

Зубов Владислав Иванович

Санкт-Петербургский государственный университет,
Россия, 199034, Санкт-Петербург, Университетская наб., 7–9

Национальный исследовательский университет «Высшая школа экономики»,
Россия, 101000, Москва, ул. Мясницкая, 20

<https://orcid.org/0000-0001-8714-7491>

vladzubov21@gmail.com

Ладнова Анастасия Алексеевна

Санкт-Петербургский государственный университет,
Россия, 199034, Санкт-Петербург, Университетская наб., 7–9

<https://orcid.org/0000-0003-0132-7016>

a.ladnova@spbu.ru

Понимание текста на родном и неродном языках: исследование с использованием методики выделения ключевых слов*

Для цитирования: Зубов В. И., Ладнова А. А. Понимание текста на родном и неродном языках: исследование с использованием методики выделения ключевых слов. *Вестник Санкт-Петербургского университета. Язык и литература.* 2025, 22 (3): 664–683. EDN XKRVES

В исследовании предпринята попытка применить методику выделения ключевых слов для анализа компетенции понимания текста на родном и неродном языках. Предполагалось, что наборы ключевых слов, выделяемые участниками, отражают уровень их понимания текста. Индивидуальные наборы ключевых слов сопоставлялись с истинными наборами, сформированными на основе слов, выделенных более чем 20 % участников. Успешность выделения ключевых слов оценивалась по количеству совпадений с истинным набором. Проверялись гипотезы о совпадении истинных наборов ключевых слов для текстов на разных языках и о связи успешности выделения ключевых слов с успешностью ответов на вопросы по содержанию текста. Были проведены три эксперимента, в которых участвовали носители русского языка, читавшие тексты на русском и английском языках, а также носители китайского, работавшие с текстами на китайском и русском языках. Во всех трех экспериментах была выявлена положительная корреляция между успешностью выделения ключевых слов и успешностью ответов на вопросы на неродном языке, что подтверждает гипотезу о том, что оба задания оценивают степень понимания текста. Предположение о совпадении истинных наборов ключевых слов для одного и того же текста на разных языках подтвердилось как для русскоязычных, так и для китайскоязычных участников. Успешность ответов на вопросы и уровень владения иностранным языком оказались связаны для носителей обоих языков, однако успешность выделения ключевых слов не зависела от уровня языковой компетенции. Это позволяет предположить, что навык выделения ключевых

* Работа выполнена при поддержке СПбГУ «Механизмы чтения и интерпретации текста на родном и неродном языках: междисциплинарное экспериментальное исследование с использованием методов регистрации движения глаз, визуальной аналитики и технологий виртуальной реальности», шифр проекта 124032900009-2.

слов менее связан с языковой компетенцией и может использоваться для оценки понимания текста без учета лингвистических факторов. Результаты исследования подтверждают эффективность методики выделения ключевых слов для оценки понимания текста и подчеркивают различия с предыдущими работами. Перспективой является дальнейшее сопоставление результатов в зависимости от используемых методик.

Ключевые слова: чтение, понимание, методика выделения ключевых слов, второй язык.

Подходы к описанию процесса понимания текста

Понимание текста представляет собой сложный многоуровневый процесс, на который влияют различные факторы. Среди них можно выделить уровень языковой компетенции читателя (например, беглость чтения и морфологическое осознание), мотивацию, а также когнитивные способности (такие как рабочая память и умение контролировать процесс чтения) [Lervåg, Melby-Lervåg 2022]. Существуют разные уровни понимания текста: от поверхностного, включающего способность уловить основную идею, тему и ключевую информацию, до более глубокого, при котором читатель может понять намерения автора текста [Базылев, Красильникова 2011]. Обычно для оценки понимания текста используется вопросно-ответная методика. Хотя этот метод широко применяется, у него есть свои недостатки. Умение отвечать на вопросы по тексту зависит не только от способности понимать текст, но и от других навыков, таких как языковая компетенция, уровень интеллекта и общие знания. Исследования показывают, что некоторые читатели могут успешно справляться со стандартизованными вопросами к текстам, не обязательно основываясь на глубоком понимании текста, а используя другие стратегии или подсказки, встроенные в сами вопросы [Nation 2022: 260].

В отечественной психолингвистике для исследования процессов восприятия и понимания текста широко применяется метод выделения ключевых слов (КС), как, например, в статье [Петрова и др. 2017]. Этот метод был подробно описан в работе [Мурзин, Штерн 1991]: участникам эксперимента предлагают выделить КС из текста, при этом каждый составляет свой собственный набор КС. Некоторые КС будут совпадать, а другие различаться, что объясняется как общим пониманием текста, так и индивидуальными различиями в восприятии содержания и задачи по выделению КС [Мурзин, Штерн 1991: 74]. Наиболее частотные КС отображают «общее в восприятии текста — его “инвариантный смысл”» (Л. А. Черняховская, цит. по: [Мурзин, Штерн 1991: 77]). Выделенные КС используются как инструмент оценки понимания текста школьниками, а также как опора для последующего пересказа текста [Грудева, Губушкина 2020]. Анализируя КС, выделенные школьниками на основе прослушанного текста, авторы приходят к выводу о том, что КС отражают степень понимания участниками прослушанного текста.

В зарубежных исследованиях метод выделения КС и автоматически выделенные наборы КС чаще всего применяются в смежных с лингвистикой областях, таких как интеллектуальный анализ текста, информационный поиск и обработка естественного языка [Beliga et al. 2015: 1]; так, наборы КС используются в качестве метаданных для улучшения содержания документов, упрощения их классификации, систематизации, индексации и создания кратких обзоров текстовой информации, что облегчает поиск и настройку рекомендаций для пользователей [Firoozeh et al. 2020: 260]. Однако имеются исследования, где метод выделения КС применяется

для анализа понимания текста, напр.: [de Bruin et al. 2011; Engelen et al. 2018]. В работе [Engelen et al. 2018] было установлено, что успешность выделения ключевых слов не коррелировала с количеством правильных ответов на вопросы по содержанию текста. Экспертный анализ КС, выбранных учениками после прочтения текста, показал, что успешность в выделении КС не предсказывала точность ответов на вопросы, но было высказано предположение, что КС все же могут отражать уровень понимания текста и требуются дальнейшие исследования в этом направлении.

Поскольку набор КС к тексту является его смысловым ядром, «отражает инвариантный смысл» [Мурзин, Штерн 1991: 77], можно предположить, что для текстов с одинаковым содержанием, но написанных на разных языках, читатели будут выделять схожие или даже одинаковые наборы КС. Настоящее исследование представляет собой попытку применить методику выделения КС к исследованию компетенции понимания текста на родном и неродном языках, а также сравнить эту методику с вопросно-ответной методикой. Мы видим потенциал данной методики при оценке уровня понимания текста как на родном, так и на неродном языке, поскольку, как отмечают авторы методики Л. В. Сахарный и А. С. Штерн, простота процедуры выделения КС делает легким проведение эксперимента с применением этой методики и обработку данных, а также дает возможность выдвижения «ряда гипотез о более общих механизмах построения и осознания текстов» [Сахарный, Штерн 2006: 54].

В связи с этим в настоящем исследовании проверялись следующие гипотезы:

- 1) истинные наборы КС для текстов на разных языках совпадают, поскольку инвариантный смысл текста нелингвоспецифичен;
- 2) успешность выделения КС связана с успешностью ответов на вопросы по содержанию текста, поскольку оба задания нацелены на оценку компетенции «понимание текста».

