



Научно-исследовательский журнал «Вестник филологических наук / Philological Sciences Bulletin»

<https://vfn-journal.ru>

2024, Том 4, № 1 / 2024, Vol. 4, Iss. 1 <https://vfn-journal.ru//archives/category/publications>

Научная статья / Original article

Шифр научной специальности: 5.9.8. Теоретическая, прикладная и сравнительно-сопоставительная лингвистика (филологические науки)

УДК 81'33

¹ Марико М.Л.

¹ Казанский (Приволжский) федеральный университет

Сопоставительное исследование учебников по обществознанию и документов ООН на русском языке

Аннотация: в данном исследовании мы анализировали корпус из 60 000 слов, представленный русскими текстами ООН и четырьмя (4) комплектами школьных учебников (6 и 9 классы) по обществознанию, написанных [2, с. 63-88] и [10, с. 240]. Двадцать (20) текстов были извлечены из документов Организации Объединенных Наций и сорок (40) текстов – из учебников по обществознанию. Объем выборки не превышал 1000 слов в каждом тексте. Цель исследования – выявить корреляцию между документами ООН и учебниками по обществознанию Никитина и Боголюбова. Для достижения цели мы использовали Rulingva, вычислительный инструмент, доступный в Интернете, чтобы проанализировать данные, выбрав уровни Флеша-Кинкейда, лексическую плотность и лексическое разнообразие (TTR). Результаты показали, что по уровню читабельности (FKGL) 6-е классы идут практически параллельно с 9-ми классами; тексты варьировались от 7,95 до 9,37. FKGL русских текстов ООН варьировались от 11,75 до 16,49, поэтому оказались более сложными, чем учебники по обществознанию. Кроме того, лексическая плотность образцов, взятых из русских текстов ООН и учебников по обществознанию, была очень похожа и составляла от 66% до 74%; более высокий процент лексической плотности является показателем понятности. Аналогичным образом, лексическое разнообразие (TTR) всего корпуса варьировалось от 0,55 до 0,68; более высокое значение TTR также является показателем понимания текста. Мы пришли к выводу, что существует корреляция между русскими текстами ООН и учебниками по обществознанию Никитина и Боголюбова по показателям лексической плотности и лексического разнообразия (TTR).

Ключевые слова: читабельность (FKGL), лексическое разнообразие (TTR), лексическая плотность, русские тексты ООН, учебные материалы, Rulingva, сложность текста

Для цитирования: Марико М.Л. Сопоставительное исследование учебников по обществознанию и документов ООН на русском языке // Вестник филологических наук. 2024. Том 4. № 1. С. 50 – 58.

Поступила в редакцию: 13 октября 2023 г.; Одобрена после рецензирования: 14 декабря 2023 г.; Принята к публикации: 9 февраля 2024 г.

¹ Mariko M.L.

¹ Kazan (Volga Region) Federal University

A comparative study of textbooks on social studies and Russian UN texts

Abstract: in this research we examined a 60.000 word-based corpus represented by Russian UN texts and four (4) sets of school textbooks (6th grades and 9th grades) on social studies written by [2, p. 63-88] and [10, p. 240]. Twenty (20) texts were extracted from the documents of the United Nations and forty (40) texts from the textbooks on social studies. The samples were limited to 1000 words per text. The purpose of the research was to discover the correlation between the documents of the United Nations and the textbooks on social studies by Nikitin and Bogolubov. To meet our target, we used Rulingva, a computational tool available online, to analyze the data by selecting Flesch Kincaid Grade Level, Lexical density, and lexical diversity (TTR). The findings revealed that the 6th grades run in parallel with the 9th grades in terms of readability (FKGL); the texts varied between 7.95 and 9.37.

FKGL of Russian UN texts varied from 11.75 to 16.49, thus proved to be more complex than the textbooks on social studies. Furthermore, the lexical density of the samples extracted from Russian UN texts and textbooks on social studies were very similar ranging from 66 % to 74%; a higher percentage of lexical density is an indication of understandability. Likewise, the lexical diversity (TTR) of the whole corpus varied from 0.55 to 0.68; a higher value of TTR is also an indication of text comprehension. We concluded that there is a correlation between Russian UN texts and the textbooks on social studies by Nikitin and Bogolubov in terms of lexical density and lexical diversity (TTR).

Keywords: readability (FKGL), lexical diversity (TTR), lexical density, Russian UN texts, educational textbooks, Rulingva, text complexity

For citation: Mariko M.L. A comparative study of textbooks on social studies and Russian UN texts. *Philological Sciences Bulletin*. 2024. 4 (1). P. 50 – 58.

