

Научная статья

УДК 81-13

DOI 10.25205/1818-7935-2024-22-3-84-97

Оценка надежности лексических списков для автоматизированной оценки уровня владения русским языком как иностранным

Антон Юрьевич Вахранев

Национальный исследовательский университет «Высшая школа экономики»
Москва, Россия

ag15barto@gmail.com, <https://orcid.org/0009-0006-8721-1872>

Аннотация

Оценка языковой грамотности играет важную роль в образовании, но она часто оценивается субъективно, что может внести предвзятость и несоответствия в результаты оценки. В связи с этим многие исследователи предлагают использовать автоматизированные и полуавтоматизированные методы оценки на основе лингвистических характеристик текстов. В данном исследовании рассматривается применимость доступных списков лексики в качестве инструментов для автоматической оценки уровня владения русским языком студентами.

Существует несколько различных видов лексических списков, такие как списки по частотности слов и минимальные списки лексики. В этом исследовании анализируются четыре популярных русскоязычных лексических списка, которые используются в образовании и анализе лексических знаний. Предполагается, что тексты студентов, оцененные на более низких уровнях языковой грамотности, будут в большей степени содержать часто употребляемые слова и слова низкого уровня, соответствующие распределению лексических элементов по частотности или уровню владения языком, представленному в этих списках. Студенты, оцененные на более высоких уровнях грамотности, наоборот, будут использовать менее часто встречающиеся и более сложные лексические единицы. Анализ корреляции между этими ресурсами и текстами, созданными студентами, позволяет получить понимание о целесообразности использования лексических списков для оценки уровня владения русским языком.

Для анализа корреляции между выбранными лексическими списками и текстами студентов используются собственные скрипты Python. Кроме того, был применен метод кластерного анализа, известный как анализ главных компонент (PCA), чтобы проверить гипотезу о том, что тексты студентов на одном и том же уровне грамотности имеют тенденцию использовать схожий базовый словарь с некоторой степенью вариации.

Исследование поднимает важные вопросы о том, насколько эффективно можно использовать лексические списки для оценки языковой грамотности. Полученные результаты могут стать основой для создания более точных и всесторонних методов оценки лексической грамотности студентов, изучающих русский язык.

Ключевые слова:

языковая грамотность, автоматическая оценка, лексический анализ, русский язык, образование

Для цитирования

Вахранев А. Ю. Оценка надежности лексических списков для автоматизированной оценки уровня владения русским языком как иностранным // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2024. Т. 22. № 3. С. 84–97. DOI 10.25205/1818-7935-2024-22-3-84-97

© Вахранев А. Ю., 2024

ISSN 1818-7935

Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2024. Т. 22, № 3
Vestnik NSU. Series: Linguistics and Intercultural Communication, 2024, vol. 22, no. 3

Assessment of the Reliability of Lexical Lists for Automated Evaluation of Proficiency in Russian as a Foreign Language

Anton Y. Vakhranov

National Research University Higher School of Economics
Moscow, Russian Federation

ag15barto@gmail.com, <https://orcid.org/0009-0006-8721-1872>

Abstract

The assessment of language proficiency plays a crucial role in education, but it often relies on subjective evaluation methods, which can result in bias and inconsistencies in the results. To address this challenge, many researchers advocate for the use of automated and semi-automated assessment methods based on linguistic characteristics of texts. In this study, we explore the applicability of available vocabulary lists as tools for automatically evaluating the proficiency levels of students learning Russian.

There exist several types of lexical lists, including frequency-based word lists and minimum vocabulary lists. In this research, we analyze four popular Russian-language lexical lists commonly employed in education and lexical knowledge analysis. We hypothesize that texts produced by students assessed at lower language proficiency level will predominantly contain frequently used words and low-level vocabulary items corresponding to the distribution of lexical elements by frequency or language proficiency level represented in these lists. Conversely, students assessed at a higher proficiency level are expected to employ less commonly used and more complex lexical units. By examining the correlation between these resources and student-generated texts, we aim to gain insight into the suitability of using lexical lists for evaluating proficiency in Russian.

For the analysis of correlations between the selected lexical lists and student texts, we employ custom Python scripts. Additionally, we utilize a cluster analysis method known as Principal Component Analysis (PCA) to test the hypothesis that students at the same proficiency level tend to use a similar basic vocabulary with some degree of variation.

This study raises important questions about the effectiveness of using lexical lists for assessing language proficiency. The findings may serve as a foundation for developing more accurate and comprehensive methods for evaluating lexical proficiency among students learning the Russian language.

Keywords:

language proficiency, automated assessment, lexical analysis, Russian language, education

For citation

Vakhranov A. Y. Assessment of the reliability of lexical lists for automated evaluation of proficiency in Russian as a foreign language *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2024, vol. 22, no. 3, pp. 84–97. DOI 10.25205/1818-7935-2024-22-3-84-97

Введение

Оценка языковой компетенции является важной и ответственной задачей в образовании. Однако практика оценки часто полагается на качественные общие описания и субъективные оценки, что может внести предвзятость и несогласованность в результаты оценки. Для преодоления критики субъективности и обеспечения более объективной и надежной оценки ряд исследователей предложили использовать протоколы автоматической и полуавтоматической оценки языка на основе предположения о том, что тексты разного качества и/или разных уровней владения языком характеризуются узорами использования и распределения объективных языковых признаков (см. [Lu, 2010]). Несмотря на то что объем исследований, направленных на выявление конкретных коррелятов уровня владения языком или критериальных признаков, растет (см. [Hawkins & Buttery, 2010; Callies & Götz, 2015]), остается много теоретических и методологических вопросов. В данном исследовании мы рассматриваем пригодность доступных словарных списков (включая списки частотности и списки минимального словаря) как инструментов для автоматической оценки уровней владения русским языком студентами.