Для проверки этих гипотез было проведено три лингвистических эксперимента с применением методики выделения КС в текстах на трех типологически различных языках — русском, английском и китайском. В экспериментах участвовали носители русского и китайского языков. Носители русского языка выделяли КС в текстах на русском и английском языках, носители китайского языка — в текстах на русском и китайском языках.

Экспериментальное исследование понимания текста носителями русского языка с помощью метода выделения КС (эксперименты 1 и 2)

Материалы исследования

Материалами исследования стали четыре пары прозаических текстов научно-популярного стиля на русском и английском языках: в первом тексте рассказывается о тесте Тьюринга («Алан Тьюринг»), во втором — о том, как животные способны предчувствовать землетрясения («Землетрясение»), третий посвящен значению и происхождению жеста «шака» («Шака»), четвертый повествует о древнеримском боге Янусе («Янус»).

Тексты «Алан Тьюринг» и «Землетрясение» использовались в эксперименте 1, целью которого было проверить экспериментальную процедуру и предварительно

протестировать исследовательские гипотезы. Тексты были взяты из региональных этапов Всероссийской олимпиады школьников¹ за 9-й класс 2010 и 2016 гг., переведены на русский язык с помощью переводчика DeepL², после чего перевод был отредактирован вручную. Для каждого текста были составлены пять вопросов по содержанию. После обработки результатов эксперимента 1 было обнаружено, что участники субъективно оценили тексты как достаточно простые. Текст «Землетрясение» оказался значимо более легким с точки зрения успешности ответов на вопросы по содержанию, поэтому в эксперименте 2 решено было использовать более сложные тексты, уже протестированные в изучении процессов чтения на разных языках. Тексты «Янус» и «Шака» и ряд вопросов к ним были заимствованы из исследования [Kuperman et al. 2022]. К четырем разработанным в исследовании [Kuperman et al. 2022] вопросам были добавлены еще четыре, нацеленные на оценку понимания разных уровней глубины; таким образом, для каждого текста было подготовлено по восемь вопросов.

В табл. 1 представлены параметры текстов: количество слов, количество предложений и два индекса читабельности — индекс Р.Флеша (Flesch — Kincaid) и индекс SMOG. Читабельность — это мера того, насколько легко текст воспринимается и понимается читателем. Индексы рассчитываются автоматически на основе длины слов и предложений, что позволяет количественно оценить сложность текста. Методики, лежащие в основе Flesch — Kincaid и SMOG, являются валидированными и широко используемыми в научной и прикладной практике, а их надежность подтверждена в ряде эмпирических исследований (см., напр.: [Friedman, Hoffman-Goetz 2006: 369–370]). Читабельность текстов на русском языке была определена с помощью сервиса «Оценка понятности текста»³, на английском языке — с помощью сервиса «Readable»⁴.

Таблица 1. Параметры текстов, использованных в качестве материала для первого этапа исследования

Текст	Количество слов	Количество предложений	Индекс Flesch — Kincaid	Индекс SMOG
«Янус»	149	9	8,98	9,73
«Janus»	182	10	10,40	12,70
«Шака»	142	7	14,51	13,42
«Shaka»	183	6	14,00	15,90
«Алан Тьюринг»	153	12	9,56	8,51
«Alan Turing»	214	12	8,60	8,40
«Землетрясение»	157	12	13,14	12,02
«Earthquake»	182	12	9,80	12,20

¹ Информационный сайт об олимпиадах и других мероприятиях для школьников. <https://olimpiada.ru> (дата обращения: 23.10.2023).

² Онлайн-сервис для перевода DeepL. <https://deepl.com/ru/translator> (дата обращения: 23.10.2023).

³ Инструмент оценки понятности текстов на русском языке. <https://plainrussian.ru> (дата обращения: 23.10.2023).

⁴ Инструмент оценки понятности текстов на английском языке. <https://readable.com> (дата обращения: 23.10.2023).

Все стимульные материалы и результаты исследования доступны на платформе Open Science Framework (OSF)⁵.

Участники исследования. К участию в исследовании привлекались добровольцы, для которых русский язык был родным (L1), а английским — изучаемым (L2). Все участники перед исследованием заполнили анкету, в которой указали свой пол, возраст, а также субъективно оценили уровень владения английским языком по шкале CEFR. В исследование включались результаты респондентов, которые указали уровень владения английским не ниже B1 по шкале CEFR. Всего было привлечено 247 участников.

В эксперименте 1135 участников (88 женщин и 47 мужчин в возрасте от 12 до 68 лет, средний возраст — 21 год) читали пары текстов «Алан Тьюринг» и «Earthquake» или «Alan Turing» и «Землетрясение»⁶. Распределение участников по уровням владения английским языком следующее: B1 — 71 участник, B2 — 46 участников, C1 — 14 участников, C2 — 4 участника.

В эксперименте 2112 участников (92 женщины и 20 мужчин в возрасте от 18 до 59 лет, средний возраст — 23 года) читали пары текстов «Янус» и «Shaka» и «Janus» и «Шака». Все участники указали уровень владения английским языком B2—C2 по шкале CEFR (B2 — 66 участников, C1 — 38 участников, C2 — 7 участников; один участник не указал уровень). Кроме того, участники дополнительно сообщили о количестве лет, в течение которых они изучали английский язык ($M = 13,14$).

Методика и ход экспериментов с носителями русского языка. Каждому участнику предлагалось прочитать два текста на разные темы (один на родном языке (L1), другой на неродном языке (L2)), оценить его сложность по шкале от 3 до −3, где 3 — очень простой, а −3 — очень сложный, выписать из текстов по десять КС и ответить на вопросы по содержанию. КС было определено как «наиболее важное с точки зрения содержания текста слово». В эксперименте 1 для текстов «Алан Тьюринг» и «Землетрясение» было пять вопросов по содержанию: три вопроса с выбором нескольких вариантов ответа и два вопроса с выбором ответа «да» или «нет». В эксперименте 2 для текстов «Янус» и «Шака» было восемь вопросов по содержанию: два вопроса с выбором нескольких вариантов ответа, два вопроса с развернутым ответом и четыре вопроса с выбором ответа «да» или «нет». Во время ответа на вопросы участники не могли возвращаться к тексту. Участники могли записывать ответы на любом языке. Все инструкции и вопросы были написаны на том же языке, на котором был написан текст. Порядок предъявления текстов был случайным.

Предварительная обработка данных

Обработка КС, которые выписали участники, в обоих экспериментах проводилась следующим образом. Различные формы слова сводились к форме, наиболее часто встречающейся в ответах (*ответ* (10), *ответа* (7), *ответы* (24) → *ответы* (41)), словосочетания, не являющиеся устойчивыми, делились на слова, входящие в них (*не почувствовать разницы* → *не почувствовать + разницы*; *предчувствие животных* → *предчувствие+ животных*); союзы, предлоги и вспомогательные гла-

⁵ Проект «Изучение понимания текста». Open Science Framework. osf.io/34d5w/ (дата обращения: 22.10.2024).

⁶ Язык названий текстов соответствует языку, на котором написаны тексты.

голы удалялись (*разница в общении* → *разница + общение*; *могут выглядеть растерянными* → *выглядеть + растерянными*); имена собственные не делились и приводились во всех ответах участников к полному варианту (*Алан, Тьюринг, Алан Тьюринг* → *Алан Тьюринг; Cambridge, Cambridge University* → *Cambridge University*).