The article was submitted: October 13, 2023; Approved after reviewing: December 14, 2023; Accepted for publication: February 9, 2024.

Введение

Некоторые тексты читаются легко. Другие – трудно. Это совершенно естественно. Сложность заключается в том, чтобы разработать объективное средство измерения текстов по тому, как они читаются. Это одна из задач, которая послужила мотивом для нашего исследования. Таким образом, мы можем рассматривать русские тексты ООН и учебники по обществознанию как публичные документы. Сложность чтения многих публичных документов слишком высока для значительной части населения. Например, юридические документы, медицинские документы и трудовые договоры также являются прекрасными примерами сложных текстов, которые трудно понять большинству населения.

Аналогичным образом, анкеты и опросы содержат большой процент вопросов, которые вызывают трудности с пониманием у значительной части населения [3, с. 335; 6, с. 3-22]. Надежность и достоверность результатов этих опросов ставится под сомнение, когда вопросы содержат сложные слова, двусмысленное значение, сложный синтаксис или содержание, которое чрезмерно нагружает когнитивные ресурсы. В результате страдают и люди, и общество. Стоит отметить, что русские тексты ООН и общественные документы единично посвящены теме прав человека, что делает их специфическими. Таким образом, эти тексты также могут относиться к публичным документам или рассматриваться как юридические документы, политические документы.

Под трудностью понимания понимается то, насколько читабельность текста соответствует уровню понимания текста отдельными читателями [1, с. 416]. Мы можем определить читабельность на основе таких характеристик, как количество слогов в слове, количество слов в предложении, сложность слов и сложность языка [8, с. 12-15]. Изначально формулы читабельности были разработаны для учебников начальной школы, чтобы убедиться, что книги не слишком сложны для восприятия детьми [4, с. 74]. Одним из первых инструментов читабельности, созданных для этой цели, была формула Flesch Reading Ease (FRE) [5, с. 221-233]. Эта формула основана на 100-балльной системе, где более высокий балл показывает, что текст легче для понимания. В 1975 году эта мера была дополнена, в результате чего появилась формула Flesch-Kincaid Reading Grade Level (FKGL) [11, с. 149-154]. Формула FKGL для определения читабельности является, вероятно, наиболее широко используемой в различных дисциплинах и оценивает текст на основе уровня школьного образования в США в диапазоне от 1 до 12. Формула, используемая для определения FKGL, основана на двух основных критериях. Первый – синтаксическая сложность, которая измеряется путем оценки количества слов в предложении. Сложность слова – второй компонент, количественно измеряется количеством слогов в слове [11, с. 149-154].

Показатели лексического разнообразия предположительно связаны как с трудностью текста, так и с его связностью. Лексическое разнообразие увеличивает сложность, поскольку каждое уникальное слово представляет новую информацию, которую необходимо закодировать и интегрировать в контекст дискурса. Наиболее известным показателем лексического разнообразия является соотношение между словами (type-token ratio, TTR), [12, с. 183]. Это количество уникальных слов в тексте (т. е. типов), деленное на общее количество слов (т. е. токенов) в тексте.

Лексическая плотность – часть лингвистики, изучающая устную и письменную речь. Лексическая плотность — это соотношение лексических единиц (слов содержания) к общему объему дискурса [7, с. 109]. Лексическая плотность также может быть определена путем деления количества лексических единиц (слов

содержания) в сложной клаузе на количество клауз в комплексе. Лексическая плотность показывает сложность слов в тексте. Она также описывает развитие лексических единиц в письменном языке.

Цель исследования является изучение сложности русского языка (учебники по обществознанию) и текстов Организации Объединенных Наций.

Материалы и методы исследований

Мы отобрали данные из русских текстов ООН, размещенных на сайте ООН в разделе (https://www.ohchr.org/en/ohchr_homepage), а также из четырех (4) комплектов школьных учебников (6 и 9 классы) по обществознанию, написанных [2, с. 63-88] и [10, с. 240]. Учебники Никитина и Боголюбова рекомендованы Министерством образования Российской Федерации и предназначены для учащихся 6 и 9 классов. После отбора данных мы использовали бесплатный онлайн-вычислительный инструмент *Rulingva* (<https://Rulingva.kpfu.ru/>) для анализа корпуса, который включал 60.000 слов, то есть 10.000 слов из каждого учебника и 20.000 слов из русских текстов ООН. Таким образом, мы создали 60 текстов, объемом 1000 слов на текст.