Списки слов могут различаться по своей природе и назначению. Списки частотности составляются на основе частоты встречаемости слов в корпусе общего языка, в то время

как списки минимального словаря являются педагогическими инструментами, разработанными для иностранных/вторых языков (L2) с целью направления обучения и оценки на разных уровнях владения языком. Используемые не только в образовательных целях, таких как разработка учебных программ и материалов, эти ресурсы также широко используются при анализе лексических знаний; однако их эффективность как объективных показателей владения языком должна быть проверена и протестирована на практике.

В данном исследовании анализируются четыре русских словарных списка, обычно используемых в образовательных целях, и оцениваются эти списки с точки зрения их использования в (полу)автоматической оценке лексической сложности. Эти списки включают: Visualize Russian Tool (далее VRT; см. [Clancy, 2014–2023]), SMARTool (см. [Janda et al., 2018–2023]), «Лексический минимум для русского как иностранного языка» (далее LM; см. [Golubeva, 2015]) и KELLY (см. [Kilgarriff et al., 2014]). Предполагается, что тексты студентов, оцененные на более низком уровне владения, будут в основном содержать чрезвычайно частые и низкоуровневые лексические элементы, соответствующие распределению лексических элементов по частотности или уровням владения в списке. В отличие от этого, тексты студентов, оцененные на более высоких уровнях владения, предположительно будут содержать менее частые и более высокоуровневые лексические элементы. Исследование корреляции между этими ресурсами и реальными письменными текстами, созданными студентами, позволяет получить представление об эффективности словарных списков при оценке владения языком.

В данном исследовании используются скрипты Python, разработанные внутри учреждения, для анализа корреляции между выбранными словарными списками и текстами учащихся; кроме того, исследование использует метод кластеризации, известный как анализ главных компонент (PCA), чтобы проверить гипотезу о том, что тексты учащихся на одном уровне владения склонны использовать схожий базовый словарь с некоторой степенью вариации.

Исследование связи между словарными списками и владением языком, возможно, позволит внести вклад в область оценки языка и предоставить информацию об ограничениях и потенциале использования автоматических методов для оценки владения языком. Эти результаты помогут разработать более точные и всесторонние инструменты измерения лексической сложности и владения русским языком студентами.

1. Обзор литературы

Данный раздел предоставляет обзор соответствующих исследований, связанных с оценкой владения русским языком и использованием русских частотных списков и списков минимального словаря. Он также выявляет невыясненные вопросы в этой области и подчеркивает важность текущего исследования.

Автоматическая оценка владения языком, включая оценку лексической компетенции, представляет собой сложную задачу, которая зависит от объективных и надежных мер. А. Н. Лапошина (2018) сосредоточилась на исследовании возможностей автоматизации определения сложности текста на русском языке для студентов, изучающих его как второй язык. В эксперименте результаты математической модели были сравнены с оценками экспертов, мнениями студентов, результатами тестов и официальными требованиями к владению русским языком. Автоматизация процесса определения уровня текста предоставляет быстрый и удобный доступ к подходящим материалам для чтения как для учителей, так и для студентов. Результаты эксперимента показали, что средняя скорость чтения для уровня B1 составляла от 40 до 50 слов в минуту при чтении в процессе обучения и от 80 до 100 слов в минуту для общего понимания содержания. Процент слов, отсутствующих в минимальном лексиконе для уровня B1 (5–7 %), значительно превышал норму, однако последующая обратная связь от студентов и учителей указывала на то, что им встречается меньше незнакомых слов, чем можно было бы ожидать.

Исследование предполагает, что умение применять модели словообразования и предыдущий языковой опыт студентов могут способствовать их способности понимать сложные слова.

А. Н. Лапошина и М. Ю. Лебедева (2021) изучили языковую доступность текстов для студентов, изучающих русский как второй язык, на разных уровнях владения, сравнив их с минимальными лексиконами. Тексты, классифицированные как уровень CEFR A1, показали более высокий процент слов, которых не было в минимальных лексиконах, чем ожидалось. Они разработали инструмент Текстометр для автоматического анализа сложности текста на шкале CEFR, предоставляя полезную информацию для адаптации текстов к образовательным задачам, словарным спискам и другим целям. Инструмент учитывает более 800 текстов и использует машинное обучение и обработку естественного языка. Текстометр может быть полезен учителям, авторам и издателям при выравнивании текстов по уровням владения и образовательным целям. Модель опирается на практический опыт и материалы из реального мира, с которыми сталкиваются студенты. Несмотря на многообещающий характер, все еще существуют некоторые ограничения, среди которых зависимость от словарных списков, которые не были подвергнуты эмпирической проверке.

Предыдущие исследования (см. [Ляшевская и др., 2010]) по автоматическому анализу текстов и морфологическим парсерам для русского языка рассматривают различные вызовы, которые возникают при попытке автоматически измерить различные параметры текстов, такие как категоризация частей речи, оптимальный размер словарей, стратегии обработки многозначных слов. К этому списку вызовов следует добавить вопрос о том, насколько надежны имеющиеся словарные списки для автоматической оценки лексической сложности текста на русском языке. В этом исследовании анализируются корреляции между лексическими элементами, охваченными имеющимися словарными списками, и уровнями владения языком, присвоенными студенческим текстам, оцененными людьми-оценщиками; через это будет определяться эффективность словарных списков в понимании уровня владения языком.

2. Данные и методы

В данном исследовании была собрана выборка экзаменационных работ студентов, обучающихся в Миддлбери Летней Русской школы и изучающих русский язык как второй язык. Выборка включает более тысячи эссе, категоризированных на восемь уровней владения экспертами, начиная с уровня Novice Mid и заканчивая Advanced High в соответствии с шкалой ACTFL (см. рекомендации American Council on the Teaching of Foreign Languages по оценке уровня владения языком), что соответствует уровням с A1 по B2 в шкале CEFR (см. шкалу владения языком в рамках Общеввропейских рекомендаций по языкам). Эссе распределены по уровням ACTFL следующим образом:

Таблица 1

Распределение студенческих работ по уровням ACTFL

Table 1

Distribution of student work by ACTFL level	
Уровень	Количество эссе
Novice Mid	13
Novice High	50
Intermediate Low	70
Intermediate Mid	371
Intermediate High	254
Advanced Low	218
Advanced Mid	114
Advanced High	13

Несмотря на то что полное совпадение между уровнями владения на шкалах ACTFL и CEFR отсутствует, в предыдущих исследованиях были зарегистрированы положительные соответствия [см. Tschirner et al., 2012]. В данном исследовании выделяются четыре набора данных обучающихся, по одному на каждый уровень CEFR, от A1 до B2 (63, 441, 472 и 127 эссе в каждом наборе соответственно).