В соответствии с инструкцией подсчета КС, описанной в работе [Мурзин, Штерн 1991: 74], для каждого КС, встретившегося в ответах участников, была рассчитана абсолютная (m) и относительная (p) частота встречаемости: $p = m/n$, где n — количество участников. Группа КС, получивших наибольшую относительную частоту встречаемости, составила истинный набор КС, в который были включены все КС, которые выписали как минимум 20 % участников. Для текста «Янус» на русском языке в истинный набор КС вошли 18 слов, на английском языке — 17 слов; для текста «Шака» на русском языке — 17 слов, на английском языке — 21 слово. Для текста «Алан Тьюринг» на обоих языках, а также для текста «Землетрясение» на русском языке в истинный набор КС вошли 14 слов. Для текста «Землетрясение» на английском языке в истинный набор КС вошли 15 слов. Все списки истинных наборов КС с указанием абсолютной частотности КС доступны по ссылке с материалами исследования, приведенной выше.

Далее для каждого участника рассчитывался показатель успешности выделения КС для каждого текста по шкале от 0 до 10. Если в индивидуальном наборе КС конкретного участника встречалось слово из истинного набора КС, данному КС участника присваивался один балл. Максимальное количество баллов, которое участники могли получить за успешность выделения КС за один текст, — 10. Это ограничение шкалы было введено для обеспечения единобразия оценивания и предотвращения завышения результата за счет формальных совпадений. В частности, некоторые участники вместо одного КС указывали словосочетания, не являющиеся устойчивыми или лексически цельными (*не почувствовать + разницы*), что позволяло охватить сразу два или более ключевых элемента. Ограничение в десять баллов исключает возможность необоснованного повышения оценки за счет дробления или расширенной формулировки ответов и делает систему оценивания сопоставимой для всех участников.

Если ни одно КС из набора конкретного участника не входило в истинный набор КС, такой участник получал 0 баллов за успешность выделения КС.

Успешность ответов на вопросы по содержанию текста определялась по количеству правильных ответов. Каждый правильный ответ оценивался одним баллом. При оценке вопросов, подразумевающих развернутый ответ, использовались ключи со всеми возможными вариантами ответа. Максимальное количество баллов, которое можно было получить за успешность ответов на вопросы по одному тексту, — 8 в эксперименте 1 или 5 в эксперименте 2.

Результаты

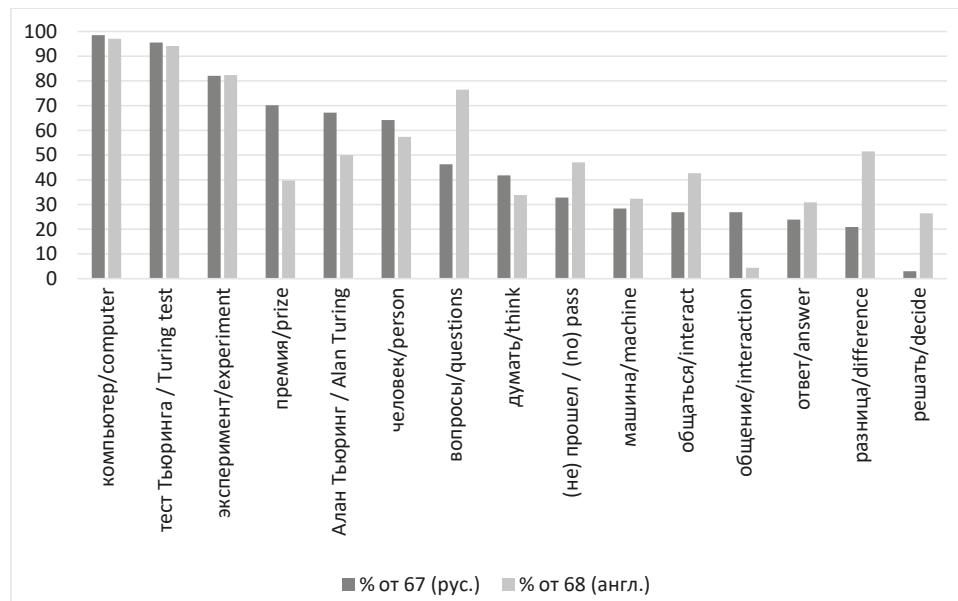
Во-первых, было проверено, насколько сложными для чтения оказались тексты. Тексты на иностранном языке оценивались респондентами как более сложные, чем тексты на родном языке (значение критерия *t*-Стьюдента $t = 4,86$, $p < 0,001$), средние оценки приведены в табл. 2. В то же время статистически значимой корреляции между субъективной оценкой сложности текста и индексами читабельности не об-

наружено (значение коэффициента корреляции r -Пирсона $r = -0,07$, $p = 0,304$ для индекса по формуле Flesch — Kincaid; $r = 0,059$, $p = 0,381$ для SMOG), что можно объяснить либо недостатками формул индекса читабельности для разных языков, либо тем, что сложность текста является более объемным понятием, чем читабельность, и не имеет прямого отношения к читабельности [Hiebert 2011: 39–40].

Таблица 2. Средняя оценка субъективной сложности текстов «Янус», «Шака», «Алан Тьюринг» и «Землетрясение» носителями русского языка

Текст	Язык текста	
	Русский	Английский
«Алан Тьюринг»	1,9	2,2
«Землетрясение»	1,7	2,4
«Янус»	2,4	3,0
«Шака»	2,2	2,9

Во-вторых, были сопоставлены истинные наборы КС на разных языках для четырех текстов. Для этого была сопоставлена абсолютная частота встречаемости одинаковых КС в текстах на разных языках (*жест* (55) — *gesture* (45) и т. п.) и была обнаружена значимая корреляция между истинными наборами КС во всех четырех парах текстов: «Алан Тьюринг» / «Alan Turing» — $r = 0,782$, $p < 0,001$; «Землетрясение» / «Earthquake» — $r = 0,856$, $p < 0,001$; «Янус» / «Janus» — $r = 0,605$, $p = 0,005$, «Шака» / «Shaka» — $r = 0,651$, $p = 0,001$. Гистограммы с сопоставлением истинных наборов КС для четырех пар текстов см. на рис. 1–4.



Rис. 1. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Алан Тьюринг» на русском и английском языках

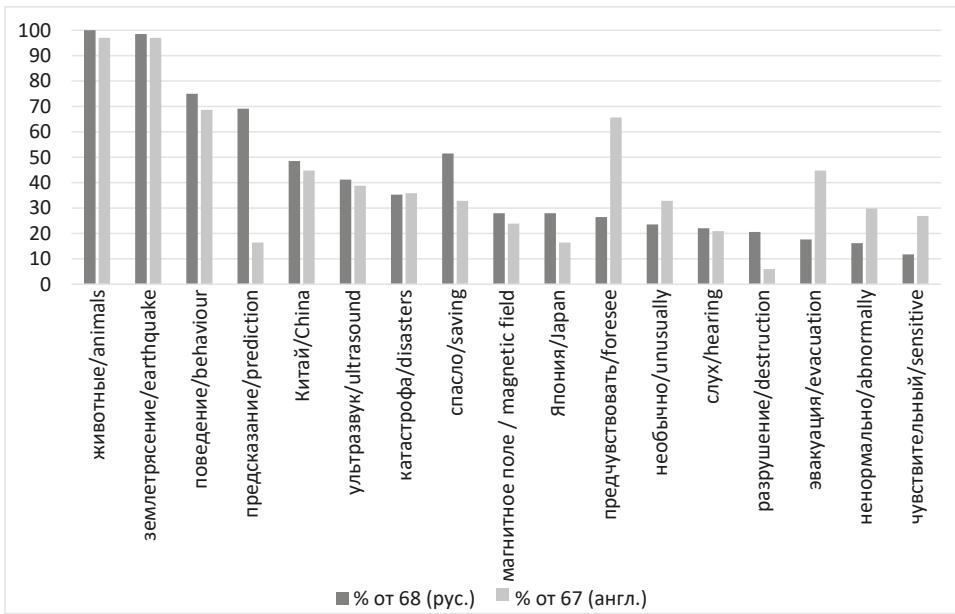


Рис. 2. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Землетрясение» на русском и английском языках