Тем не менее, мы стремились выявить сходства и различия между русскими текстами ООН и текстами по обществознанию Никитина и Боголюбова. Для достижения поставленной цели мы выбрали показатели Flesch-Kincaid Grade Level (*FKGL*), лексической плотности и лексической плотности (*TTR*). Результаты представлены на рисунках ниже, после чего следует обсуждение, которое мы представили в шесть (6) этапов.

Результаты и обсуждения

В данном исследовании мы использовали для анализа данные из документов ООН и четырех (4) учебников по обществознанию 6 и 9 классов Никитина и Боголюбова. После отбора данных мы использовали программу *Rulingva* для анализа корпуса, который состоял из 60 000 слов, то есть 10 000 слов из каждого учебника и 20 000 слов из русских текстов ООН. Таким образом, мы создали 60 текстов по 1000 слов на каждый текст. Нашей целью было выявить и найти сходства и различия между документами ООН и учебниками по обществознанию Никитина и Боголюбова. Поэтому для определения сложности текста мы выбрали Flesch-Kincaid Grade Level (*FKGL*), лексическую плотность и лексическое разнообразие (*TTR*).

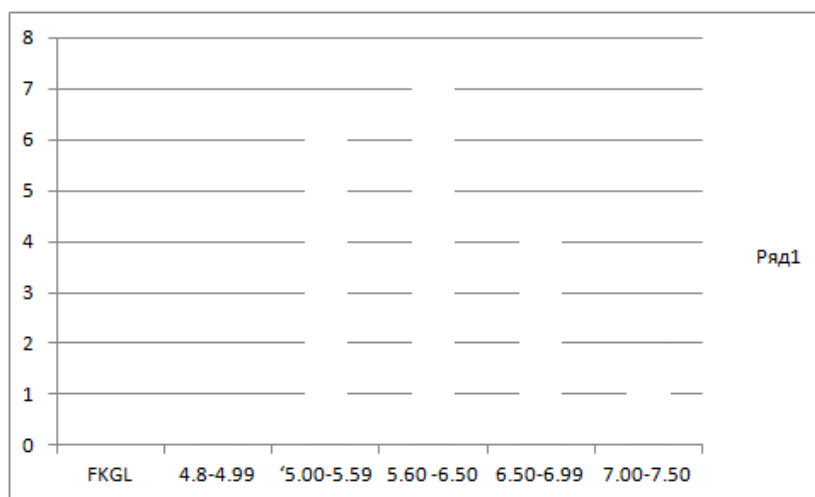


Рис. 1. Учебники 6 класс _Никитин и Боголюбов (Flesh-Kincaid Grade Level).

Fig. 1. Textbooks 6th grade Nikitin and Bogolyubov (Flesch-Kincaid Grade Level).

На этом графике показан уровень владения *FKGL* по учебникам 6 класса Никитина и Боголюбова. Показатели варьируются от 5,00 до 7,50, что говорит о том, что для понимания этих текстов учащимся необходимо не менее 5 лет образования.

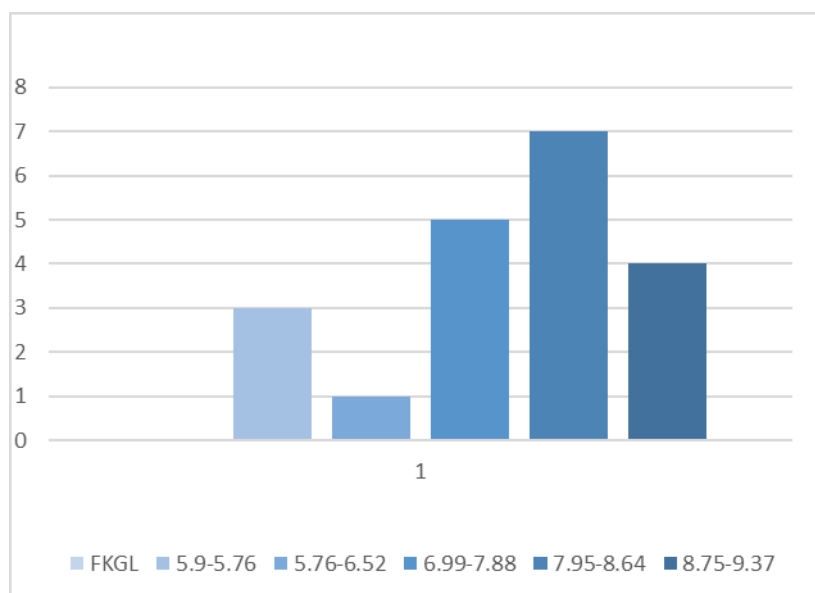


Рис. 2. Учебник 9 класс _Никитин и Боголюбов (Flesh-Kincaid Grade Level).
Fig. 2. Textbook 9th grade Nikitin and Bogolyubov (Flesch-Kincaid Grade Level).