Для оценки уровня владения лексикой русского языка студентами на уровнях A1-B2 сравниваются четыре лексических ресурса: VRT, SMARTool, LM и KELLY.

Список VRT содержит более 28 000 лемм и служит ценным ресурсом для студентов из США. Однако этот список не сбалансирован по жанрам и стилям.

Список SMARTool включает в себя 6214 существительных, прилагательных и глаголов, представляющих уровни A1, A2, B1 и B2 по шкале CEFR. Этот ресурс является компиляцией лексических элементов из Нового частотного словаря русского языка [см. Lyashevskaya & Sharoff, 2009] и существующих списков минимального лексикона [LM, см. Golubeva, 2015]. Он был тщательно оценен профессиональными учителями и исследователями.

Список минимального лексикона для русского как иностранного языка (LM) включает 5 734 слова, выбранных на основе их использования в повседневном общении. Он охватывает широкий спектр тем и служит стандартным ресурсом для тестирования уровня владения языком в России.

Списки KELLY основаны на большом объеме текстов, доступных для этих языков в сети «Интернет». Эти списки дают представление об использовании языка и его особенностях, учитывая частоту встречаемости различных частей речи в корпусе текстов.

Хотя эти списки различаются по структуре и предполагаемой цели, они все созданы с учетом потребностей изучающих русский язык и ориентированы на определенные уровни владения. Возникает вопрос о том, какой из этих списков может служить надежным инструментом для автоматической оценки лексической сложности текста.

На первом этапе данного исследования были проанализированы корреляции между словами в четырех словарных списках. Для этой задачи использован коэффициент Сёренсена – Дайса (далее – коэффициент S&D) [см. Sorensen, 1948; Dice Lee R., 1945]. Этот непараметрический коэффициент корреляции подходит для ординальных данных, таких как списки слов разной длины. Он измеряет степень сходства и согласования между рейтингами двух списков, учитывая как направление, так и расстояние несогласий:

$$K_{0,-1} = \frac{2P(A \cap B)}{P(A) + P(B)}. \quad (1)$$

Эффективность коэффициента S&D в анализе отношений между различными аспектами производства текста была продемонстрирована в ряде исследований [см. Looman & Campbell, 1960; Ondov, et al., 2016].

Затем были сопоставлены фактические словари, используемые студентами в эссе на разных уровнях владения, с соответствующими словарными списками. Это сравнение позволило понять, насколько списки отражают фактическое использование лексики на каждом уровне владения.

Наконец, была проверена гипотеза о том, что студенты на определенных уровнях владения склонны использовать сходные «резервуары» словарного запаса. Для этой части анализа был использован метод главных компонент (PCA) для каждого уровня CEFR. Этот процесс позволил визуализировать, формируют ли экзаменационные работы отдельные кластеры, коррелирующие с уровнями владения языком.

В целом, собранные данные и методология включают в себя экзаменационные работы студентов, четыре словарных списка, коэффициент Сёренсена – Дайса для анализа корреляции слов и визуализацию с использованием РСА. Эти методы позволяют оценивать уровни владения языком, понимать использование лексики и выявлять закономерности в данных.

3. Анализ и результаты для каждого метода отдельно

Цель этого раздела – представить результаты, полученные с помощью каждого метода, использованного в исследовании. Рассмотрим основные результаты и их последствия, а также оценим, подтверждаются ли или опровергаются исследовательские гипотезы.

Табл. 2 предоставляет обзор размеров списков, использованных в анализе. Эти размеры варьировались в зависимости от уровня владения, изменяясь от 581 до 12 596 элементов.

Таблица 2

Размеры словарных списков, разделенные по уровню владения языком

Table 2

The size of the vocabulary lists split for proficiency level

Уровни	VRT	Kelly	SMARTool	LM
1	1879	938	754	994
2	2713	1060	1048	581
3	11249	1563	1081	941
4	12596	3424	3331	3218
	28437	6985	6214	5734

3.1. Корреляция Сёренсена – Дайса между словарными списками

В этом разделе сравниваются между собой списки, использованные в исследовании. Было обнаружено пересечение между списками, особенно между SMARTool и LM, как показано на рис. 1. Эти перекрытия, вероятно, объясняются тем, что SMARTool частично основан на LM. Однако важно отметить, что использованные в анализе списки не были исчерпывающими и могли не полностью отражать словарь, используемый студентами, так как студенты могут использовать лексику, не входящую в данные списки. Размеры списков были относительно небольшими и предоставляли лишь частичное представление фактического словаря.

Учитывая сильные и слабые стороны каждого ресурса, рекомендуется, чтобы учащиеся и учителя тщательно рассмотрели свои конкретные потребности и цели при выборе ресурса. В этом отношении SMARTool представляется наиболее полным и полезным. Совмещение нескольких ресурсов может обеспечить более глубокое понимание словаря, необходимого для обучения или оценки.