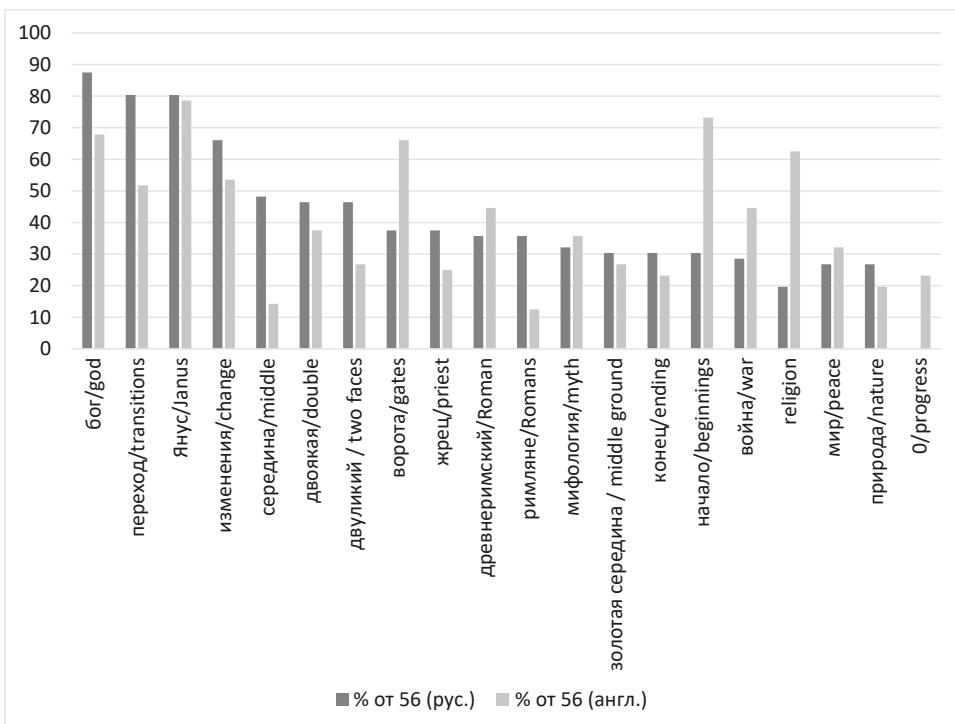


Рис. 3. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Янус» на русском и английском языках

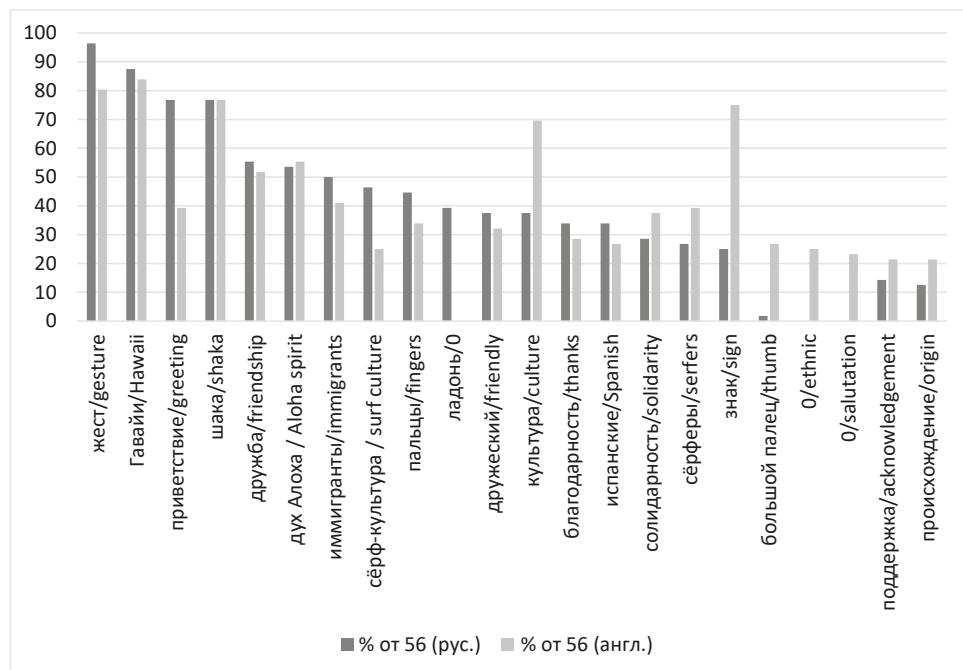


Рис. 4. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Шака» на русском и английском языках

В табл. 3 представлено количество респондентов, читавших тот или иной текст, количество выделенных КС и количество уникальных КС (так как участники могли выписать больше десяти КС для одного текста, поскольку иногда выписывали словосочетания (например, *peace and war* или *война и мир*), общее количество КС для текста может превышать предполагаемое максимальное число КС, ср. для текста «Алан Тьюринг» на русском языке 67 участников предложили 683 КС, несмотря на то что в задании максимально можно было предложить десять КС).

Таблица 3. Количество респондентов, выделенных респондентами КС, уникальных КС для текстов на русском и английском языках

Параметр	Текст							
	«Алан Тьюринг»		«Землетрясение»		«Янус»		«Шака»	
	Русский язык	Английский язык	Русский язык	Английский язык	Русский язык	Английский язык	Русский язык	Английский язык
Количество респондентов	67	68	68	67	56	56	56	56
Количество выделенных КС	683	690	704	737	634	640	613	647
Количество уникальных КС	80	72	83	84	63	68	62	65

Таблица 4. Результаты линейной регрессии: влияние предикторов на успешность выделения КС ключевых слов и ответы на вопросы на L2 (эксперимент 1)

Зависимая переменная	Предиктор	Оценка параметра	Стандартная ошибка	Значение t-критерия	p
Успешность выделения КС на L2 ($R^2 = 0,23$)	Свободный коэффициент	2,19	1,40	1,56	0,12
	Успешность выделения КС на L1	0,29	0,09	3,07	0,003*
	Ответы на вопросы на L1	0,17	0,10	1,62	0,11
	Ответы на вопросы на L2	0,17	0,08	1,99	0,05*
	Уровень языка	0,08	0,23	0,35	0,73
	Текст («Шака»)	1,09	0,28	3,92	<0,001*
Ответы на вопросы на L2 ($R^2 = 0,27$)	Свободный коэффициент	-1,04	1,65	-0,63	0,53
	Ответы на вопросы на L1	0,46	0,11	3,99	<0,001*
	Успешность выделения КС на L1	-0,08	0,11	-0,71	0,48
	Успешность выделения КС на L2	0,22	0,11	1,99	0,05*
	Уровень языка	0,89	0,26	3,44	<0,001*
	Текст («Шака»)	0,25	0,34	0,72	0,47

Примечание: эксперимент 1, «Янус» и «Шака», русский (L1) и английский (L2) языки, звездочкой отмечены статистически значимые предикторы.

