

В связи с растущей популярностью изучения русского языка и компьютерных технологий, оказывающих сильное влияние на образование, существует необходимость в программе, которая создает задания различной сложности на основе любых загруженных текстов на русском языке. Именно этим фактом обуславливается актуальность данной работы.

Разработанная компьютерная программа позволяет автоматически генерировать несколько типов упражнений для изучения русского языка как иностранного. Возможность пользователя данной программы самому выбирать текст для будущих заданий имеет несколько преимуществ. Во-первых, человек, изучающий иностранный язык может сам выбирать уровень сложности текстов. Во-вторых, пользователь может загружать тексты определенной тематики, что способствует улучшению не только тренируемых языковых навыков, но и пополнению словарного запаса по отдельным темам. И в-третьих, использование программы повышает интерес к изучению языка и делает его более легким и интересным, так как человек сам генерирует задания и выбирает из них наиболее подходящие для себя.

Разработанная компьютерная программа написана на языке программирования Python с использованием двух внешних библиотек по автоматической обработке текстов: NLTK (<http://www.nltk.org>) и `pymorphy2` (<http://pymorphy2.readthedocs.io/en/latest/>). Библиотека NLTK была использована для деления предложений и токенизации в нескольких типах упражнений, библиотека `pymorphy2` – для морфологического анализа слов.

В разработанной нами программе генерируется несколько типов упражнений:

- определение рода существительных;
- постановка глаголов в правильную временную форму;
- выбор прилагательного в нужном роде и числе;
- разделение слитного текста;
- поиск ошибок в написании отдельных слов.

## Оценка результатов

Одним из этапов разработки данной программы было её тестирование. Случайным образом выбирались 5 текстов (длина от 200 до 300 слов) разной тематики, которые загружались в программу. На выходе генерировались упражнения с ответами. Основная задача данного этапа заключалась в максимально точной и объективной оценке создаваемых упражнений. Результаты получились следующими:

1. В задании на определение рода отдельных слов доля правильных пропусков от общего числа пропусков оказалась равна 87%;
2. В упражнениях на постановку глагола во временную форму, а также выбор числительных, прилагательных и местоимений в нужном роде, числе и падеже процент правильных пропусков, генерируемых программой, был равен 96%;
3. Задание, в котором необходимо разделить слитный текст, расставив пробелы и знаки препинания, при каждой проверке генерировалось без ошибок;
4. В последнем типе упражнения, касающемся правильного написания слов, процент слов с ошибками от общего числа слов в тексте равен 6%.

Таким образом, оценка позволила выявить, что код, написанный для данной программы не совершенен, так как результаты в ходе тестирования оказались ниже 100%. Следовательно, необходима доработка и усовершенствование программы, что и будет следующей целью развития данного проекта.

Задания, поддерживаемые программой, могут быть использованы как для самостоятельного изучения языка, так и для применения на занятиях с преподавателем. Разумеется, разработанная программа для русскоязычных текстов не претендует на то, чтобы стать самостоятельным и единственным средством для изучения языка и улучшения языковых навыков. Тем не менее,

она будет не только интересна, но и полезна будущим пользователям, а также поможет им освоить русский язык на достаточно высоком уровне.

Что касается дальнейших направлений развития, то в первую очередь подразумевается усовершенствование данной программы (повышение точности генерации упражнений, использование лексических минимумов для разных уровней языка, разделение текстов по уровням сложности), а также добавление новых типов упражнений.

### **Библиография**

1. Горисев С.А. Интеллектуальный лингвопроцессорный комплекс для обучения русскому языку как иностранному / С.А. Горисев, А.В. Койнов, С.В. Лисинин, А.П. Савинов // Научно-методическая конференция «Уровневая подготовка специалистов: государственные и международные стандарты инженерного образования», 26-30 марта 2013 г. – Томск: Изд-во Томского политехнического университета, 2013. – С. 315-316.
2. Antonsen, L., Johnson, R., Trondsterud, T., and Uibo, H. (2013). Generating Modular Grammar Exercises with Finite-State Transducers. Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning at NODALIDA 2013 27–38.
3. Malafeev A.Yu. Exercise Maker: Automatic Language Exercise Generation. In: Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27–30 мая 2015 г.). Вып. 14 (21): В 2 т. Т. 1: Основная программа конференции. — М.: Изд-во РГГУ, 2015. С. 441-452.
4. Aldabe, I., Lacalle, M. L. de, Maritxalar, M., Martinez, E., and Uria, L. (2006). ArikIturri: An Automatic Question Generator Based on Corpora and NLP Techniques. In Intelligent Tutoring Systems, M. Ikeda, K. D. Ashley, and T.-W. Chan, eds. (Springer Berlin Heidelberg), pp. 584–594.

5. Almeida, J. J., Araujo, I., Brito, I., Carvalho, N., Machado, G. J., Pereira, R. M. S., and Smirnov, G. (2013). PASSAROLA: High-Order Exercise Generation System. In 2013 8th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–5.
6. Burstein, J., and Marcu, D. (2005). Translation Exercise Assistant: Automated Generation of Translation Exercises for native-Arabic Speakers Learning English. In Proceedings of HLT/EMNLP on Interactive Demonstrations, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 16–17.
7. Dickinson, M., and Herring, J. (2008). Developing Online ICALL Exercises for Russian. In Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications, (Stroudsburg, PA, USA: Association for Computational Linguistics), pp. 1–9.