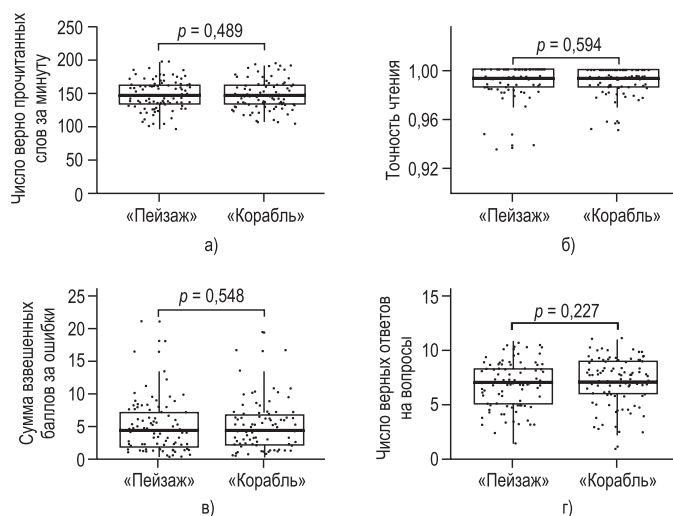
Значимых различий между двумя текстами ни в средних значениях, ни в дисперсиях во всех четырех метриках не обнаружено. Тест Уилкоксона показал $p > 0,05$, тест Левена — $p > 0,05$ при сравнении дисперсий. Кроме того, метрики скорости чтения, точности чтения, взвешенных баллов за ошибки и количества верных ответов на вопросы двух текстов значимо и положительно коррелировали между собой (табл. 13, рис. 4 и 5).

Таблица 13. Сравнение медиан (тест Уилкоксона) и дисперсий выборок (тест Левена), а также результаты корреляционного анализа по критерию Спирмена между двумя текстами ($n = 96$)

	Тест Уилкоксона		Тест Левена		Корреляция Спирмена	
	W	p	W	p	r_s	p
Скорость чтения	1961	0,489	0,35	0,554	0,9*	< 0,001
Точность чтения	1184	0,594	0,42	0,517	0,33*	0,001
Количество взвешенных баллов за ошибки	2342	0,5485	0,29	0,589	0,40*	< 0,001
Количество верных ответов на вопросы	1960	0,2273	0,05	0,831	0,44*	< 0,001

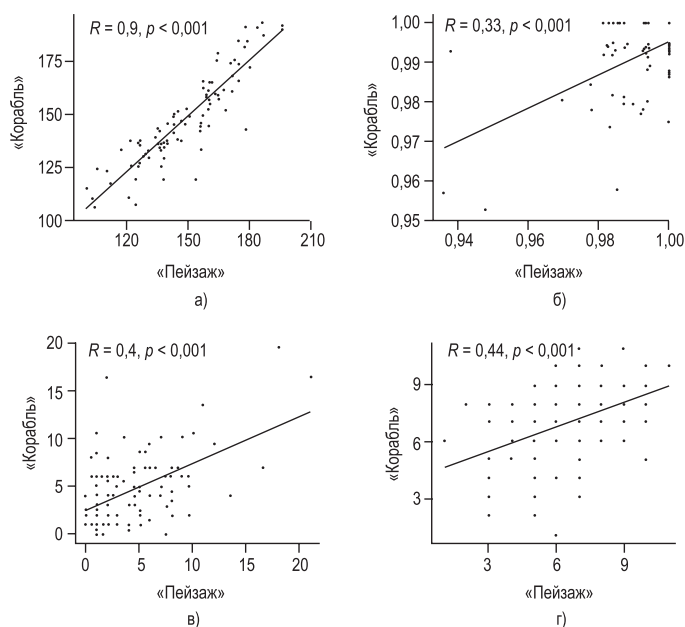
Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном числе сравнений (четыре пары сравнения для каждого текста; $p < 0,013$), отмечены *.

Рис. 4. Сравнение результатов чтения двух текстов по всем основным метрикам при помощи критерия Уилкоксона



а) скорость чтения (количество верно прочитанных слов за первую минуту чтения); б) точность чтения (доля верно прочитанных слов от всех прочитанных слов за первую минуту чтения); в) сумма взвешенных баллов за ошибки; г) количество верных ответов на вопросы

Рис. 5. Коэффициент корреляции Спирмена для всех основных метрик чтения текстов и уровень значимости



а) скорость чтения (количество верно прочитанных слов за первую минуту чтения); б) точность чтения (доля верно прочитанных слов от всех прочитанных слов за первую минуту чтения); в) сумма взвешенных баллов за ошибки; г) количество верных ответов на вопросы

**4.2.1. Влияние
порядка предъявления текстов
на скорость и точность чтения и на
уровень понимания прочитанного**

Для исследования влияния порядка предъявления текстов на скорость и точность чтения, а также на уровень понимания прочитанного, мы использовали данные тех 15 участников (средний возраст 19,2 года, $SD = 1,0$), для которых процедура тестирования была непреднамеренно нарушена (см. также раздел 3.2). Все эти участники читали текст «Пейзаж» первым и отвечали на вопросы к нему. Затем им предлагался для чтения один из листов слов и один из листов псевдослов [Дорофеева и др., 2021]. После этого участники читали текст «Неблагодарная ель» [Корнев, Ишимова, 2010] или текст «Корабль» и отвечали на вопрос к одному из этих текстов, затем читали оставшийся лист слов и псевдослов, и в конце тестирования читали оставшийся текст и отвечали на вопросы к нему. Кроме того, в этом анализе использовались данные 96 участников из основной выборки.

Мы сравнили результаты участников, читавших текст «Пейзаж» первым, с результатами участников, которые читали его в середине или в финале тестирования, при помощи теста Манна — Уитни. Количество верных ответов на вопросы были значительно ниже, если текст читался в начале тестирования. На скорость и точность чтения (рассчитанную и как отношение числа верно прочитанных слов к числу всех прочитанных слов, и как количество взвешенных баллов за ошибки) порядок предъявления текстов не влиял (табл. 14).

Таблица 14. **Результаты чтения текста «Пейзаж» (минимум — максимум (медиана; среднее $\pm$ стандартное отклонение)) при разном порядке тестирования и сравнение при помощи теста Уилкоксона**

	Текст читался первым ($n = 15$)	Текст читался не первым ($n = 96$)	W	p
Скорость чтения	109–155 (132; 133,73 $\pm$ 13,99)	96–196 (145,5; 147,48 $\pm$ 22,04)	996	0,018
Точность чтения	0,99–1 (1; 1,00 $\pm$ 0,01)	0,94–1 (1; 0,99 $\pm$ 0,01)	673	0,668
Взвешенные баллы за ошибки	0–9 (3; 3,17 $\pm$ 2,46)	0–21 (4; 4,59 $\pm$ 4,02)	847	0,274
Количество верных ответов на вопросы	1–7 (5; 4,47 $\pm$ 2,13)	1–11 (7; 6,80 $\pm$ 2,16)	1119	< 0,001*

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном числе сравнений (4 пары; $p < 0,013$), отмечены *.