Таблица 5. Результаты линейной регрессии: влияние предикторов на успешность выделения КС и ответы на вопросы на L2 (эксперимент 2)

Зависимая переменная	Предиктор	Оценка параметра	Стандартная ошибка	Значение t-критерия	p
Успешность выделения КС на L2 ($R^2 = 0,14$)	Свободный коэффициент	2,22	1,12	1,99	0,05*
	Успешность выделения КС на L1	0,25	0,10	2,66	0,01*
	Ответы на вопросы на L1	0,14	0,09	1,55	0,12
	Ответы на вопросы на L2	0,16	0,07	2,21	0,03*
	Уровень языка	0,25	0,16	1,50	0,14
	Текст («Землетрясение»)	-0,56	0,29	-1,92	0,06
Ответы на вопросы на L2 ($R^2 = 0,09$)	Свободный коэффициент	4,06	1,31	3,10	0,002*
	Ответы на вопросы на L1	0,19	0,11	1,79	0,08
	Успешность выделения КС на L1	0	0,12	0,04	0,97
	Успешность выделения КС на L2	0,23	0,10	2,21	0,03*
	Уровень языка	0,18	0,20	0,93	0,36
	Текст («Землетрясение»)	0,93	0,34	2,72	0,01*

Примечание: эксперимент 2, «Алан Тьюринг» и «Землетрясение», русский (L1) и английский (L2) языки, звездочкой отмечены статистически значимые предикторы.

Далее для оценки влияния ряда факторов на успешность выделения КС и ответов на вопросы по содержанию текста на L2 использовалась линейная регрессия. Анализ проводился в среде R (версия 4.3.2) с использованием стандартного паке-

та stats. В качестве зависимых переменных выступили успешность выделения КС и количество правильных ответов на вопросы на L2, в качестве предикторов (т.е. независимых переменных, объясняющих вариации в зависимых) — успешность выделения КС и ответов на вопросы на L1, уровень языка и текст (см. табл. 4–5).

При выделении КС на L2 значимый вклад в успешность выполнения задания оказывала успешность выделения КС на L1, количество верных ответов на L2, а также тип текста для пары текстов «Янус» и «Шака» (для текста «Шака» успешность выделения КС выше).

Для предсказания ответов на вопросы на L2 для обеих пар текстов значимым оказалось количество верно выделенных КС на L2. Для пары текстов «Янус» и «Шака» значимым предиктором также является количество верных ответов на вопросы на L1 и уровень языка участников (чем выше уровень языка участников, тем больше правильных ответов на вопросы они дают). Для пары текстов «Алан Тьюринг» и «Землетрясение» значимым предиктором оказался тип текста: участники значимо лучше отвечали на вопросы по содержанию текста «Землетрясение».

Обсуждение результатов экспериментов с носителями русского языка

В обоих экспериментах с носителями русского языка была обнаружена положительная корреляция между абсолютной частотой встречаемости КС, выделенных на родном языке, и абсолютной частотой встречаемости КС, выделенных на неродном, английском, языке. Подобный результат позволяет предположить, что истинные наборы КС будут отражать содержание текста, его смысловое ядро вне зависимости от языка, на котором написан текст.

В обоих экспериментах с носителями русского языка обнаружена положительная связь между успешностью выделения КС и правильностью ответов на вопросы по содержанию текста при чтении текста на неродном языке. Это подтверждает предположение о том, что индивидуальные наборы КС будут в определенной степени отражать степень понимания прочитанного текста, этот результат не совпадает с результатом исследования [Engelen et al. 2018].

Было также обнаружено, что успешность выделения КС на неродном языке связана с успешностью выделения КС на родном языке, что позволяет предположить наличие универсального навыка работы с содержанием текста вне зависимости от языка, на котором написан текст. Данный результат может служить подтверждением предположения о наличии универсальных стратегий работы со смыслом текста вне зависимости от формата последнего, см. обсуждение в работе [Петрова и др. 2017: 155]. Важно подчеркнуть и то, что успешность выделения КС в одном из экспериментов была связана с содержанием текста (для текста «Шака» участники выделяли КС более успешно, чем для текста «Янус»).

В эксперименте с текстами «Янус» и «Шака» уровень владения иностранным языком вносит вклад в то, насколько правильно участники отвечают на вопросы на неродном языке, чего не было обнаружено для КС, т.е. при выделении КС носителями русского языка в текстах на английском языке языковая компетенция роли не играет. Это позволяет сделать предположение о том, что, с одной стороны, традиционный метод оценки понимания текста с помощью ответов на вопросы по содержанию связан с языковой компетенцией читателя, что необходимо учитывать

при использовании этой методики оценки понимания прочитанного, с другой стороны, методика набора КС свободна от этого ограничения.

Связь между успешностью выделения КС и успешностью ответов на родном и неродном языках в обоих экспериментах позволяет сделать вывод о том, что методика выделения КС, предлагаемая в исследовании [Мурзин, Штерн 1991: 73–91], может быть использована как дополнительный способ оценки понимания прочитанного как на родном, так и на неродном языке.

Экспериментальное исследование понимания текста носителями китайского языка с помощью метода выделения ключевых слов (эксперимент 3).

Для того чтобы проверить гипотезу о совпадении КС для текстов на разных языках, был проведен эксперимент 3, в который мы включили тексты на китайском языке, являющимся типологически далеким от русского и английского языков.

Материал исследования

Тексты «Янус» и «Шака» были переведены носителем китайского языка, владеющим русским и английским языками, на китайский язык и вычитаны тремя носителями китайского языка. Во время перевода носитель китайского языка пользовался версиями текстов на русском и английском языках. Переводы были предложены двум другим носителям китайского языка для экспертной оценки качества перевода текстов.

Методика и ход эксперимента

Условия эксперимента 3 соответствовали условиям эксперимента с русскоязычными участниками, в котором использовались тексты «Янус» и «Шака». Все заполненные участниками эксперимента протоколы были переведены на русский язык носителем китайского языка, владеющим русским и английским языками. При анализе КС применялись описанные ранее принципы, за исключением того, что анализировались переводы КС на русский язык.

Участники

В эксперименте 3 приняли участие 88 человек (52 женщины и 36 мужчин в возрасте от 18 до 34 лет, средний возраст — 23 года), участниками стали добровольцы, студенты и аспиранты СПбГУ, для которых китайский язык был родным (L1), а русский — изучаемым (L2). Участники эксперимента 3 владели сертификатами ТРКИ, о чем указывали в анкетах. Распределение участников по уровням владения русским языком следующее: ТРКИ-1 (B1 по шкале CEFR) — 59, ТРКИ-2 (B2) — 20, ТРКИ-3 (C1) — 9.

Результаты

Тексты на иностранном языке оценивались респондентами как более сложные, чем тексты на родном языке ($t = 11,04, p < 0,001$) (табл. 6).

Для текста «Янус» на русском и китайском языках в истинный набор КС вошли 22 слова, для текста «Шака» на русском языке — 22 слова, на китайском языке — 21 слово. Гистограммы с сопоставлением истинных наборов КС для двух пар текстов см. на рис. 5–6.