На втором графике видно, что Flesch Kincaid Grade Level (FKGL) учебников 9 класса Никитина и Боголюбова находится в пределах требуемого уровня для этих учебников, то есть почти 5 текстов из 10 варьируются от 5,76 до 7,88. кроме того, 7 текстов из 10 варьируются от 7,95 до 9,37.

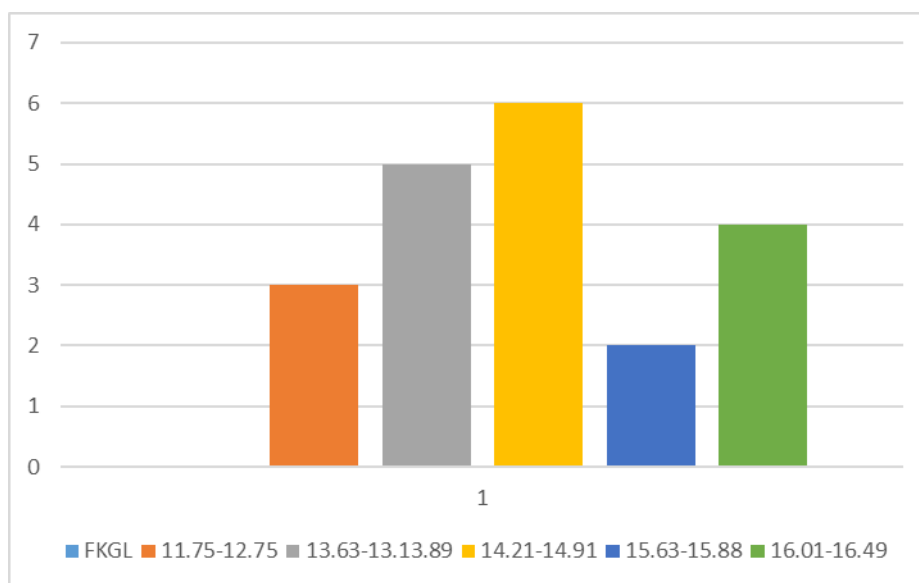


Рис. 3. Русские тексты ООН_ Flesch Kincaid Grade Level (FKGL).
Fig. 3. Russian texts UN_ Flesch Kincaid Grade Level (FKGL).

График показывает, что Flesch Kincaid Grade Level (FKGL) текстов Russian UN варьируется между 11.75 и 16.49.

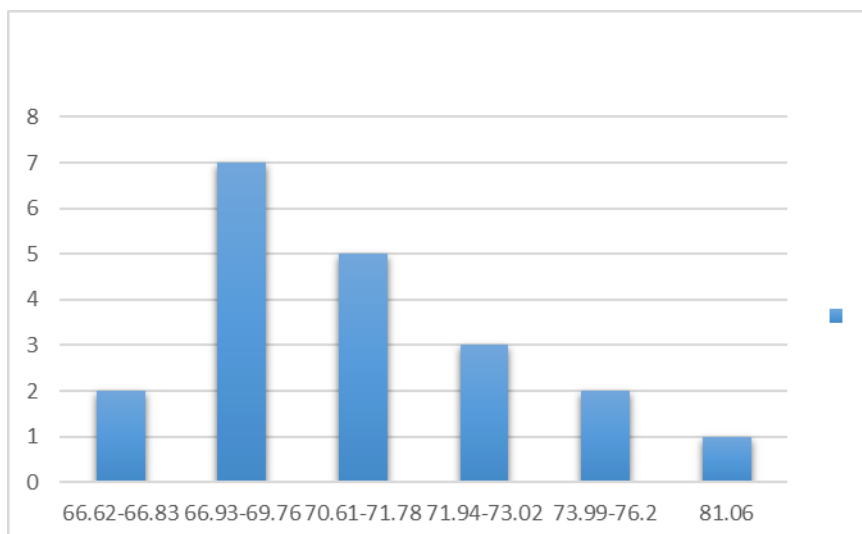


Рис. 4. Учебники_Никитин+Боголюбов 9 класс разделены на группы (Лексическая плотность).
Fig. 4. Textbooks_Nikitin+Bogolyubov grade 9 are divided into groups (Lexical density).

График показывает, что тексты 9-х классов Никитина и Боголюбова варьируются между 66% и 81%. 7 текстов из 10 варьируются от 66,93 до 71%.