Такой же анализ проведен для текста «Неблагодарная ель». Как и при использовании текста «Пейзаж», порядок предъявления влиял на успешность понимания текста: понимание было значительно менее успешным, если текст располагался в начале батареи. При этом порядок тестирования не влиял на скорость чтения и точность чтения (табл. 15).

	Текст читался первым ($n = 96$)	Текст читался не первым ($n = 15$)	W	p
Скорость чтения	108–222 (160; 163,38 ± 26,03)	130–187 (156; 156,00 ± 18,56)	830	0,343
Точность чтения	0,93–1 (1; 0,99 ± 0,01)	0,97–1 (1; 0,99 ± 0,01)	642	0,500
Взвешенные баллы за ошибки	0–34 (5; 6,50 ± 5,96)	0–8,5 (2; 2,87 ± 2,39)	999	0,013
Количество верных ответов на вопросы	0–10 (5; 4,99 ± 2,21)	6–10 (8; 7,53 ± 1,25)	238	< 0,001*

Таким образом, мы обнаружили, что порядок предъявления влияет на успешность понимания текста (понимание является значительно менее успешным, если текст располагается в начале батареи) и не оказывает значимого влияния на скорость и точность чтения.

Для оценки внешней валидности нового теста на чтение мы провели анализ корреляций между результатами нового теста и СМИНЧ на основании критерия Спирмена. Обнаружена значимая высокая корреляция скорости чтения текста «Неблагодарная ель» со скоростью чтения как текста «Пейзаж» ($r_s = 0,82, p < 0,001$), так и текста «Корабль» ($r_s = 0,86, p < 0,001$). Корреляции по количеству верных ответов на вопросы оказались ниже, но остались значимыми после применения поправки Бонферрони для обоих текстов ($r_s = 0,5, p < 0,001$ для текста «Пейзаж», $r_s = 0,38, p < 0,001$ для текста «Корабль»). Кроме того, для текста «Корабль» количество взвешенных баллов за ошибки показало значимую корреляцию с количеством взвешенных баллов за ошибки при чтении текста «Неблагодарная ель» ($r_s = 0,43, p < 0,001$) (табл. 16 и 17).

	«Неблагодарная ель»				
		Скорость чтения	Точность чтения	Сумма взвешенных баллов за ошибки	Количество верных ответов на вопросы
«Пейзаж»	Скорость чтения	0,82 (<0,001)*	0,04 (0,691)	0,11 (0,273)	0,07 (0,528)
	Точность чтения	0,02 (0,854)	0,21 (0,043)	-0,14 (0,167)	-0,04 (0,674)

Окончание табл. 16

«Неблагодарная ель»		Скорость чтения	Точность чтения	Сумма взвешенных баллов за ошибки	Количество верных ответов на вопросы
	Сумма взвешенных баллов за ошибки	0,08 (0,445)	–0,11 (0,303)	0,26 (0,012)	–0,02 (0,879)
	Количество верных ответов на вопросы	0,04 (0,688)	0,11 (0,276)	0,02 (0,812)	0,5 (<0,001)*

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном количестве сравнений (16 пар для каждого текста; $p < 0,003$), отмечены *. Для всех пар сравнений число участников равнялось 96, число степеней свободы во всех случаях равно 94.

Таблица 17. Коэффициент корреляции Спирмена (в скобках указан уровень значимости) между результатами СМИНЧ и нового теста (текст «Корабль»)

		«Неблагодарная ель»			
		Скорость чтения	Точность чтения	Сумма взвешенных баллов за ошибки	Количество верных ответов на вопросы
«Корабль»	Скорость чтения	0,86 (<0,001)*	0,09 (0,393)	0,11 (0,273)	0,03 (0,805)
	Точность чтения	0,04 (0,667)	0,26 (0,009)	–0,25 (0,013)	0,01 (0,922)
	Сумма взвешенных баллов за ошибки	0,09 (0,408)	–0,26 (0,011)	0,43 (<0,001)*	0,27 (0,007)
	Количество верных ответов на вопросы	0,02 (0,837)	–0,02 (0,839)	0,11 (0,295)	0,38 (<0,001)*

Примечание. Результаты, значимые на уровне $p = 0,05$, выделены жирным. Мы применили поправку Бонферрони. Результаты, сохранившие свою значимость при данном числе сравнений (16 пар; $p < 0,003$), отмечены *. Для всех пар сравнений число участников равнялось 96, число степеней свободы во всех случаях равно 94.

4.4. Контрольные диагностические уровни

В российской практике в данной области тестирования для расчета контрольных диагностических уровней традиционно используется подход с учетом среднего и стандартных отклонений [Корнев, Ишимова, 2010; Бурдына и др., 2023; Лезина и др., 2023]. Поскольку наши данные не имеют нормального распределения, предпочтительно использовать для определения контрольных уровней процентиля. Мы построили таблицы с расчетами и на основе средних (табл. 18 и 20), и на основе процентилей (табл. 19 и 21). И в табл. 19, и в табл. 21 расчеты проводились так, что 10-й процентиль отсекает 10% участников с самыми худшими результатами, а 95-й процентиль — 5% участников с самыми лучшими результатами. Во всех таблицах для большей информативности приведены результаты с точностью до двух знаков после запятой. При использовании тестов на практике скорость чтения и количество верных ответов на вопросы для каждого участника измеряются в целых числах, поэтому приведенные в таблицах значения рекомендуется округлять до целых чисел.

Таблица 18. Контрольные уровни для оценки скорости чтения, точности чтения как доли верно прочитанных слов и понимания прочитанного, рассчитанные на основании средних и стандартных отклонений

Метрика	Текст	Среднее	Медиана	Стандартное отклонение	Среднее минус одно стандартное отклонение	Среднее минус полтора стандартных отклонения	Среднее минус два стандартных отклонения
Скорость чтения (слов в минуту)	«Пейзаж»	147,48	145,5	22,04	125,43	114,41	103,39
	«Корабль»	146,38	143,5	21,41	124,97	114,26	103,56
Точность чтения (доля верно прочитанных слов)	«Пейзаж»	0,99	0,99	0,01	0,98	0,97	0,97
	«Корабль»	0,99	0,99	0,01	0,98	0,98	0,97
Понимание прочитанного (количество верных ответов на вопросы)	«Пейзаж»	6,8	7	2,16	4,64	3,56	2,48
	«Корабль»	7,09	7	2,11	4,98	3,92	2,87

Таблица 19. Контрольные уровни для оценки скорости чтения, точности чтения как доли верно прочитанных слов и понимания прочитанного, рассчитанные на основании процентилей

Метрика	Текст	5-й процентиль	7-й процентиль	10-й процентиль	95-й процентиль
Скорость чтения (слов в минуту)	«Пейзаж»	109	115,25	122	181
	«Корабль»	113,75	118,3	121	187,75
Точность чтения (доля верно прочитанных слов)	«Пейзаж»	0,978	0,98	0,982	1
	«Корабль»	0,977	0,978	0,98	1
Понимание прочитанного (количество верных ответов на вопросы)	«Пейзаж»	3	3	4	10
	«Корабль»	3	3,65	4	10