Таблица 6. Средняя оценка субъективной сложности текстов «Янус», «Шака» носителями китайского языка

Текст	Язык текста	
	Русский	Китайский
«Янус»	5,2	2,8
«Шака»	5	2,1

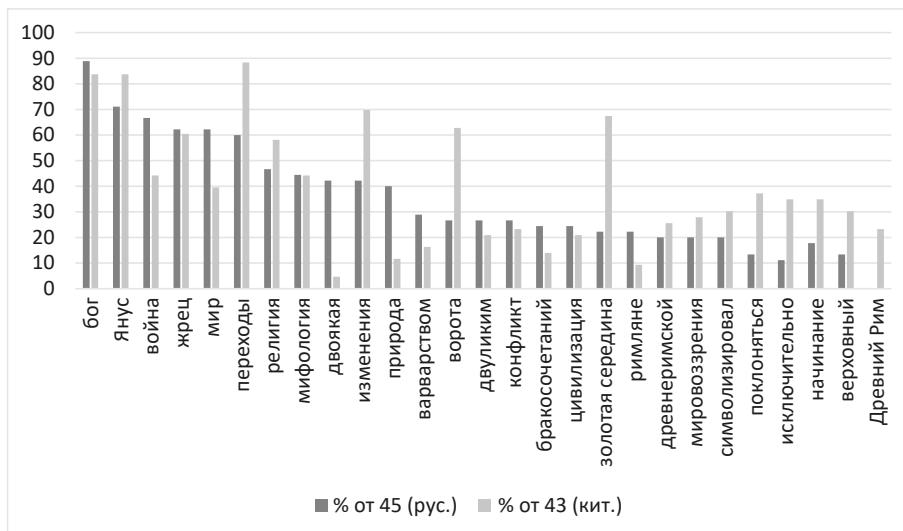


Рис. 5. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Янус» на русском и китайском языках (с переводом на русский язык)

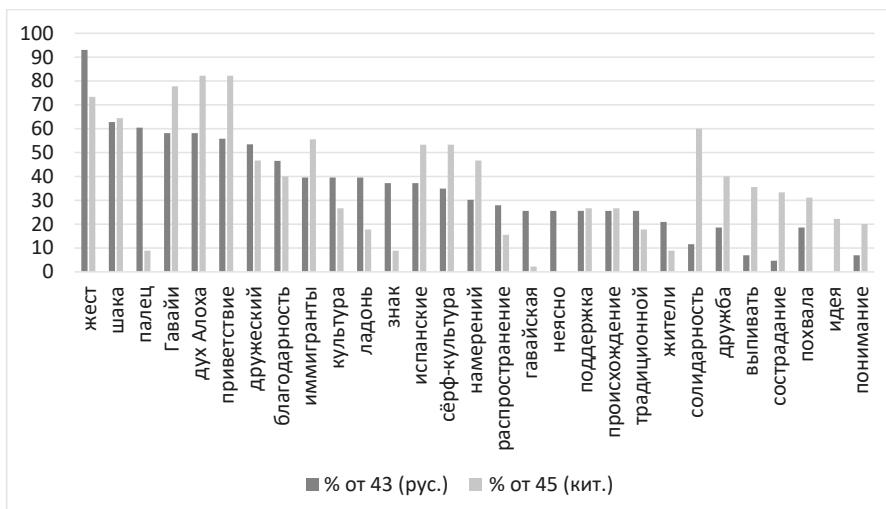


Рис. 6. Процент участников эксперимента, выписавших слово в качестве ключевого, для текста «Шака» на русском и китайском языках (с переводом на русский язык)

Были сопоставлены истинные наборы КС на китайском и русском языках для обоих текстов. Была обнаружена средняя корреляция между истинными наборами КС: для текста «Янус» $r = 0,608$, $p < 0,001$; для текста «Шака» $r = 0,522$, $p = 0,004$.

Больший разброс КС наблюдается для текстов на русском языке, чем парных им текстов на китайском языке, однако это различие не является значимым (табл. 7).

Таблица 7. Количество респондентов, выделенных респондентами КС, уникальных КС для текстов на русском и китайском языках

Параметр	Текст			
	«Янус»		«Шака»	
	Китайский язык	Русский язык	Китайский язык	Русский язык
Количество респондентов	43	45	45	43
Количество выделенных КС	573	547	583	517
Количество уникальных КС	62	72	63	65

Так же, как и в эксперименте с носителями русского языка, для оценки влияния ряда факторов на успешность выделения КС и ответов на вопросы по содержанию текста на L2 использовалась линейная регрессия. В качестве зависимых переменных выступили успешность выделения КС и количество правильных ответов на вопросы на L2, в качестве предикторов — успешность выделения КС и ответов на вопросы на L1, уровень языка (для участников, не указавших точный уровень, использовались пропущенные значения) и текст (табл. 8).

Таблица 8. Результаты линейной регрессии: влияние предикторов на успешность выделения КС и ответы на вопросы на L2 (эксперимент 3)

Зависимая переменная	Предиктор	Оценка параметра	Стандартная ошибка	Значение <i>t</i> -критерия	<i>p</i>
Успешность выделения КС на L2 ($R^2 = 0,3$)	Свободный коэффициент	1,89	1,63	1,16	0,250
	Успешность выделения КС на L1	0,68	0,15	4,49	<0,001*
	Ответы на вопросы на L1	-0,06	0,11	-0,55	0,590
	Ответы на вопросы на L2	0,31	0,08	3,93	<0,001*
	Уровень языка	-0,43	0,25	-1,74	0,090
	Текст («Шака»)	0,45	0,34	1,34	0,180
Ответы на вопросы на L2 ($R^2 = 0,37$)	Свободный коэффициент	-3,31	2,09	-1,58	0,120
	Ответы на вопросы на L1	0,31	0,14	2,27	0,025*
	Успешность выделения КС на L1	-0,14	0,22	-0,66	0,510
	Успешность выделения КС на L2	0,51	0,13	3,93	<0,001*
	Уровень языка	1,29	0,30	4,37	<0,001*
	Текст («Шака»)	-1,28	0,42	-3,04	0,003*

Примечание: «Янус» и «Шака», русский и китайский языки, звездочкой отмечены статистически значимые предикторы.

Для носителей китайского языка, так же, как и для носителей русского языка, успешность выделения КС на L2 связана с успешностью выделения КС на L1 и количеством верных ответов на L2. Чем лучше участники выделяют КС в тексте на L2, тем успешнее они отвечают на вопросы по тексту на L2. Содержание текста не влияет на успешность выделения КС на L2, что не согласуется с результатами, показанными в аналогичном эксперименте с носителями русского языка.

Взаимосвязь между успешностью ответов на вопросы по содержанию текста на неродном языке и уровнем L2 такая же, как у русскоязычных читателей этой же пары текстов: чем выше уровень языка, тем лучше участники справляются с этим заданием. Также была выявлена взаимосвязь между успешностью ответов на вопросы на родном и неродном языках. Содержание текста оказало значимое влияние на успешность ответов на вопросы на L2: участники значимо лучше отвечали на вопросы по содержанию текста «Шака», что не было обнаружено у русскоязычных читателей этой же пары текстов.

Обсуждение результатов эксперимента с носителями китайского языка

По результатам эксперимента 3 подтверждается несколько предположений, сделанных на предыдущем этапе исследования. Во-первых, подтверждается предположение о том, что истинный набор КС отражает содержание текста, поскольку также была обнаружена значимая положительная корреляция между абсолютной частотой встречаемости КС, выделенных на родном языке, и абсолютной частотой встречаемости КС, выделенных на неродном, китайском языке. Во-вторых, положительная связь между успешностью выделения КС и количеством правильных ответов для текстов на неродном языке свидетельствует о том, что индивидуальный набор КС все же в определенной степени отражает понимание текста читающим. В-третьих, уровень владения иностранным языком так же, как и в предыдущих экспериментах, вносит вклад в то, насколько правильно участники отвечают на вопросы на неродном языке. Чем выше уровень неродного языка, тем успешнее участники отвечают на вопросы. Таким образом, успешность ответов на вопросы по содержанию текста на неродном языке связана с языковой компетенцией читателя, а успешность выделения КС не связана.

Выводы

В данном исследовании была предпринята попытка применить методику выделения КС к исследованию компетенции понимания текста на родном и неродном языках. Мы предположили, что наборы слов, которые участники выделяют в текстах как ключевые, могут отражать то, насколько хорошо участники понимают текст. Индивидуальные наборы КС участников сопоставлялись с истинными наборами КС, выделенными на основе слов, которые выписали более 20 % участников. Успешность выделения КС у отдельного участника определялась по количеству совпадений КС из индивидуального набора и КС из истинного набора.