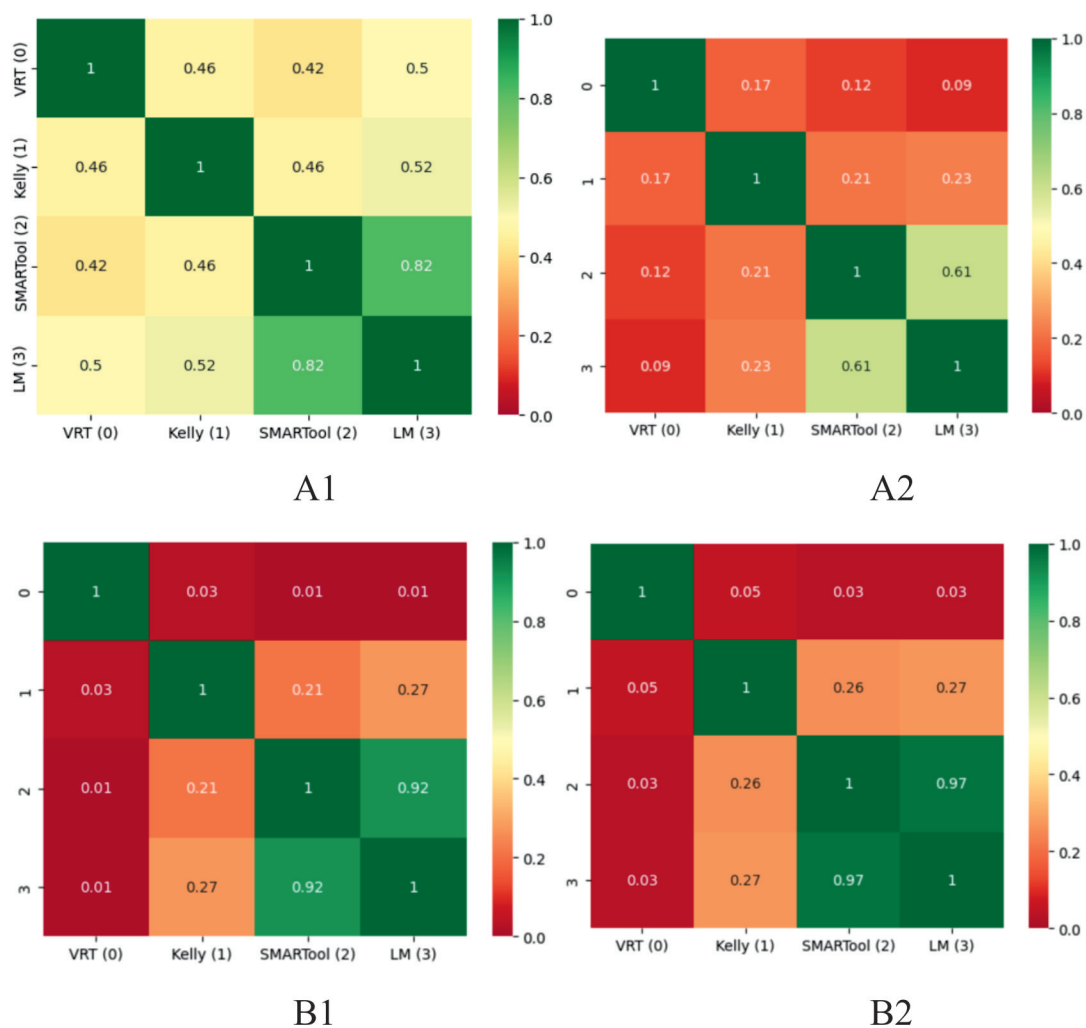


Рис. 1. Корреляция Сёренсена – Дайса между словарными списками для всех четырех уровней
 Fig. 1. Sørensen – Dice correlation between vocabulary lists for all the four levels

3.2. Списки слов как инструменты для оценки лексической сложности в текстах студентов

Второй исследуемый вопрос заключается в следующем: хотя списки не коррелируют между собой, может ли быть найден такой список, который наилучшим образом охватывает лексическое использование студентов, изучающих русский как второй язык? Будучи найден, он мог бы служить хорошей отправной точкой при измерении лексической грамотности. Для обеспечения достоверности результатов было проведено тщательное ручное исследование экстремальных выбросов, в которых менее 45 % лемм принадлежали всем соответствующим уровням. Эти выбросы оказались реальными русскими словами.

Во-первых, был обнаружен ряд неподходящих лемм, которые не фигурировали в каком-либо из списков, но присутствовали в данных студентов. Количество этих лемм варьируется от средних 5,55 до 48,69 %, в зависимости от списка, как показано в табл. 3.

Таблица 3

Среднее количество несоответствующих лемм для каждого уровня, %

Table 3

An average number of unmatched lemmas for each level, %				
Уровень	KELLY	VRT	SMARTool	LM
A1	5,81	5,55	16,6	5,71
A2	7,2	7,29	28,38	6,74
B1	10,57	11,87	37,29	10,09
B2	17,08	22,13	48,69	17,66

Во-вторых, анализ корреляции между словарем продвинутых студентов и нашими списками показал степень согласованности, изменяющуюся от 50 до 100 % в зависимости от конкретного списка. Эти результаты позволяют сделать вывод, что списки слов могут быть ценными инструментами для оценки уровня владения языком студентов, изучающих русский как второй язык, на продвинутом уровне. Имеет смысл использовать списки слов совместно с другими потенциальными показателями уровня владения языком, которые планируется изучить в дальнейшем (средняя длина токена, количество служебных частей речи, TTR и т. д.), на всех уровнях. Для студентов с нижними уровнями владения степень согласованности была значительно ниже, варьируя от 25 до 93 %. Таким образом, несмотря на практичность списков слов, они могут быть менее надежным методом оценки для студентов с более низким уровнем владения. Включение дополнительных показателей улучшит точность и всесторонность оценки уровня владения языком для студентов на всех этапах изучения русского языка.

В то же время распределение слов из заранее составленных списков среди текстов разных уровней владения позволяет предположить, что слова действительно характеризуют уровни владения, хотя и не линейно. Ниже показаны диаграммы (рис. 2), представляющие четыре текста на четырех разных уровнях владения. Как видно, доля слов уровня A1 постепенно уменьшается, в то время как количество слов, отнесенных к более продвинутым уровням владения, показывает восходящий тренд.