Таблица 20. Контрольные уровни для оценки количества взвешенных баллов за ошибки, рассчитанные на основании средних и стандартных отклонений

Метрика	Текст	Среднее	Медиана	Стандартное отклонение	Среднее плюс одно стандартное отклонение	Среднее плюс полтора стандартных отклонения	Среднее плюс два стандартных отклонения
Количество взвешенных баллов за ошибки	«Пейзаж»	4,59	4	4,02	8,61	10,62	12,63
	«Корабль»	4,83	4	3,72	8,55	10,41	12,27

Таблица 21. Контрольные уровни для оценки количества взвешенных баллов за ошибки, рассчитанные на основании процентилей

Метрика	Текст	5-й процентиль	7-й процентиль	10-й процентиль	95-й процентиль
Количество взвешенных баллов за ошибки	«Пейзаж»	11,25	9,68	9,25	4
	«Корабль»	10,5	10	9,5	4

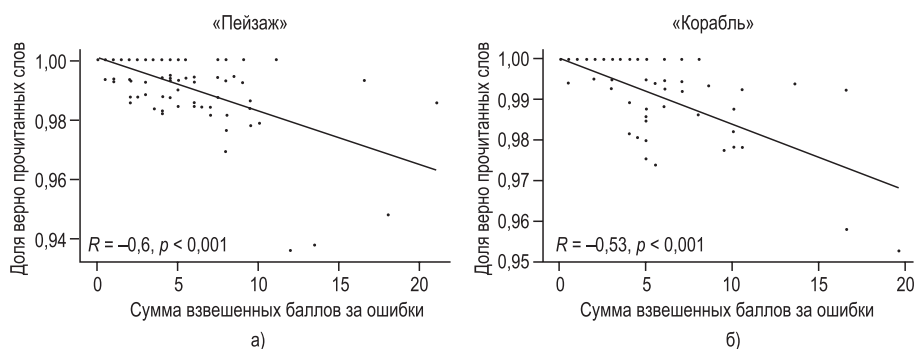
Высокие баллы по скорости чтения, точности чтения как доли верно прочитанных слов и понимания прочитанного как количества верных ответов на вопросы свидетельствуют о высоком уровне навыка чтения. При этом высокие баллы по точности чтения, рассчитанной как сумма взвешенных баллов за ошибки, напротив, показывают более низкий уровень навыка. Поэтому для удобства использования контрольные уровни по взвешенным баллам за ошибки приводятся в отдельных таблицах (табл. 20 и 21).

Мы рекомендуем в качестве диагностических использовать контрольные уровни, рассчитанные на основе процентилей и представленные в табл. 19 и 21. В пользу такого выбора можно привести два аргумента. Во-первых, по большинству метрик в нашей выборке данные не имеют нормального распределения, и если при значительном увеличении количества участников выборки есть вероятность, что распределение данных по скорости чтения станет соответствовать нормальному, то в случае с оценкой точности чтения имеющееся распределение с большой вероятностью сохранится. Во-вторых, полученные показатели имеют достаточно большие стандартные отклонения, и при использовании подхода, основанного на средних и стандартных отклонениях, контрольные уровни по некоторым метрикам могут иметь отрицательные значения (см., например, обсуждение способа рассчитывать контрольные уровни в [Лезина и др., 2023]), что непригодно для практического применения.

Таким образом, показателями в пределах нормы следует считать результаты выше уровня 10-го процентиля, показателями риска следует считать результаты от уровня 10-го процентиля до 7-го процентиля включительно, нарушением — показатели ниже уровня 7-го процентиля, тяжелым нарушением — показатели ниже уровня 5-го процентиля. Результаты выше уровня 95-го процентиля могут использоваться для выявления участников с самым высоким уровнем навыков чтения (сверхуспешных) (табл. 19 и 21).

Для оценки точности чтения в нашем исследовании используются две метрики: точность чтения, рассчитанная как доля верно прочитанных слов от всех прочитанных слов, и сумма взвешенных баллов за ошибки. Результаты с использованием этих двух метрик значимо коррелировали между собой ($rs = -0,60$, $p < 0,001$ для текста «Пейзаж»; $rs = -0,53$, $p < 0,001$ для текста «Корабль») (рис. 6). Поскольку подсчет взвешенных баллов за ошибки является гораздо более трудоемкой задачей, чем оценка точности чтения как доли верно прочитанных слов, на практике при необходимости экономии времени может быть рекомендована оценка точности чтения как доли верно прочитанных слов. Однако оценка по сумме взвешенных баллов за ошибки более информативна, она позволяет обнаружить отличия там, где при более грубой оценке точности чтения они нивелируются. Взвешенные баллы за ошиб-

Рис. 6. Корреляция двух метрик точности чтения



а) текст «Пейзаж»; б) текст «Корабль»

ки рекомендуется использовать в исследованиях техники чтения или при диагностике, если интерес представляет именно техника чтения и важно отследить тонкие различия.

5. Обсуждение результатов

Благодаря тому что разработанные для нового диагностического инструментария тексты «Пейзаж» и «Корабль» не различаются по лингвистическим характеристикам, ни скорость чтения, ни точность чтения, ни сумма взвешенных баллов за ошибки, ни количество верных ответов на вопросы при использовании первой и второй версий теста не показали значимых различий ни в средних значениях, ни в дисперсиях выборок. Помимо этого, результаты использования первой версии теста значимо положительно коррелировали с результатами второй версии теста. Следовательно, опираясь на определение эквивалентности версий теста [Henning, 1987], можно утверждать, что составленные нами две версии теста на чтение параллельны.

Внешняя валидность разработанного теста оценивалась на основании анализа корреляций между результатами нашего теста и результатами ранее стандартизированного теста из СМИНЧ. Поскольку ввиду отсутствия теста для оценки навыков чтения у взрослых, который можно было бы использовать в качестве «золотого стандарта», мы использовали тест для оценки навыков чтения у детей, о результатах оценки внешней валидности следует говорить с осторожностью. Тем не менее по меньшей мере три из четырех использованных нами метрик (скорость чтения, сумма взвешенных баллов за ошибки и количество верных ответов на вопросы) показали значимые корреляции в данном случае.

Мы рассмотрели два способа оценивать точность чтения: как долю верно прочитанных слов от всех прочитанных за одну минуту слов и как сумму взвешенных баллов за ошибки. Первый способ менее трудозатратен, поэтому, скорее всего, именно он

будет использоваться на практике. Однако расчет взвешенных баллов за ошибки позволяет провести более подробный анализ полученных данных и повысить диагностическую валидность теста: выявить тех участников с дислексией, у кого скорость чтения и уровень понимания прочитанного оказался в пределах нормы, а точность чтения нарушена [Dorofeeva et al., 2019]. Поэтому именно взвешенная оценка ошибок может быть рекомендована для детальной диагностики, а также для использования в исследованиях механизмов нарушения чтения.