В первую очередь проверялась гипотеза о том, что истинные наборы КС для текстов на разных языках будут совпадать в силу того, что КС отражают инвариант-

ный смысл текста. Также проверке подверглось предположение о связи успешности выделения КС и успешности ответов на вопросы по содержанию текста. Было проведено три эксперимента: два эксперимента с носителями русского языка, которые читали тексты и выполняли задания на русском (родном) и английском (неродном) языках, а также эксперимент с носителями китайского языка, которые читали тексты и выполняли задания на китайском (родном) и русском (неродном) языках.

Предположение о том, что истинные наборы КС для одного текста на разных языках будут совпадать, подтвердилось как для носителей русского языка, изучающих английский, так и для носителей китайского языка, изучающих русский, что подтверждается значимой положительной корреляцией между абсолютными значениями количества КС из истинных наборов для всех пар текстов, т. е. одинаковые по содержанию тексты, написанные на разных языках, обладают для читателей одним и тем же смысловым ядром.

Во всех трех экспериментах была обнаружена положительная корреляция между успешностью выделения КС и успешностью ответов на вопросы на неродном языке. Такой результат подтверждает гипотезу о том, что оба задания проверяют степень понимания текста, что не подтверждает результаты об отсутствии связи между успешностью выделения ключевых слов и успешностью ответов на вопросы исследования [Engelen et al. 2018: 299].

Для носителей как русского языка (в одном из двух экспериментов), так и китайского успешность ответов на вопросы по содержанию текста на неродном языке оказалась связана с языковой компетенцией читателя: чем выше уровень языка, тем успешнее участники отвечают на вопросы по содержанию текста. В то же время успешность выделения КС не связана с языковой компетенцией участников ни в одном эксперименте, что позволило сделать предположение о том, что навык выделения КС менее связан с языковой компетенцией носителя. Такое свойство методики может быть использовано в тех случаях, когда при оценке понимания текста требуется исключить лингвистические факторы (разные уровни языковых компетенций).

Результаты исследования позволяют сделать вывод об эффективности использования методики набора КС для оценки степени понимания текста на родном и неродном языках и рекомендовать ее как дополнительный способ оценки понимания прочитанного. Различие с полученными ранее результатами может быть связано с процедурой использования методики. Так, в работе [Engelen et al. 2018] успешность выделения ключевых слов оценивали эксперты, в то время как в настоящей работе использовалась процедура, описанная в исследовании [Мурзин, Штерн 1991: 73–91], и индивидуальные наборы КС сопоставлялись с истинными, что, на наш взгляд, позволяет более объективно оценить успешность выполнения задания. Перспективой исследований в этом направлении может стать сопоставление результатов оценки успешности выделения КС в зависимости от процедуры использования этой методики.

Поскольку традиционной методикой измерения компетенции понимания текста считаются ответы на вопросы по содержанию, существует большое количество эмпирических данных, описывающих вклад в понимание прочитанного ряда факторов, связанных, например, с морфологическим осознанием, словарным запасом, невербальным интеллектом, памятью, вниманием, метакогнитивными навыками и т. д. (напр., [Li, Kirby 2014; Lervåg, Melby-Lervåg 2022] и мн. др.). Включение в по-

добные исследования еще одного измерения, оценивающего понимание текста с помощью методики набора ключевых слов, выглядят весьма перспективным.

Еще одним направлением будущих исследований может стать проверка предположения наличия универсальных стратегий работы со смыслом текста вне зависимости от языка, на котором написан текст. Полученные в текущей работе результаты в некоторой степени подтверждают предположения об универсальных стратегиях, сделанные в работе [Петрова и др. 2017: 155], однако в будущих исследованиях предлагается проверить эту гипотезу с привлечением текстов на разных языках, а также в разных знаковых системах (в том числе с сочетанием вербальных и невербальных компонентов).

Основными ограничениями текущего исследования стоит считать особенности процедуры работы. Во-первых, было использовано только четыре текста и две пары языков L1–L2 (русский — английский; китайский — русский); во-вторых, не был учтен порядок записи КС, который может оказаться значимым для оценки степени понимания текста; в-третьих, оценка уровня иностранного языка участников эксперимента 1 и эксперимента 2 основывалась на субъективных представлениях участников. Если первое и второе ограничения возможно устраниТЬ при помощи модификации процедуры обработки данных и использования большего числа стимульных текстов, то последнее ограничение связано с отсутствием универсальных мультиязыковых инструментов оценки языковых компетенций. Существующие тесты на определение уровня владения разными языками, по всей видимости, могут различаться в зависимости от языка.

* * *

Авторы благодарны Се Жои за помощь в переводе текстов на китайский язык и обработке ответов, а также всем, кто способствовал проведению эксперимента и привлечению участников.

Литература

- Базылев, Красильникова 2011 — Базылев В. Н., Красильникова В. Г. Измерение компетенции «понимание текста». *Вопросы психолингвистики*. 2011, (14): 42–47.
- Грудева, Губушкина 2020 — Грудева Е. В., Губушкина А. А. Набор ключевых слов и устный пересказ как вторичные тексты (на материале вторичной речевой деятельности учащихся 6-х классов). *Вестник Череповецкого государственного университета*. 2020, 2 (95): 56–72.
- Мурзин, Штерн 1991 — Мурзин Л. Н., Штерн А. С. *Текст и его восприятие*. Свердловск: Изд-во Урал. ун-та, 1991.
- Петрова и др. 2017 — Петрова Т. Е., Риехакайнен Е. И., Кузнецова А. С., Мараев А. В., Шаталов М. А. Выделение ключевых слов в вербальных и невербальных паттернах. *Социо- и психолингвистические исследования*. 2017, (5): 149–156.
- Сахарный, Штерн 2006 — Сахарный Л. В., Штерн А. С. Набор ключевых слов как тип текста. В сб. ...*Слово отзовется: памяти А. С. Штерн и Л. В. Сахарного*. Доценко Т. И., Ерофеева Е. В., Овчинникова И. Г., Чугаева Т. Н. (ред.). Пермь: Перм. гос. ун-т, 2006. С. 50–59.
- Beliga et al. 2015 — Beliga S., Meštrović A., Martinčić-Ipšić S. An overview of graph-based keyword extraction methods and approaches. *Journal of Information and Organizational Sciences*. 2015, 39 (1): 1–20.
- de Bruin et al. 2011 — de Bruin A. B., Thiede K. W., Camp G., Redford J. Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*. 2011, 109 (3): 294–310.

- Engelen et al. 2018 — Engelen J. A. A., Camp G., van de Pol J., de Bruin A. B. H. Teachers' monitoring of students' text comprehension: Can students' keywords and summaries improve teachers' judgment accuracy? *Metacognition and learning*. 2018, (13): 287–307.
- Firoozeh et al. 2020 — Firoozeh N., Nazarenko A., Alizon F., Daille B. Keyword extraction: Issues and methods. *Natural Language Engineering*. 2020, 26 (3): 259–291.
- Friedman, Hoffman-Goetz 2006 — Friedman D. B., Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education & Behavior*. 2006, 33 (3): 352–373.
- Hiebert 2011 — Hiebert E. H. Beyond single readability measures: Using multiple sources of information in establishing text complexity. *Journal of Education*. 2011, 191 (2): 33–42.
- Kuperman et al. 2022 — Kuperman V., Siegelman N., Schroeder S., Acartürk C., Alexeeva S., Amenta S., Bertram R., Bonandrini R., Brysbaert M., Chernova D., Da Fonseca S. M., Dirix N., Duyck W., Fella A., Frost R., Gattei C. A., Kalaitzi A., Löö K., Marelli M., Nisbet K., Papadopoulos T. C., Protopapas A., Savo S., Shalom D. E., Slioussar N., Stein R., Sui L., Taboh A., Tønnesen V., Usal K. A. Text reading in English as a second language: Evidence from the multilingual eye-movements corpus. *Studies in Second Language Acquisition*. 2022, 45 (1): 1–35.
- Lervåg, Melby-Lervåg 2022 — Lervåg A., Melby-Lervåg M. Modeling the development of reading comprehension. In: *The science of reading: A handbook*. Snowling M. J., Hulme C., Nation K. (eds). Hoboken: Wiley-Blackwell, 2022. P. 280–297.
- Li, Kirby 2014 — Li M., Kirby J. R. Unexpected poor comprehenders among adolescent ESL students. *Scientific Studies of Reading*. 2014, 18 (2): 75–93.
- Nation 2022 — Nation K. Children's reading comprehension difficulties. In: *The science of reading: A handbook*. Snowling M. J., Hulme C., Nation K. (eds). Hoboken: Wiley-Blackwell, 2022. P. 298–322.