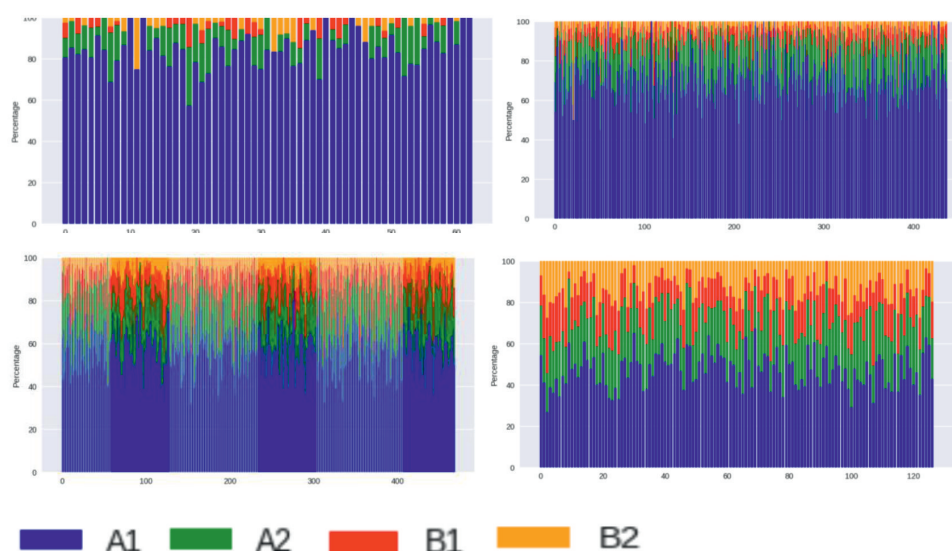


Рис. 2. Соответствие слов между реальными текстами (вверху слева – работы уровня A1, вверху справа – A2, внизу слева – B1, внизу справа – B2) и списком Келли

Fig. 2. Word match between real works (top left – A1 works, top right – A2, bottom left – B1, bottom right – B2) and Kelly's list

Остается вопрос: как можно выявить разницу в словарном запасе между разными уровнями, если списки не полностью коррелируют друг с другом и не отражают в полной мере фактически используемые студентами слова? Этот вопрос будет рассмотрен в следующем разделе.

3.3. Анализ главных компонент текстов студентов

Теоретически можно предположить, что лексические репертуары студентов настолько разнообразны, что невозможно обнаружить какие-либо связи. Однако это предположение выглядит невероятным, и в данном исследовании делается попытка доказать, что тексты, написанные студентами на одном и том же уровне владения, похожи по использованию лексики. Для проверки этой гипотезы был использован метод главных компонент (PCA) [см. Abdi и Williams, 2010], который, в целом, направлен на выявление кластеризации гетерогенных и многомерных данных. Данная гипотеза состоит в том, что студенты на уровне X не используют одни и те же слова, но их использование следует определенному тренду, который можно уловить с помощью PCA. Этот аналитический инструмент принципиально выгоден своей способностью интерпретировать данные, сохраняя максимальное количество информации, что впоследствии облегчает визуализацию многомерных данных. В соответствии с этой гипотезой текстовые композиции, написанные студентами на одном и том же уровне владения, демонстрируют сходство в использовании лексики. Однако это сходство многомерно, так как каждый студент неизбежно использует свой собственный набор слов. Таким образом, нет ни одной пары текстов, которые обнаруживают полное сходство, даже на начальном уровне.

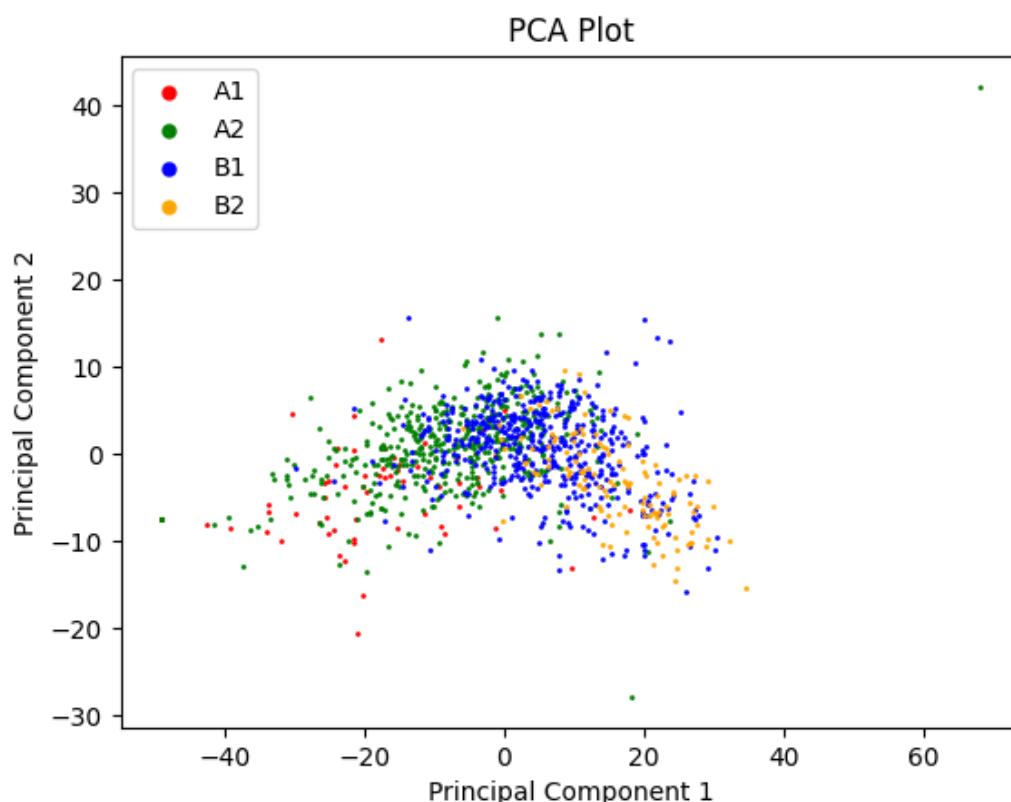


Рис. 3. Главные компоненты (PCA) для списка Келли
Fig.3. PCA for Kelly's list

Представленная на рис. 3 диаграмма PCA отображает разброс эссе и предполагает явное смещение в использовании лексики в этих эссе. Очевидно, что различные уровни владения, такие как A1 (обозначены красными точками) и B2 (обозначены желтыми точками), проявляют выраженное разделение с небольшим количеством выбросов. Тем временем уровни A1-A2, A2-B1 и B1-B2 демонстрируют последовательное перекрытие.