Скорость чтения, в отличие от других рассчитанных нами метрик, имеет тенденцию к положительной корреляции с возрастом участников. Поскольку мы рассматривали довольно узкий возрастной диапазон, контрольные уровни рассчитаны для всего диапазона полностью. Однако наличие корреляции скорости чтения с возрастом означает, что для диагностики нарушений чтения в других возрастных группах требуются рассчитанные именно для них уровни отсечки. Сбор нормативных данных для других возрастных групп может быть задачей следующих исследований.

Мы предлагаем оценивать понимание текста на основании открытых послетекстовых вопросов. При этом в дальнейшем может быть полезно дополнить тест другими способами оценки понимания текста, например, при помощи пересказа или ключевых слов, а также сравнить диагностическую валидность разных способов оценки понимания прочитанного.

Результаты исследования свидетельствуют о значимости порядка предъявления заданий для оценки успешности понимания прочитанного: количество верных ответов на вопросы к тексту, который предъявлялся первым, было значимо ниже, чем к текстам, которые предъявлялись в середине или в конце тестирования. Возможно, участники не могли в полной мере осознать сразу степень детальности, с которой нужно запоминать содержание текста для ответов на вопросы. Поэтому влияние порядка предъявления заданий необходимо учитывать, если тест используется для оценки навыка понимания прочитанного текста. Таким образом, для того чтобы приведенные в данной статье уровни отсечки по метрике «количество верных ответов на вопросы» можно было использовать в диагностических целях, необходимо предъявлять сначала какой-либо иной текст, к примеру, из СМИНЧ, и вопросы к нему, давая тем самым участнику возможность лучше понять процедуру тестирования. Однако этот результат мы получили на основании данных только 15 участников, поэтому он не может считаться окончательным. Вопрос о том, влияет ли порядок предъявления текстов на их понимание, заслуживает отдельного, более тщательного рассмотрения.

Предыдущие исследования скорости чтения показали, что уровень невербального интеллекта коррелирует со скоростью

чтения у детей [Johann, Könen, Karbach, 2020] и пожилых взрослых [Nelson, Willison, 1991]. В данном исследовании подобной корреляции не обнаружено. Одно из возможных объяснений такого результата — относительно небольшой разброс показателей с уровнем невербального интеллекта в нашей выборке.

6. Заключение

Цель исследования состояла в разработке теста для оценки навыков чтения у взрослых носителей русского языка, поскольку на момент проведения исследования стандартизированные инструменты, позволяющие оценить навыки чтения у русскоязычных взрослых и диагностировать возможные нарушения, отсутствовали. Чтобы разработанный инструментарий можно было использовать не только для первичной диагностики, но и для повторного тестирования, составлены два текста, имеющих сходную синтаксическую структуру и не обнаруживающих значимых различий по психолингвистическим параметрам входящих в них слов (длине, числу слогов и частотности). Каждый текст сопровождался набором вопросов на понимание, причем наборы вопросов также имели одинаковую структуру. Разработанный тест позволяет оценить как скорость и точность чтения, так и успешность понимания прочитанного текста. Благодаря тщательной лингвистической сбалансированности двух версий теста участники не показали значимых различий ни в скорости чтения, ни в точности чтения, ни в понимании прочитанного. Таким образом, создан инструмент, который позволит не только провести первичную диагностику нарушений чтения у взрослых, но и оценить динамику развития навыков чтения или эффект от коррекционной работы через некоторый промежуток времени.

В качестве основных пользователей разработанного диагностического инструмента мы видим исследователей, изучающих механизмы чтения и его нарушения у взрослых, и специалистов-практиков. В научных исследованиях предлагаемый инструмент может использоваться для обоснованного отбора в исследования взрослых участников с успешно сформированным навыком чтения или с нарушенным чтением, а также для оценки влияния различных факторов на скорость или точность чтения. Специалисты-практики — логопеды, нейропсихологи, дефектологи, неврологи и др. — с помощью данного инструмента смогут оценить навык чтения любого взрослого носителя русского языка и дать обоснованный ответ на запрос клиента или документ о наличии или отсутствии нарушений чтения. Важно подчеркнуть, что принятие решения о постановке диагноза «дислексия» требует комплексного подхода, и тестирование навыка чтения является необходимой, но не единственной его составляющей. Диагностику должен проводить специалист, имеющий соответствующую

квалификацию. Широкое применение разработанного нами инструментария для самостоятельного тестирования навыков чтения не предусмотрено. Более того, открытое распространение текстов данного теста и знакомство с ними широкой аудитории нежелательно, так как может снизить диагностическую эффективность инструмента, ведь метрики скорости, точности чтения и понимания прочитанного могут существенно меняться при повторном, а тем более многократном прочтении любого подобного материала.

Среди ограничений проведенного исследования можно выделить относительно небольшой объем выборки и отсутствие надежного критерия для оценки внешней валидности. С одной стороны, небольшая и довольно однородная выборка могла стать причиной того, что мы не обнаружили корреляций показателей развития навыка чтения с уровнем интеллекта, которую можно было ожидать на основании результатов других исследований. С другой стороны, узкий возрастной диапазон позволил нам рассчитать диагностические уровни уже на таком объеме, хотя бы для этой возрастной группы — молодых взрослых.

Полученные данные могут быть очень полезны для диагностики нарушений чтения у студентов высших учебных заведений, для которых навык чтения чрезвычайно важен. Отсутствие своевременной диагностики дислексии может приводить к существенным негативным последствиям у студентов с трудностями чтения [Cavalli et al., 2017; Konur, 2002]. При этом наличие инструментария для обоснованной диагностики дислексии могло бы помочь вузам принять решение о мерах помощи таким студентам. В мировой практике используются такие способы поддержки, как выделение дополнительного времени для выполнения контрольных или экзаменационных заданий студентам с дислексией, использование специального программного обеспечения (например, для проверки орфографии), при наличии возможностей у вузов — предоставление студентам с дислексией аудио- или видеoverсий заданий, а также материалов для изучения, при наличии ресурсов и специалистов — организация платного или бесплатного дополнительного коррекционного обучения, направленного на компенсацию дефицита [Tops et al., 2022]. Учитывая, что большинство мер потребовали бы дополнительных ресурсов от вузов, мы ожидаем, что наиболее реалистичными первыми шагами в плане помощи студентам с дислексией могут быть дополнительное время на контрольных работах и экзаменах, а также практика неснижения оценок за специфические ошибки при письме студентам с подтвержденной дислексией в случае, если ответ верен по существу, а основным навыком, тренируемым на том или ином курсе, не является именно чтение или письмо.

Наиболее актуальными направлениями дальнейшей работы с тестом для оценки навыков чтения у взрослых является, во-пер-

вых, сбор данных в группе людей с дислексией, который позволит оценить диагностическую валидность теста. Однако добиться этого достаточно непросто, поскольку из-за отсутствия соответствующих стандартизированных диагностических инструментов численность русскоговорящих взрослых с надежно диагностированной дислексией невелика. Во-вторых, поскольку обнаружена тенденция к корреляции скорости чтения с возрастом, необходим сбор нормативных данных в других возрастных группах, в которых диагностические уровни могут отличаться от приведенных нами. В-третьих, требуется провести повторную разметку данных и оценить межэкспертную надежность теста, что особенно важно для взвешенных баллов за ошибки.