Статья поступила в редакцию 23 сентября 2024 г.

Статья рекомендована к печати 19 мая 2025 г.

Vladislav I. Zubov

St. Petersburg State University,
7–9, Universitetskaya nab., St. Petersburg, 199034, Russia
HSE University,
20, ul. Myasnitskaya, Moscow, 101000, Russia
<https://orcid.org/0000-0001-8714-7491>
vladzubov21@gmail.com

Anastasiia A. Ladnova

St. Petersburg State University,
7–9, Universitetskaya nab., St. Petersburg, 199034, Russia
<https://orcid.org/0000-0003-0132-7016>
a.ladnova@spbu.ru

Assessing text comprehension in native and non-native languages using keyword extraction method: Evidence from Russian, English, and Chinese*

For citation: Zubov V. I., Ladnova A. A. Assessing text comprehension in native and non-native languages using keyword extraction method: Evidence from Russian, English, and Chinese. *Vestnik of Saint Petersburg University. Language and Literature*. 2025, 22 (3): 664–683. EDN XKRVES (In Russian)

The study attempts to apply the keyword extraction method to analyse text comprehension competence in both native and non-native languages. The hypothesised that the sets of key-

* The work was supported by St. Petersburg State University “Mechanisms of reading and interpreting text in native and non-native languages: an interdisciplinary experimental study using eye movement recording methods, visual analytics and virtual reality technologies”, project code 124032900009-2.

words identified by participants reflect their level of text comprehension. Individual sets of keywords were compared with “true sets”, which were formed based on the words selected by more than 20 % of participants. We tested hypotheses regarding the match of “true keyword sets” for texts in different languages, and the correlation between the success of keyword extraction and performance on content-related questions. Three experiments were conducted involving native Russian speakers reading texts in Russian and English, and native Chinese speakers working with texts in Chinese and Russian. In all three experiments, a positive correlation was found between the success of keyword extraction and the success of answering questions in L2, confirming the hypothesis that both tasks assess the degree of text comprehension. The success in answering questions and L2 proficiency level were related for all participants, but the success of keyword extraction did not depend on language proficiency. This suggests that the skill of keyword extraction is less related to language competence and may be used to assess text comprehension regardless of linguistic factors. The results of the study confirm that the keyword extraction method is effective for evaluating text comprehension and highlight differences from previous research, which may be attributed to the procedures used. Future research could involve further comparison of results depending on the methods applied.

Keywords: reading, comprehension, keyword extraction method, second language.

References

- Базылев, Красильникова 2011 — Bazylev V.N., Krasil'nikova V.G. Measuring text comprehension competence. *Voprosy psicholinguistiki*. 2011, (14): 42–47. (In Russian)
- Грудева, Губушкина 2020 — Grudeva E.V., Gubushkina A.A. Set of keywords and oral retelling as secondary texts (by the material of the 6th grade students' secondary speech activities). *Bulletin of the Cherepovets State University*. 2020, 2 (95), 56–72. (In Russian)
- Мурзин, Штерн 1991 — Murzin L.N., Stern A.S. *Text and its perception*. Sverdlovsk: Izdatel'stvo Ural'skogo universiteta Publ., 1991. (In Russian)
- Петрова и др. 2017 — Petrova T.E., Riekhakainen E.I., Kuznetsova A.S., Maraev A.V., Shatalov M.A. Keywords extraction in verbal and non-verbal patterns. *Sotsio- i psicholinguisticheskie issledovaniia*. 2017, (5): 149–156. (In Russian)
- Сахарный, Штерн 2006 — Sakharnyi L.V., Shtern A.S. Keyword set as a type of text. In: ...*Slovo otzovet-sia: pamiati A.S. Shtern i L.V. Sakharnogo*. Dotsenko T.I., Erofeeva E.V., Ovchinnikova I.G., Chu-gaeva T.N. (eds). Perm': Permskii gumanitarnyi universitet Publ., 2006. P.50–59. (In Russian)
- Beliga et al. 2015 — Beliga S., Meštrović A., Martinčić-Ipšić S. An overview of graph-based keyword extraction methods and approaches. *Journal of Information and Organizational Sciences*. 2015, 39 (1): 1–20.
- de Bruin et al. 2011 — de Bruin A.B., Thiede K.W., Camp G., Redford J. Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*. 2011, 109 (3): 294–310.
- Engelen et al. 2018 — Engelen J.A.A., Camp G., van de Pol J., de Bruin A.B.H. Teachers' monitoring of students' text comprehension: Can students' keywords and summaries improve teachers' judgment accuracy? *Metacognition and learning*. 2018, (13): 287–307.
- Firoozeh et al. 2020 — Firoozeh N., Nazarenko A., Alizon F., Daille B. Keyword extraction: Issues and methods. *Natural Language Engineering*. 2020, 26 (3): 259–291.
- Friedman, Hoffman-Goetz 2006 — Friedman D.B., Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education & Behavior*. 2006, 33 (3): 352–373.
- Hiebert 2011 — Hiebert E.H. Beyond single readability measures: Using multiple sources of information in establishing text complexity. *Journal of Education*. 2011, 191 (2): 33–42.
- Kuperman et al. 2022 — Kuperman V., Siegelman N., Schroeder S., Acartürk C., Alexeeva S., Amenta S., Bertram R., Bonandrini R., Brysbaert M., Chernova D., Da Fonseca S.M., Dirix N., Duyck W., Fella A., Frost R., Gattei C.A., Kalaitzi A., Lõo K., Marelli M., Nisbet K., Papadopoulos T.C., Protopapas A., Savo S., Shalom D.E., Slioussar N., Stein R., Sui L., Taboh A., Tønnesen V., Usal K.A. (2022). Text rea-

- ding in English as a second language: Evidence from the multilingual eye-movements corpus. *Studies in Second Language Acquisition*. 2022, 45 (1): 1–35.
- Lervåg, Melby-Lervåg 2022 — Lervåg A., Melby-Lervåg M. Modeling the development of reading comprehension. In: *The science of reading: handbook*. Snowling M.J., Hulme C., Nation K. (eds). Hoboken: Wiley-Blackwell, 2022. P. 280–297.
- Li, Kirby 2014 — Li M., Kirby J. R. Unexpected poor comprehenders among adolescent ESL students. *Scientific Studies of Reading*. 2014, 18 (2): 75–93.
- Nation 2022 — Nation K. Children's reading comprehension difficulties. In: *The science of reading: handbook*. Snowling M.J., Hulme C., Nation K. (eds). Hoboken: Wiley-Blackwell, 2022. P. 298–322.

Received: September 23, 2024

Accepted: May 19, 2025