В целом результаты частично подтверждают исследовательскую гипотезу. Несмотря на то что график представляет собой континуум, было замечено явное различие между двумя наборами уровней владения (A1 и A2 против B1 и B2), но не было обнаружено четкого различия между A1 и A2 или между B1 и B2. Это указывает на необходимость дальнейшего исследования и изучения факторов, влияющих на использование лексики и уровни владения у студентов, изучающих русский как второй язык.

Анализ и результаты, представленные в этом разделе, способствуют пониманию оценки словаря и уровней владения языком среди студентов, изучающих русский как второй язык. Они подчеркивают важность учета нескольких показателей и ресурсов при оценке уровня владения языком и демонстрируют представления для разработки новых учебных материалов и ресурсов для изучения языка.

4. Обсуждение и перспективы

Сравнение списков показало слабые корреляции, кроме списков SMARTool и LM. Это находка вызывает удивление, так как базовая лексика должна быть относительно схожей в этих списках. Кроме того, отсутствие сильных корреляций указывает на то, что эти списки лексики могут быть не прямо связаны с текстами эссе, используемыми в исследовании, так как значительное количество слов в текстах не входит в эти списки, что подчеркивает необходимость комплексного подхода к оценке языковой компетенции, который включает различные лингвистические признаки, помимо словаря.

Анализ реальных студенческих эссе с использованием графиков PCA продемонстрировал возможность кластеризации, но только в том смысле, что можно различить уровни A и B. Создание отдельных кластеров для A1-A2, A1-B1 и B1-B2 при сравнении их с каждым из четырех списков оказалось невозможным, хотя можно сделать вероятностное предположение. Эти результаты указывают на то, что существующие списки не являются надежными инструментами для оценки языковой компетенции.

Включение педагогических соображений:

1. Роль списков лексики в обучении языку.

Несмотря на то что данное исследование вызывает вопросы о надежности списков лексики для оценки языковой компетенции, важно учитывать их роль в обучении языку. Списки лексики, подобные тем, которые были проанализированы, давно являются стандартом в обучении языку. Они служат руководствами для педагогов, помогая им разрабатывать учебные планы, выбирать учебные материалы и устанавливать учебные цели. Эти списки предоставляют структурированный каркас, который позволяет учащимся систематически расширять свой словарь.

2. Баланс между лексической компетенцией и навыками общения.

Одним из ключевых вызовов в обучении языку является поиск правильного баланса между лексической компетенцией и навыками общения. Результаты исследования показывают, что применение только списков лексики может быть недостаточным для точной оценки языковой компетенции. Это означает, что преподавателям языка следует применять более целостный подход к обучению.

3. Интеграция списков лексики в учебный план.

Несмотря на то что списки лексики плохо подходят для оценки языковой компетенции, преподавателям стоит продолжать использовать эти списки, например, в качестве справочных материалов при планировании уроков и оценок.

4. Будущие направления в методике обучения языку.

В будущих исследованиях планируется исследовать другие параметры, которые могут обеспечить более точную оценку уровней компетенции студентов. К таким параметрам можно отнести длину слова, количество морфем, количество слогов, разнообразие частей речи, использованных слов и другие. Создав меру, учитывающую каждый из этих параметров с определенными весами, следует стремиться получить многофакторную и, следовательно, более точную оценку языковой компетенции.

Исследования будут направлены на решение ограничений данной работы и расширение методов оценки языковой компетенции.

Хотя списки лексики сами по себе могут быть недостаточными для автоматизированной оценки языковой компетенции, нет никаких сомнений, что они остаются ценными инструментами для обучения языку. Преподавателям языка стоит продолжать использовать их как справочные материалы, внедряя при этом более холистический подход к обучению, который акцентирует практическое использование языка и навыки общения. Выводы этого исследования ценны для развития практик оценки языковой компетенции и учебных материалов для поддержки учащихся на пути к освоению языка

4.1. Ссылки и будущие исследования в области методики обучения

В будущих исследованиях важно глубже интегрировать педагогические соображения. Это может включать в себя следующее.

Исследование эффективных стратегий обучения лексике: изучение и разработка стратегий обучения, способствующих усвоению и применению лексики.

Оценка влияния списков лексики на обучение языку: проведение исследований, направленных на оценку того, как использование списков лексики влияет на результаты обучения языку. Это может включать в себя сравнение прогресса учащихся, которые интенсивно используют списки лексики, с теми, кто полагается на другие методы.

Интеграция технологий: исследование того, как технологии, такие как языковые приложения и онлайн-платформы, могут улучшить эффективность списков лексики в обучении языку. Это может включать в себя создание цифровых инструментов, которые динамически адаптируют списки лексики в зависимости от успехов учащегося.

Интегрируя эти педагогические соображения в будущие исследования, возможно в будущем создать инструмент, который позволит автоматически оценивать уровень знания языка и отдельные языковые параметры, что в конечном итоге будет полезно как преподавателям, так и учащимся.

Заключение

Цель настоящей работы заключалась в исследовании эффективности списков частотных слов и лексических минимумов в качестве надежных инструментов автоматизированной оценки языковой компетенции. Через всесторонний анализ выбранных списков лексики и их сравнение с фактическим использованием языка в текстах, написанных студентами, изучающими русский язык, данное исследование стало отправной точкой к дальнейшему развитию точных и всесторонних методов оценки языковой компетенции.