Благодарности

Статья подготовлена в ходе проведения работы в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики».

Авторы благодарят всех, кто принял участие в исследовании при сборе нормативных данных.

Литература

1. Бурдына А.В., Колтунцева В.Е., Никифорова Н.В., Марченко Е.Н., Сергеева О.А., Алексеева С.В. (2023) Оценка навыков чтения слов и псевдослов у русскоговорящих подростков 14–17 лет. Материалы конференции «Когнитивная наука в Москве: новые исследования» (Москва, 2023, 21–22 июня), сс. 558–563.
2. Валгина Н.С. (2003) *Теория текста*. М.: Логос.
3. Доброго А.С., Петрова Т.Е. (2016) Особенности восприятия статических и динамических текстов: экспериментальное исследование на материале русского языка. *Актуальные проблемы языкознания*, т. 1, сс. 33–34.
4. Дорофеева С.В., Гринько И.Ю., Перевощикова Т.Д., Драгой О.В. (2021) Разработка тестов на чтение слов и псевдослов для оценки навыков чтения у русскоговорящих детей. Материалы конференции «Когнитивная наука в Москве: новые исследования» (Москва, 2021, 23–24 июня), сс. 502–507.
5. Здорова Н.С., Старовойтова В.Н., Лопухина А.А., Джонбобоева Д.Б., Парешина Е.А., Шестакова Е.Р., Гмырина В.А., Дмитрова Е.П., Драгой О.В. (2025) *LexiMetr*: тест для оценки навыков чтения у младших школьников на русском языке. *Вопросы образования / Educational Studies Moscow*, № 2, сс. 73–100. <https://doi.org/10.17323/vo-2025-21693>
6. Корнев А.Н. (1997) *Нарушения чтения и письма у детей*. СПб.: МиМ.
7. Корнев А.Н., Ишимова О.А. (2010) *Методика диагностики дислексии у детей*. СПб.: Политехнический университет.
8. Лапошина А.Н., Лебедева М.Ю. (2021) Текстометр: онлайн-инструмент определения уровня сложности текста по русскому языку как иностранному. *Русистика*, т. 19, № 3, сс. 331–345. <https://doi.org/10.22363/2618-8163-2021-19-3-331-345>
9. Лапошина А.Н., Лебедева М.Ю., Берлин Хенис А.А. (2022) Влияние частотности слов текста на его сложность: экспериментальное исследование читателей младшего школьного возраста методом айтрекинга. *Russian Journal of Linguistics*, т. 26, № 2, сс. 493–514. <https://doi.org/10.22363/2687-0088-30084>

10. Лезина А.Д., Никонова Я.Е., Черевик А.С., Шатковская И.Г., Алексеева С.В. (2023) Чтение вслух текста на русском языке как инструмент для выявления риска дислексии у подростков 14–17 лет. Материалы конференции «Когнитивная наука в Москве: новые исследования» (Москва, 2023, 21–22 июня), сс. 599–603.
11. Мاستицкий С.Э., Шитиков В.К. (2015) *Статистический анализ и визуализация данных с помощью R*. М.: ДМК.
12. Мурзин Л.Н., Штерн А.С. (1991) *Текст и его восприятие*. Свердловск: Уральский университет.
13. Филиппов К.А. (2003) *Лингвистика текста*. СПб.: Санкт-Петербургский университет.
14. Хамаганова В.М. (2002) *Структурно-семантическая и лексическая модель текста типа «описание»: проблемы семиотики и онтологии* (Дисс. д-ра филол. наук). М.: Московский педагогический государственный университет.
15. Alexeeva S., Slioussar N., Chernova D. (2018) StimulStat: A Lexical Database for Russian. *Behavior Research Methods*, vol. 50, no 6, pp. 2305–2315. <https://doi.org/10.3758/s13428-017-0994-3>
16. Altpeter E.K., Marx T., Nguyen N.X., Naumann A., Trauzettel-Klosinski S. (2015) Measurement of Reading Speed with Standardized Texts: A Comparison of Single Sentences and Paragraphs. *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 253, no 8, pp. 1369–1375. <https://doi.org/10.1007/s00417-015-3065-4>
17. Catts H.W., Kamhi A.G. (1999) *Language and Reading Disabilities*. Boston: Allyn and Bacon.
18. Cavalli E., Colé P., Leloup G., Poracchia-George F., Sprenger-Charolles L., El Ahmadi A. (2017) Screening for Dyslexia in French-Speaking University Students: An Evaluation of the Detection Accuracy of the Alouette Test. *Journal of Learning Disabilities*, vol. 51, no 3, pp. 268–282. <https://doi.org/10.1177/0022219417704637>
19. Cline F., Johnstone C.J., King T.K. (2006) *Focus Group Reactions to Three Definitions of Reading (As Originally Developed in Support of NARAP Goal 1)*. Minneapolis, MN: National Accessible Reading Assessment Projects.
20. Desrochers A., Glickman V. (2009) Criteria for the Evaluation of Reading Assessment Tools. *Encyclopedia of Language and Literacy Development*. London, ON: Canadian Language and Literacy Research Network, pp. 1–9.
21. Dorofeeva S.V., Reshetnikova V., Serebryakova M., Goranskaya D., Akhutina T.V., Dragoy O. (2019) Assessing the Validity of the Standardized Assessment of Reading Skills in Russian and Verifying the Relevance of Available Normative Data. *The Russian Journal of Cognitive Science*, vol. 6, no 1, pp. 4–24.
22. Driscoll D.P. (2012) *Reading Framework for the 2013. National Assessment of Educational Process*. Available at: <https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/reading/2013-reading-framework.pdf> (accessed 05.05.2025).
23. Flesch R. (1948) A New Readability Yardstick. *Journal of Applied Psychology*, vol. 32, no 3, pp. 221–233. <https://doi.org/10.1037/h0057532>
24. Harris J.L., Rogers W.A., Qualls C.D. (1998) Written Language Comprehension in Younger and Older Adults. *Journal of Speech Language and Hearing Research*, vol. 41, no 3, pp. 603–616. <https://doi.org/10.1044/jslhr.4103.603>
25. Hasbrouck J.E., Tindal G.A. (2006) Oral Reading Fluency Norms: A Valuable Assessment Tool for Reading Teachers. *The Reading Teacher*, vol. 59, iss. 7, pp. 636–644. <https://doi.org/10.1598/RT.59.7.3>
26. Hasbrouck J.E., Tindal G.A. (1992) Curriculum-Based Oral Reading Fluency Norms for Students in Grades 2 through 5. *TEACHING Exceptional Children*, vol. 24, iss. 3, pp. 41–44. <https://doi.org/10.1177/004005999202400310>