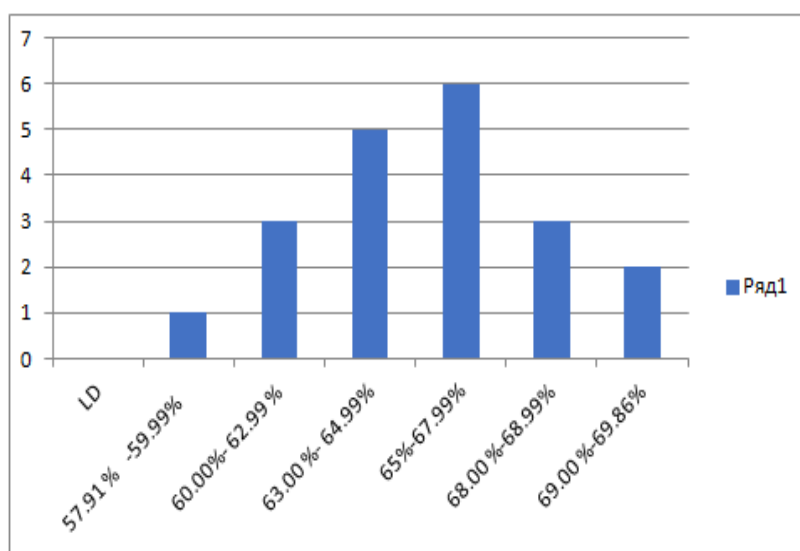


Рис. 5. Учебники_Никитин+Боголюбов разделены на группы.
Fig. 5. Textbooks_Nikitin + Bogolyubov are divided into groups.

На графике показана лексическая плотность текстов Никитина и Боголюбова для обоих 6-х классов. Тексты варьируются между 60 и 70%, что свидетельствует о том, что тексты не очень сложны для понимания шестиклассниками.

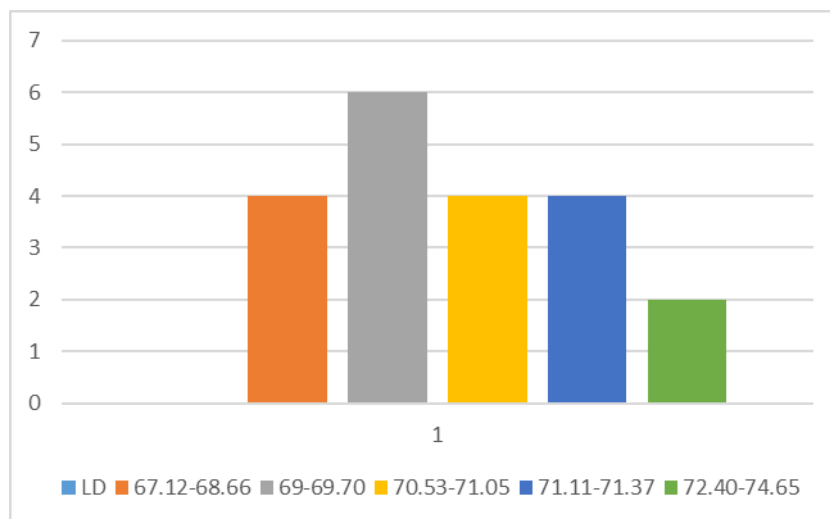


Рис. 6. Русские тексты ООН_ Лексическая плотность.
Fig. 6. Russian UN texts_ Lexical density.

На графике видно, что лексическое разнообразие вычисленных текстов варьируется между 67 и 74.

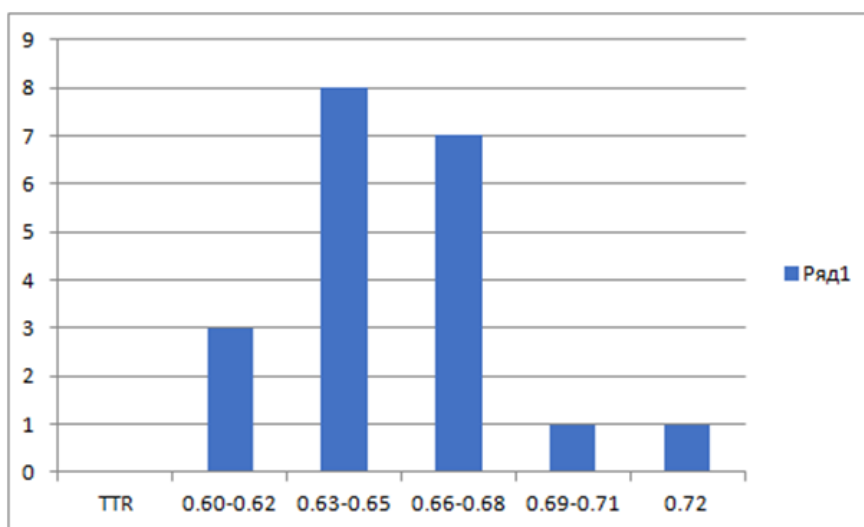


Рис. 7. Учебники_Никитина+Боглюбова (TTR) разделены на группы.
Fig. 7. Textbooks_Nikitin+Boglyubov (TTR) are divided into groups.