Результаты данного исследования показывают, что существующие списки лексики не являются полностью надежными для оценки языковой компетенции. Слабые корреляции между списками и их несоответствие фактическим текстам вызывают вопросы об их пригодности в качестве инструментов оценки. Анализ кластеризации с использованием PCA продемонстрировал потенциал различения между уровнями А и В, но оказался вызовом для более детализированных различий, таких как А1 и А2 или В1 и В2. Это указывает на необходимость более

тщательного подхода к подсчету и оценке словаря для достижения более точной оценки языковой компетенции.

Очевидно, что исключительно списки лексики недостаточны для автоматизированной оценки языковой компетенции. Вместо этого в дальнейшем необходимо исследовать альтернативные параметры и меры для улучшения процесса оценки. Интегрировав дополнительные лингвистические признаки, такие как длина слова, количество морфем, количество слогов и разнообразие частей речи, следует разработать более нюансный и всесторонний метод оценки. Эти исследования помогут выявить оптимальный способ подсчета и оценки словаря, что в конечном итоге приведет к улучшению инструментов оценки языковой компетенции.

Несмотря на то что исследование вносит свой вклад в понимание корреляций в словаре, оно подчеркивает необходимость многофакторного подхода, который учитывает различные лингвистические признаки. Проведение дальнейших исследований и усовершенствование методов оценки позволит лучше оснастить преподавателей и учащихся эффективными инструментами для более точной оценки языковой компетенции. Результаты и выводы данного исследования ценны для развития практик оценки языковой компетенции и учебных материалов для поддержки учащихся на пути к достижению компетенции в языке, так как позволяют оценить актуальность современных лексических списков для русского языка.

Список литературы

- Андрюшина Н. П., Козлова Т. В. Лексический минимум по русскому языку как иностранному. Элементарный уровень. Общее владение. 5-е изд. СПб.: Златоуст, 2015.
- Андрюшина Н. П., Козлова Т. В. Лексический минимум по русскому языку как иностранному. Базовый уровень. Общее владение. 5-е изд. СПб.: Златоуст, 2013.
- Андрюшина Н. П. и др. Лексический минимум по русскому языку как иностранному. Первый сертификационный уровень. Базовое владение. 7-е изд. СПб.: Златоуст, 2015.
- Андрюшина Н. П. и др. Лексический минимум по русскому языку как иностранному. Второй сертификационный уровень. Общее владение. 5-е изд. СПб.: Златоуст, 2015.
- Дайс Ли Р. Меры экологической ассоциации между видами. Экология, 1945. 26 (3): 297–302. [Online] DOI:10.2307/1932409. JSTOR 1932409
- Джанда Лора А. и Фрэнсис М. Тайерс. Меньше – значит лучше: почему все парадигмы дефектны, и почему это хорошо. Корпусная лингвистика и лингвистическая теория, 2018. 14(2), 33 стр. [Online] DOI.org/10.1515/cllt-2018-0031
- Килгаррифф А. и др. (Списки лексики на основе корпуса для изучающих язык в девяти языках. Языковые ресурсы и оценка, 2014. 48 стр. 121–163. [Online] DOI 10.1007/s10579-013-9251-2
- Лапошина А. Опыт экспериментального исследования сложности текстов по РКИ // Конференция «Динамика языковых и культурных процессов в современной России» [Электронный ресурс]. Вып. 6. Материалы VI Конгресса РОПРЯЛ (г. Уфа, 11–14 октября 2018 года) В: СПб.
- Лапошина А. Н., & Лебедева М. Ю. Текстометр: онлайн-инструмент определения уровня сложности текста по русскому языку как иностранному // Русская языковая студия, 2021. 19(3), 331–345. [Online] DOI: 10.22363/2618-8163-2021-19-3-331-345
- Лу Ши. Автоматический анализ синтаксической сложности вторичного языка // Международный журнал корпусной лингвистики, 2010. 15(4), 474–496.
- Луман Дж., Кэмпбелл Дж. Б. Адаптация К Соренсена (1948) для оценки аффинитета элементов в степных растительных сообществах // Экология. 1960, 41 (3): 409–416. DOI:10.2307/1933315. JSTOR 1933315

- Ляшевская О. Н., и др.** Оценка методов автоматического анализа текста: морфологические парсеры русского языка // Вычислительная лингвистика и интеллектуальные технологии. Труды ежегодной международной конференции «Диалог», 2010. Т. 9. Вып. 16. С. 318–326. М.: РГГУ.
- Ляшевская О. Н., Шарофф С. А.:** Частотный словарь современного русского языка (на основе материалов Национального корпуса русского языка). М.: Азбуковник, 2009.
- Совет Европы.** Общеевропейские рекомендации по владению иностранными языками: схема. [Online] Доступно: <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>
- Соренсен Т.** Метод формирования групп с равными амплитудами в растительной социологии на основе сходства видов и его применение для анализа растительности датских общин // Kongelige Danske Videnskabernes Selskab, 1948. 5 (4): 1–34.
- Чирнер Э., Беренфенгер О., Ваннер И.** Оценка доказательств обоснованности присвоения рейтингов CEFR собеседованию по устной речи ACTFL (OPI) и собеседованию по устной речи с помощью компьютера (OPIc) // Лейпциг: Университат Лейпциг, Институт Гердера. 2012.
- American Council on the Teaching of Foreign Languages (ACTFL)** proficiency guidelines, [Online] Доступно: <https://www.actfl.org/resources/actfl-proficiency-guidelines-2012>
- Abdi H., Lynne J. Williams.** Principal component analysis // Wiley interdisciplinary reviews: computational statistics 2, no. 4 (2010): 433–459.
- Hawkins J. A., & Buttery P.** Criterial features in learner corpora: Theory and illustrations // English Profile Journal, 2010. 1, e5.
- Bast R., et al.** (2018–2023). SMARTool. [Online] Доступно: <https://smartool.github.io/smartool-rus-eng/> (дата обращения: 26.06.2023).
- Callies M., & Götz S.** Learner corpora in language testing and assessment: Prospects and challenges. Learner corpora in language testing and assessment, 2015. 1–9.
- Clancy S.** (2014–2023). Visualizing Russian. [Online] Доступно: <https://digitalhumanities.fas.harvard.edu/project/visualizing-russian/> (дата обращения: 26.06.2023).
- Ondov Brian D. et al.** Mash: fast genome and metagenome distance estimation using MinHash // Genome biology 17.1 (2016): 1–14.
- UDPipe 1 Models.** 2021. <<https://ufal.mff.cuni.cz/udpipe/1/models>> (дата обращения: 30.06.2021).