27. Henning G. (1987) *A Guide to Language Testing: Development, Evaluation, Research*. Cambridge, MA: Newberry House.
28. Hoover W.A., Gough P.B. (1990) The Simple View of Reading. *Reading and Writing*, vol. 2, no 2, pp. 127–160. <https://doi.org/10.1007/bf00401799>
29. Ivanova M.V., Akinina Yu.S., Soloukhina O.A., Iskra E.V., Buivolova O.V., Chrabaszcz A.V., Stupina E.A., Khudyakova M.V., Akhutina T.V., Dragoy O.V. (2021) The Russian Aphasia Test: The First Comprehensive, Quantitative, Standardized, and Computerized Aphasia Language Battery in Russian. *PLOS One*, vol. 16, no 11, Article no e0258946. <https://doi.org/10.1371/journal.pone.0258946>
30. Johann V.E., Könen T., Karbach J. (2020) The Unique Contribution of Working Memory, Inhibition, Cognitive Flexibility, and Intelligence to Reading Comprehension and Reading Speed. *Child Neuropsychology*, vol. 26, no 3, pp. 324–344. <https://doi.org/10.1080/09297049.2019.1649381>
31. Keenan J.M., Betjemann R.S., Olson R.K. (2008) Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. *Scientific Studies of Reading*, vol. 12, no 3, pp. 281–300. <https://doi.org/10.1080/10888430802132279>
32. Kincaid J.P., Fishburne R.P., Rogers R.L., Chissom B.S. (1975) *Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. Research Branch Report no 8–75. Orlando, FL: Institute for Simulation and Training. <https://doi.org/10.21236/ADA006655>
33. 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*, vol. 16, no 1–2, pp. 262–284. <https://doi.org/10.1080/09541440340000213>
34. Klingner J.K. (2004) Assessing Reading Comprehension. *Assessment for Effective Intervention*, vol. 29, no 4, pp. 59–70. <https://doi.org/10.1177/073724770402900408>
35. Konur O. (2002) Assessment of Disabled Students in Higher Education: Current Public Policy Issues. *Assessment & Evaluation in Higher Education*, vol. 27, no 2, pp. 131–152. <https://doi.org/10.1080/02602930220128715>
36. 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. *Behaviour Research Methods*, vol. 51, iss. 3, pp. 1161–1178. <https://doi.org/10.3758/s13428-018-1051-6>
37. Magliano J., Millis K., Ozuru Y., McNamara D. (2007) A Multidimensional Framework to Evaluate Reading Assessment Tools. *Reading Comprehension Strategies: Theories, Interventions, and Technologies* (ed. D.S. McNamara), New York, NY: Psychology Press, pp. 107–136. <https://doi.org/10.4324/9780203810033>
38. McArthur G., Jones K., Anandakumar T., Larsen L., Castles A., Coltheart M. (2013) A Test of Everyday Reading Comprehension (TERC). *Australian Journal of Learning Difficulties*, vol. 18, no 1, pp. 35–85. <https://doi.org/10.1080/19404158.2013.779588>
39. Nelson H.E., Willison J.R. (1991) *The Revised National Adult Reading Test – Test Manual*. Windsor, UK: NFER-Nelson.
40. Nicolson R.I., Fawcett A.J. (2008) *Dyslexia, Learning, and the Brain*. Cambridge, MA: MIT. <https://doi.org/10.7551/mitpress/9780262140997.001.0001>
41. Ottosen H.F., Bønnerup K.H., Weed E., Parrila R. (2022) Identifying Dyslexia at the University: Assessing Phonological Coding Is Not Enough. *Annals of Dyslexia*, vol. 72, no 1, pp. 147–170. <https://doi.org/10.1007/s11881-021-00247-9>
42. Phillips S.M., Kelly K., Symes L. (2013) *Assessment of Learners with Dyslexic-Type Difficulties*. Los Angeles, CA: Sage.
43. Poulsen M., Gravgaard A.K.D. (2016) Who Did What to Whom? The Relationship between Syntactic Aspects of Sentence Comprehension and Text Comprehension.

- sion. *Scientific Studies of Reading*, vol. 20, no 4, pp. 325–338. <https://doi.org/10.1080/10888438.2016.1180695>
44. Şahin A. (2013) The Effect of Text Types on Reading Comprehension. *Mevlana International Journal of Education*, vol. 3, no 2, pp. 57–67. <http://dx.doi.org/10.13054/mije.13.27.3.2>
45. Share D.L. (2021) Common Misconceptions about the Phonological Deficit Theory of Dyslexia. *Brain Sciences*, vol. 11, no 11, Article no 1510. <https://doi.org/10.3390/brainsci11111510>
46. Sharma B., Jain R. (2014) Right Choice of a Method for Determination of Cut-Off Values: A Statistical Tool for a Diagnostic Test. *Asian Journal of Medical Sciences*, vol. 5, no 3, pp. 30–34. <https://doi.org/10.3126/ajms.v5i3.9296>
47. Slioussar N., Harchevnik M. (2024) Word Order and Context in Sentence Processing: Evidence from L1 and L2 Russian. *Frontiers in Psychology*, vol. 15, March, Article no 1344366. <https://doi.org/10.3389/fpsyg.2024.1344366>
48. Slioussar N., Makarchuk I. (2022) SOV in Russian: A Corpus Study. *Journal of Slavic Linguistics*, vol. 30, no 29, pp. 1–14. <https://doi.org/10.1353/jsl.2022.a923076>
49. Spector J.E. (1992) Predicting Progress in Beginning Reading: Dynamic Assessment of Phonemic Awareness. *Journal of Educational Psychology*, vol. 84, no 3, pp. 353–363. <https://doi.org/10.1037/0022-0663.84.3.353>
50. Tops W., Jansen D., Ceulemans E., Petry K., Hilton N.H., Baeyens D. (2022) Participation Problems and Effective Accommodations in Students with Dyslexia in Higher Education. *European Journal of Special Needs Education*, vol. 38, no 3, pp. 317–333. <https://doi.org/10.1080/08856257.2022.2089507>
51. Trauzettel-Klosinski S., Dietz K. (2012) Standardized Assessment of Reading Performance: The New International Reading Speed Texts IReST. *Investigative Ophthalmology & Visual Science*, vol. 53, no 9, pp. 5452–5461. <https://doi.org/10.1167/iops.11-8284>
52. Uysal P.K., Bilge H. (2019) An Investigation on the Relationship between Reading Fluency and Level of Reading Comprehension according to the Type of Texts. *International Electronic Journal of Elementary Education*, vol. 11, no 2, pp. 161–172. <https://doi.org/10.26822/iejee.2019248590>
53. Weiner I.B., Graham J.R., Naglieri J.A. (eds) (2012) *Handbook of Psychology. Vol. 10. Assessment Psychology*. Hoboken, NJ: John Wiley & Sons.
54. Wiederholt J.L., Bryant B.R. (2001) *GORT 4 Gray Oral Reading Tests Examiner's Manual*. Austin, TX: Pro-Ed.