На графике представлен Type Token Ratio (TTR) учебников 6-го класса Никитина и Боголюбова. 8 текстов из 10 варьируются от 0,63 до 0,68, что свидетельствует о том, что эти тексты легче читаются шестиклассниками.

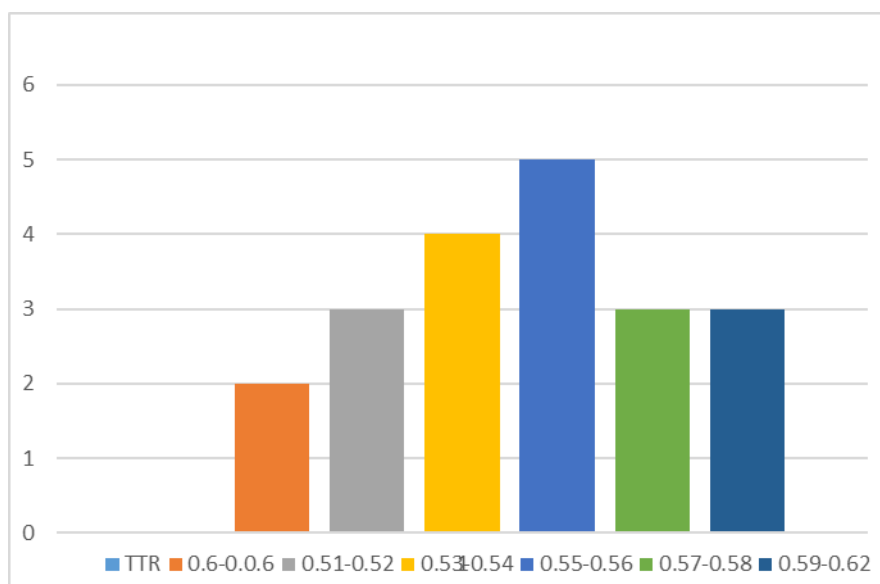


Рис. 8. Русские тексты ООН _Лексическое разнообразие (TTR).
Fig. 8. Russian UN texts _Lexical diversity (TTR).

График показывает лексическое разнообразие (*TTR*) исследуемых текстов, которое варьируется между 0,6 и 0,62. Он также показывает, что соотношение типов лексем (*TTR*) русских текстов ООН варьируется между 0,55 и 0,56.

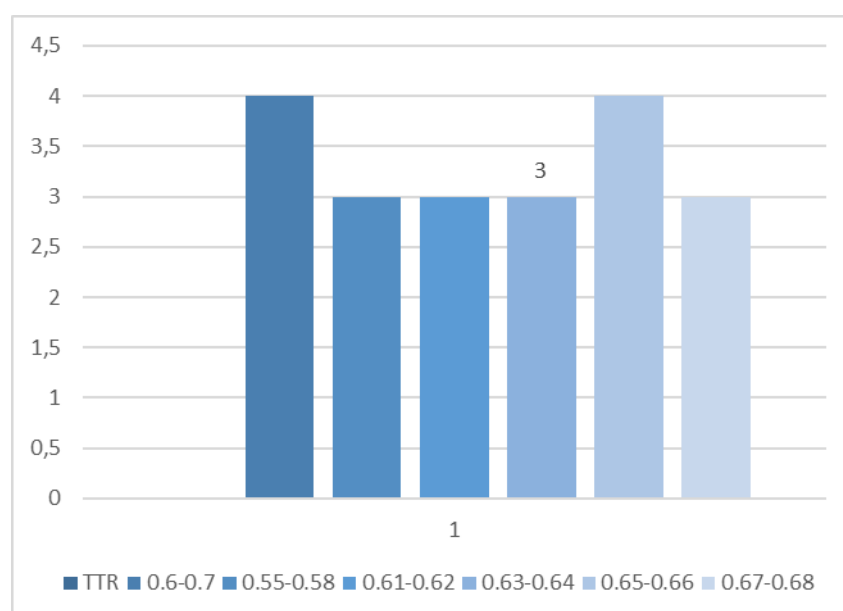


Рис. 9. Никитин+Боголюбов разделены на группы.
Fig. 9. Nikitin + Bogolyubov are divided into groups.