References

- Abdi, Herve, and Lynne J. Williams.** “Principal component analysis.” Wiley interdisciplinary reviews: computational statistics 2, no. 4 (2010): 433–459.
- American Council on the Teaching of Foreign Languages (ACTFL)** proficiency guidelines, [Online] Available: <https://www.actfl.org/resources/actfl-proficiency-guidelines-2012>
- Bast, R., et al.** (2018–2023). SMARTool. [Online] Available: <https://smartool.github.io/smartool-rus-eng/>. (Retrieved June 26, 2023).
- Callies, M., & Götz, S.** (2015). Learner corpora in language testing and assessment: Prospects and challenges. Learner corpora in language testing and assessment, 1-9.
- Clancy, S.** (2014–2023). Visualizing Russian. [Online] Available: <https://digitalhumanities.fas.harvard.edu/project/visualizing-russian/>. (Retrieved June 26, 2023).
- Common European Framework of Reference for Languages (CEFR, Council of Europe)** proficiency scale, [Online] Available: <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>
- Dice, Lee R.** (1945). “Measures of the Amount of Ecologic Association Between Species”. Ecology. 26 (3): 297–302. [Online] doi:10.2307/1932409. JSTOR 1932409.

- Golubeva, A.** (2015). *Leksičeskij minimum po russkomu kak inostrannomu* [Lexical Minimum for Russian as a Foreign Language]. Zlatoust: Saint Petersburg.
- Hawkins, J. A., & Buttery, P.** (2010). Criterial features in learner corpora: Theory and illustrations. *English Profile Journal*, 1, e5.
- Lyashevskaya, O.N., Sharoff, S.A.**: Frequency Dictionary of Modern Russian Language (based on materials from the National Corpus of the Russian Language). M.: Azbukovnik, 2009.
- Janda, Laura A. & Francis M. Tyers.** 2018. Less is More: Why All Paradigms are Defective, and Why that is a Good Thing. *Corpus Linguistics and Linguistic Theory* 14(2), 33pp. [Online] doi org/10.1515/cllt-2018-0031
- Kilgariff, A., et al.** (2014). Corpus-based vocabulary lists for language learners for nine languages. *Language Resources and Evaluation*, 48, pp. 121–163. [Online] DOI 10.1007/s10579-013-9251-2
- Laposhina, A.** (November 2018). Opyt éksperimental'nogo issledovaniya slozhnosti tekstov po RKI [Experience of Experimental Research of Text Complexity in RFL]. Conference: Dinamika yazykovykh i kul'turnykh protsessov v sovremennoy Rossii [Electronic resource]. Vyp. 6. Materialy VI Kongressa ROPRYAL (g. Ufa, 11–14 oktyabrya 2018 goda) At: SPB.
- Laposhina, A. N., & Lebedeva, M. Y.** (2021). Tekstometr: onlayn-instrument opredeleniya urovnya slozhnosti teksta po russkomu yazyku kak inostrannomu [Textometr: online tool for determining the complexity level of texts in Russian as a foreign language]. *Russian Language Studies*, 19(3), 331–345. [Online] DOI: 10.22363/2618-8163-2021-19-3-331-345.
- Looman, J.; Campbell, J.B.** (1960). "Adaptation of Sorensen's K (1948) for estimating unit affinities in prairie vegetation". *Ecology*. 41 (3): 409–416. doi:10.2307/1933315. JSTOR 1933315.
- Lu, X.** (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics*, 15(4), 474–496.
- Lyashevskaya, O. N., et al.** (2010). Otsenka metodov avtomaticheskogo analiza teksta: morfologicheskie parsery russkogo yazyka [Assessment of Text Analysis Methods: Morphological Parsers of the Russian Language]. *Computational Linguistics and Intellectual Technologies. Proceedings of the Annual International Conference "Dialogue"* (2010), Volume 9, Issue 16, pp. 318–326. Moscow: RSUH.
- Ondov, Brian D., et al.** "Mash: fast genome and metagenome distance estimation using MinHash." *Genome biology* 17.1 (2016): 1–14.
- Sorensen, T.** (1948). "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". *Kongelige Danske Videnskabernes Selskab*. 5 (4): 1–34.
- Tschirner, Erwin, Bärenfänger, Olaf & Wanner, Irmgard.** 2012. Assessing Evidence of Validity of Assigning CEFR Ratings to the ACTFL Oral Proficiency Interview (OPI) and the Oral Proficiency Interview by Computer (OPIc). Leipzig: Universität Leipzig, Herder-Institut.
- UDPipe 1 Models.** 2021. <<https://ufal.mff.cuni.cz/udpipe/1/models>> (30 June 2021).

Информация об авторе

Вахранев Антон Юрьевич, аспирант Национального исследовательского университета «Высшая школа экономики»

Information about the Author

Anton Y. Vakhranov, PhD postgraduate student, National Research University Higher School of Economics, Moscow, Russian Federation

Статья поступила в редакцию 30.09.2023;

одобрена после рецензирования 23.04.2024; принята к публикации 26.04.2024

The article was submitted 30.09.2023;

approved after reviewing 23.04.2024; accepted for publication 26.04.2024