References

- Alexeeva S., Slioussar N., Chernova D. (2018) StimulStat: A Lexical Database for Russian. *Behavior Research Methods*, vol. 50, no 6, pp. 2305–2315. <https://doi.org/10.3758/s13428-017-0994-3>
- Altpeter E.K., Marx T., Nguyen N.X., Naumann A., Trauzettel-Klosinski S. (2015) Measurement of Reading Speed with Standardized Texts: A Comparison of Single Sentences and Paragraphs. *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 253, no 8, pp. 1369–1375. <https://doi.org/10.1007/s00417-015-3065-4>
- Burdyna A.V., Koltuntseva V.E., Nikiforova N.V., Marchenko E.N., Sergeeva O.A., Alekseeva S.V. (2023) Assessment of Word and Nonword Reading Skills in Russian-Speaking Adolescents Age 14–17 Years. *Proceedings of the Conference "Cognitive Science in Moscow: New Researches" (Moscow, 2023, 21–22 June)*, pp. 558–563 (In Russian).
- Catts H.W., Kamhi A.G. (1999) *Language and Reading Disabilities*. Boston: Allyn and Bacon.
- Cavalli E., Colé P., Leloup G., Poracchia-George F., Sprenger-Charolles L., El Ahmadi A. (2017) Screening for Dyslexia in French-Speaking University Students: An Evaluation of the Detection Accuracy of the Alouette

- Test. *Journal of Learning Disabilities*, vol. 51, no 3, pp. 268–282. <https://doi.org/10.1177/0022219417704637>
- Cavalli E., Colé P., Leloup G., Poracchia-George F., Sprenger-Charolles L., El Ahmadi A. (2017) Screening for Dyslexia in French-Speaking University Students: An Evaluation of the Detection Accuracy of the Alouette Test. *Journal of Learning Disabilities*, vol. 51, no 3, pp. 268–282. <https://doi.org/10.1177/0022219417704637>
- Cline F., Johnstone C.J., King T.K. (2006) *Focus Group Reactions to Three Definitions of Reading (As Originally Developed in Support of NARAP Goal 1)*. Minneapolis, MN: National Accessible Reading Assessment Projects.
- Desrochers A., Glickman V. (2009) Criteria for the Evaluation of Reading Assessment Tools. *Encyclopedia of Language and Literacy Development*. London, ON: Canadian Language and Literacy Research Network, pp. 1–9.
- Dobrego A.S., Petrova T.E. (2016) Perception Patterns of Static and Dynamic Texts: An Experimental Study of Russian. *Aktualnye problemy yazykoznaniya*, vol. 1, pp. 33–34 (In Russian).
- Dorofeeva S.V., Grin'ko I.Ju., Perevoshchikova T.D., Dragoy O.V. (2021) Constructing of Word and Non-Word Reading Tests for Assessment of Reading Ability of Russian-Speaking Children. Proceedings of the Conference "Cognitive Science in Moscow: New Researches" (Moscow, 2021, 23–24 June), pp. 502–507 (In Russian).
- Dorofeeva S. V., Reshetnikova V., Serebryakova M., Goranskaya D., Akhutina T.V., Dragoy O. (2019) Assessing the Validity of the Standardized Assessment of Reading Skills in Russian and Verifying the Relevance of Available Normative Data. *The Russian Journal of Cognitive Science*, vol. 6, no 1, pp. 4–24.
- Driscoll D.P. (2012) *Reading Framework for the 2013. National Assessment of Educational Process*. Available at: <https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/reading/2013-reading-framework.pdf> (accessed 05.05.2025)
- Filippov K.A. (2003) *Linguistics of Text*. Saint-Petersburg: St. Petersburg University (In Russian).
- Flesch R. (1948) A New Readability Yardstick. *Journal of Applied Psychology*, vol. 32, no 3, pp. 221–233. <https://doi.org/10.1037/h0057532>
- Harris J.L., Rogers W.A., Qualls C.D. (1998) Written Language Comprehension in Younger and Older Adults. *Journal of Speech Language and Hearing Research*, vol. 41, no 3, pp. 603–616. <https://doi.org/10.1044/jslhr.4103.603>
- Hasbrouck J.E., Tindal G.A. (2006) Oral Reading Fluency Norms: A Valuable Assessment Tool for Reading Teachers. *The Reading Teacher*, vol. 59, iss. 7, pp. 636–644. <https://doi.org/10.1598/RT.59.7.3>
- Hasbrouck J.E., Tindal G.A. (1992) Curriculum-Based Oral Reading Fluency Norms for Students in Grades 2 through 5. *TEACHING Exceptional Children*, vol. 24, iss. 3, pp. 41–44. <https://doi.org/10.1177/004005999202400310>
- Henning G. (1987) *A Guide to Language Testing: Development, Evaluation, Research*. Cambridge, MA: Newberry House.
- Hoover W.A., Gough P.B. (1990) The Simple View of Reading. *Reading and Writing*, vol. 2, no 2, pp. 127–160. <https://doi.org/10.1007/bf00401799>
- Ivanova M.V., Akinina Yu.S., Soloukhina O.A., Iskra E.V., Buivolova O.V., Chrabaszcz A.V., Stupina E.A., Khudyakova M.V., Akhutina T.V., Dragoy O.V. (2021) The Russian Aphasia Test: The First Comprehensive, Quantitative, Standardized, and Computerized Aphasia Language Battery in Russian. *PLOS One*, vol. 16, no 11, Article no e0258946. <https://doi.org/10.1371/journal.pone.0258946>
- Johann V.E., Könen T., Karbach J. (2020) The Unique Contribution of Working Memory, Inhibition, Cognitive Flexibility, and Intelligence to Reading Comprehension and Reading Speed. *Child Neuropsychology*, vol. 26, no 3, pp. 324–344. <https://doi.org/10.1080/09297049.2019.1649381>