На графике представлен Type Token Ratio (*TTR*) учебников 9 класса Никитина и Боголюбова. Огромное количество текстов варьируется от 0,58 до 0,68, что говорит о том, что процентное соотношение текстов выше среднего, а значит, эти тексты легче читаются девятиклассниками.

На первом этапе исследования мы вычислили 20 текстов, отобранных из учебников для 6-х классов с помощью *Rulingva*, сравнив результаты, полученные по Flesch-Kincaid Grade Level (*FKGL*) в текстах, и выяснили, что 11 текстов из 20 варьировались от 5,00 до 6,99, как видно на первом графике выше. Это означает, что большая часть текстов относится к 6 классу. Только 2 текста относятся к 7 классу.

На втором этапе исследования мы проанализировали учебники Никитина и Боголюбова для 9 класса. На втором графике выше показано, что Flesch Kincaid Grade Level (*FKGL*) по учебникам Никитина и Боголюбова

бова для 9 класса варьировались от 5,76 до 9,37. Эти показатели означают, что тексты соотносятся с 9-м классом.

На третьем этапе нашего исследования мы сравнили данные, отобранные из учебников 6 и 9 классов по обществознанию Никитина и Боголюбова и русских текстов ООН. Первый параметр - *FKGL* - показал, что тексты варьируются от 11,75 до 16,49. Обоснование этого параметра указывает на то, что для понимания русских текстов ООН ученику требуется не менее 11 лет формального образования; 16,49 - самое высокое показание – является явным доказательством того, что русские тексты ООН более трудны для чтения, чем учебники по обществознанию. С другой стороны, данные 9-х классов совпали с требуемым уровнем читабельности, то есть 11 текстов из 20 оказались в диапазоне от 7,95 до 9,37. Кроме того, из текстов 6-го класса 11 текстов из 20 варьировались от 5,00 до 6,99, как видно на первом графике выше. Это предполагает, что большая часть текстов шла параллельно с 6-м классом. А также объясняет, что учащиеся вряд ли столкнутся с трудностями при чтении таких текстов. Аналогичным образом мы рассмотрели лексическую плотность, что позволило нам увидеть огромное сходство между учебниками 6-го, 9-го классов и русскими текстами ООН, так как в большинстве текстов варьировались от 66,62 до 74,65. Следовательно, тексты являются лексически плотными, что может облегчить их понимание.

На четвертом этапе исследования мы также провели сравнение результатов, полученных по лексической плотности в данных. Результаты, представленные на пятом (5) графике, показали, что в большинстве текстов варьировались между 63% и 68%, а это значит, что учебники по обществознанию Никитина и Боголюбова являются лексически плотными для понимания.

На пятом этапе исследования мы аналогичным образом измерили лексическое разнообразие (*TTR*) наших данных. Результаты, представленные на седьмом графике, показали, что большая часть текстов варьировались от 0,60 до 0,68. Это демонстрирует, что тексты похожи друг на друга и не являются трудными для чтения для соответствующего класса (6 класс), так как самый высокий традиционный коэффициент (*TTR*) равен 1, согласно [8, с.12-15].

На заключительном этапе исследования мы проанализировали лексическое разнообразие текстов, известное как соотношение типов лексем (*TTR*). В большинстве текстов соотношение было практически одинаковым. Они варьировались от 0,6 до 0,68. Это предполагает повторение одних и тех же слов в текстах, чтобы помочь читателю понять текст.

Выводы

В заключение следует отметить, что в данном исследовании была предпринята попытка выявить яркие особенности, существующие между русскими текстами ООН и учебниками по обществознанию Никитина и Боголюбова. Данные, отобранные из учебников по обществознанию, предназначенных для школьников, и данные, извлеченные из русских текстов ООН, показали, что многие тексты на русском языке относительно не отличаются друг от друга по лексической плотности и лексическому разнообразию (*TTR*); разница заключается в читабельности, так как *FKGL* русских текстов ООН варьируется между 11,75 и 16,49, что показывает, что читателю требуется не менее 11 лет формального образования, чтобы прочитать данные тексты (русские тексты ООН). Эти параметры могут указывать на сложность текста, как было показано в данном исследовании.