- Keenan J.M., Betjemann R.S., Olson R.K. (2008) Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. *Scientific Studies of Reading*, vol. 12, no 3, pp. 281–300. <https://doi.org/10.1080/10888430802132279>
- Khamaganova V.M. (2002) *Structural, Semantic and Lexical Model of a Text of the "Description" Type: Problems of Semiotics and Ontology* (PhD Thesis). Moscow: Moscow Pedagogical State University (In Russian).
- Kincaid J.P., Fishburne R.P., Rogers R.L., Chissom B.S. (1975) *Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Research Branch Report no 8–75*. Orlando, FL: Institute for Simulation and Training. <https://doi.org/10.21236/ADA006655>
- 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*, vol. 16, no 1–2, pp. 262–284. <https://doi.org/10.1080/09541440340000213>
- Klingner J.K. (2004) Assessing Reading Comprehension. *Assessment for Effective Intervention*, vol. 29, no 4, pp. 59–70. <https://doi.org/10.1177/073724770402900408>
- Konur O. (2002) Assessment of Disabled Students in Higher Education: Current Public Policy Issues. *Assessment & Evaluation in Higher Education*, vol. 27, no 2, pp. 131–152. <https://doi.org/10.1080/02602930220128715>
- Kornev A.N. (1997) *Reading and Writing Disorders in Children*. Saint-Petersburg: M&M (In Russian).
- Kornev A.N., Ishimova O.A. (2010) *Method of Diagnostic of Dyslexia in Children*. Saint-Petersburg: Peter the Great St. Petersburg Polytechnic University (In Russian).
- Laposhina A.N., Lebedeva M.Ju. (2021) Textometr: An Online Tool for Automated Complexity Level Assessment of Texts for Russian Language Learners. *Russian Language Studies*, vol. 19, no 3, pp. 331–345 (In Russian). <https://doi.org/10.22363/2618-8163-2021-19-3-331-345>
- Laposhina A.N., Lebedeva M.Ju., Berlin Khenis A.A. (2022) Word Frequency and Text Complexity: An Eye-Tracking Study of Young Russian Readers. *Russian Journal of Linguistics*, vol. 26, no 2, pp. 493–514. <https://doi.org/10.22363/2687-0088-30084>
- 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. *Behaviour Research Methods*, vol. 51, iss. 3, pp. 1161–1178. <https://doi.org/10.3758/s13428-018-1051-6>
- Lezina A.D., Nikonova Ya.E., Cherevik A.S., Shatkovskaya I.G., Alekseeva S.V. (2023) Reading Aloud in Russian Language as Tool for Assessment of Risk of Dyslexia in Adolescents Age 14–17. Proceedings of the Conference “Cognitive Science in Moscow: New Researches” (Moscow, 2023, 21–22 June), pp. 599–603 (In Russian).
- Magliano J., Millis K., Ozuru Y., McNamara D. (2007) A Multidimensional Framework to Evaluate Reading Assessment Tools. *Reading Comprehension Strategies: Theories, Interventions, and Technologies* (ed. D.S. McNamara), New York, NY: Psychology Press, pp. 107–136. <https://doi.org/10.4324/9780203810033>
- Mastitskiy S.E., Shitikov V.K. (2015) *Statistical Analysis and Data Visualization via R*. Moscow: DMK (In Russian).
- McArthur G., Jones K., Anandakumar T., Larsen L., Castles A., Coltheart, M. (2013) A Test of Everyday Reading Comprehension (TERC). *Australian Journal of Learning Difficulties*, vol. 18, no 1, pp. 35–85. <https://doi.org/10.1080/19404158.2013.779588>
- Murzin L.N., Štern A.S. (1991) *Text and Its Perception*. Sverdlovsk: Ural University (In Russian).
- Nelson H.E., Willison J.R. (1991) *The Revised National Adult Reading Test — Test Manual*. Windsor, UK: NFER-Nelson.

- Nicolson R.I., Fawcett A.J. (2008) *Dyslexia, Learning, and the Brain*. Cambridge, MA: MIT. <https://doi.org/10.7551/mitpress/9780262140997.001.0001>
- Ottosen H.F., Bønnerup K.H., Weed E., Parrila R. (2022) Identifying Dyslexia at the University: Assessing Phonological Coding Is Not Enough. *Annals of Dyslexia*, vol. 72, no 1, pp. 147–170. <https://doi.org/10.1007/s11881-021-00247-9>
- Phillips S.M., Kelly K., Symes L. (2013) *Assessment of Learners with Dyslexic-Type Difficulties*. Los Angeles, CA: Sage.
- Poulsen M., Gravgaard A.K.D. (2016) Who Did What to Whom? The Relationship between Syntactic Aspects of Sentence Comprehension and Text Comprehension. *Scientific Studies of Reading*, vol. 20, no 4, pp. 325–338. <https://doi.org/10.1080/10888438.2016.1180695>
- Şahi n A. (2013) The Effect of Text Types on Reading Comprehension. *Mevlana International Journal of Education*, vol. 3, no 2, pp. 57–67. <http://dx.doi.org/10.13054/mije.13.27.3.2>
- Share D.L. (2021) Common Misconceptions about the Phonological Deficit Theory of Dyslexia. *Brain Sciences*, vol. 11, no 11, Article no 1510. <https://doi.org/10.3390/brainsci11111510>
- Sharma B., Jain R. (2014) Right Choice of a Method for Determination of Cut-Off Values: A Statistical Tool for a Diagnostic Test. *Asian Journal of Medical Sciences*, vol. 5, no 3, pp. 30–34. <https://doi.org/10.3126/ajms.v5i3.9296>
- Slioussar N., Harchevnik M. (2024) Word Order and Context in Sentence Processing: Evidence from L1 and L2 Russian. *Frontiers in Psychology*, vol. 15, March, Article no 1344366. <https://doi.org/10.3389/fpsyg.2024.1344366>
- Slioussar N., Makarchuk I. (2022) SOV in Russian: A Corpus Study. *Journal of Slavic Linguistics*, vol. 30, no 29, pp. 1–14. <https://doi.org/10.1353/jsl.2022.a923076>
- Spector J.E. (1992) Predicting Progress in Beginning Reading: Dynamic Assessment of Phonemic Awareness. *Journal of Educational Psychology*, vol. 84, no 3, pp. 353–363. doi:10.1037/0022-0663.84.3.353
- Tops W., Jansen D., Ceulemans E., Petry K., Hilton N. H., Baeyens D. (2022) Participation Problems and Effective Accommodations in Students with Dyslexia in Higher Education. *European Journal of Special Needs Education*, vol. 38, no 3, pp. 317–333. <https://doi.org/10.1080/08856257.2022.2089507>
- Trauzettel-Klosinski S., Dietz K. (2012) Standardized Assessment of Reading Performance: The New International Reading Speed Texts IReST. *Investigative Ophthalmology & Visual Science*, vol. 53, no 9, pp. 5452–5461. <https://doi.org/10.1167/iops.11-8284>
- Uysal P.K., Bilge H. (2019) An Investigation on the Relationship between Reading Fluency and Level of Reading Comprehension according to the Type of Texts. *International Electronic Journal of Elementary Education*, vol. 11, no 2, pp. 161–172. <https://doi.org/10.26822/iejee.2019248590>
- Valgina N.S. (2003) *Theory of Text*. Moscow: Logos (In Russian).
- Weiner I.B., Graham J.R., Naglieri J.A. (eds) (2012) *Handbook of Psychology. Vol. 10. Assessment Psychology*. Hoboken, NJ: John Wiley & Sons.
- Wiederholt J.L., Bryant B.R. (2001) *GORT 4 Gray Oral Reading Tests Examiner's Manual*. Austin, TX: Pro-Ed.
- Zdorova N.S., Staroverova V.N., Lopukhina A.A., Dzhonboboieva D.B., Pareshina E.A., Shestakova E.R., Gmyrina V.A., Dmitrova E.P., Dragoy O.V. (2025) Reading Assessment Tool “LexiMetr” for Russian-Speaking Primary School Students. *Vo-prosy obrazovaniya / Educational Studies Moscow*, no 2, pp. 73–100 (In Russian). <https://doi.org/10.17323/vo-2025-21693>