Стоит также отметить, что хотя русские тексты ООН обычно читают политики, данное исследование показало, что ими могут пользоваться девятиклассники, так как существует корреляция между текстами по выделенным параметрам. Лексическая плотность показывает сложность слов в тексте. Она также описывает развитие лексических единиц в письменном языке. Лексическое разнообразие также известно как лексическое богатство, поэтому оно увеличивает сложность, так как каждое отдельное слово представляет новую информацию, которую необходимо закодировать и интегрировать в контекст дискурса. Мы полагаем, что данное исследование может помочь специалистам в области образования лучше выбирать учебники, предназначенные для студентов.

Список источников

1. Conrad F.G., Schober M.F. (Eds.). *Envisioning the survey interview of the future*. New York: Wiley, 2007. 416 p.
2. Benjamin R.G. Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty // *Educational Psychology Review*. 2012. № 24 (1). P. 63 – 88.
3. Bogolubov L.N. *Social Studies. 11th grade (basic level) . Textbook*. 3rd ed. Moscow, 2017. 335 p.
4. DuBay W.H. *The principles of readability*. Costa Mesa: Impact Information, 2004. 74 p.

5. Flesch R. A new readability yardstick // Journal of Applied Psychology. 1948. № 32 (3). P. 221 – 233.
6. Graesser A.C., Cai Z., Louwerse M., Daniel F. (2006) Question Understanding Aid (QUAID): A web facility that helps survey methodologists improve the comprehensibility of questions // Public Opinion Quarterly. 2006. № 70. P. 3 – 22.
7. Halliday M.A.K. Spoken and written language. Geelong Victoria: Deakin Univ. Press, 1985. 109 p.
8. McClure G.M. Readability formulas: Useful or useless? // IEEE Transactions on Professional Communication, PC. 1987. № 30 (1). P. 12 – 15.
9. Malvern D., Richards B., Chipere N. et al. Lexical diversity and language development: quantification and assessment. Basingstoke, UK: Palgrave Macmillan, 2004. 288 p.
10. Nikitin A.F., Gribanova G.I., Skorobogatko A.V., Martyanov D.S Social Studies. Textbook. Drofa, 2017. 240 p.
11. Thomas G., Hartley R.D., Kincaid J.P. Test-retest and inter-analyst reliability of the Automated Readability Index, Flesch Reading Ease Score, and the Fog Count // Journal of Reading Behavior. 1975. № 7 (2). P. 149 – 154.
12. Templin M.C. Certain Language Skills in Children: Their Development and Interrelationships, Vol. 10. Minneapolis, MN: University of Minnesota Press, 1957. 183 p.

References

1. Conrad F.G., Schober M.F. (Eds.). Envisioning the survey interview of the future. New York: Wiley, 2007. 416 p.
2. Benjamin R.G. Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. Educational Psychology Review. 2012. № 24 (1). P. 63 – 88.
3. Bogolubov L.N. Social Studies. 11th grade (basic level) . Textbook. 3rd ed. Moscow, 2017. 335 p.
4. DuBay W.H. The principles of readability. Costa Mesa: Impact Information, 2004. 74 p.
5. Flesch R. A new readability yardstick. Journal of Applied Psychology. 1948. № 32 (3). P. 221 – 233.
6. Graesser A.C., Cai Z., Louwerse M., Daniel F. (2006) Question Understanding Aid (QUAID): A web facility that helps survey methodologists improve the comprehensibility of questions. Public Opinion Quarterly. 2006. № 70. P. 3 – 22.
7. Halliday M.A.K. Spoken and written language. Geelong Victoria: Deakin Univ. Press, 1985. 109 p.
8. McClure G.M. Readability formulas: Useful or useless? IEEE Transactions on Professional Communication, PC. 1987. № 30 (1). P. 12 – 15.
9. Malvern D., Richards B., Chipere N. et al. Lexical diversity and language development: quantification and assessment. Basingstoke, UK: Palgrave Macmillan, 2004. 288 p.
10. Nikitin A.F., Gribanova G.I., Skorobogatko A.V., Martyanov D.S Social Studies. Textbook. Drofa, 2017. 240 p.
11. Thomas G., Hartley R.D., Kincaid J.P. Test-retest and inter-analyst reliability of the Automated Readability Index, Flesch Reading Ease Score, and the Fog Count. Journal of Reading Behavior. 1975. № 7 (2). P. 149 – 154.
12. Templin M.C. Certain Language Skills in Children: Their Development and Interrelationships, Vol. 10. Minneapolis, MN: University of Minnesota Press, 1957. 183 p.

Информация об авторе

Марико М.Л., преподаватель-исследователь, Казанский (Приволжский) федеральный университет

© Марико М.Л., 2024