

# ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ INFOCOMMUNICATION TECHNOLOGIES

УДК 004.89

DOI 10.52575/2712-746X-2024-51-1-189-200

## Лексический тренажер по иностранному языку для студентов технических специальностей МГТУ им. Н.Э. Баумана

Бутенко Ю.И., Солошенко К.А.

Московский государственный технический университет им. Н.Э. Баумана,  
Россия, 105005, Москва, Рубцовская наб., 2/18, стр. 2  
E-mail: iubutenko@bmstu.ru

**Аннотация.** В статье представлен метод формирования лексического тренажера на основе лингводидактического ресурса «English in Digital Age» для оказания содействия студентам технических специальностей МГТУ им. Н.Э. Баумана в освоении лексики при подготовке к практическим занятиям по английскому языку и английскому языку в профессиональной сфере в период обучения специальности. Проанализированы факторы, обуславливающие актуальность создания лексического тренажера по освоению специальной лексики. При реализации лексического тренажера для студентов технических специальностей использованы методы лексикографического описания, осуществлено моделирование структуры словарной статьи, получен программный лексикографический продукт на языке программирования Python. Предложен метод обработки текстов учебных пособий для извлечения основной лексики каждого модуля, который предусматривает анализ композиционной структуры всего текста учебного пособия, а также возможность извлечения многокомпонентных терминов. Показана структура словарной статьи лексического тренажера, в которой в компактном виде отражены основные грамматические особенности слова, транскрипция и его переводные эквиваленты. Продемонстрированы интерфейс лексического тренажера и основные виды заданий, предназначенные для освоения лексического материала как по всем, так и по отдельным модулям дисциплины «Иностранный язык». Проанализированы возможности использования разных лингвистических ресурсов в аспекте освоения лексики с привязкой к конкретному учебному материалу. Электронный учебный словарь предназначен для студентов технических специальностей МГТУ им. Н.Э. Баумана, изучающих дисциплину «Иностранный язык».

**Ключевые слова:** электронный словарь, лексический тренажер, словарная статья, структура словарной статьи, интерфейс лексического тренажера

**Для цитирования:** Бутенко Ю.И., Солошенко К.А. 2024. Лексический тренажер по иностранному языку для студентов технических специальностей МГТУ им. Н.Э. Баумана. Экономика. Информатика. 51(1): 189–200. DOI 10.52575/2712-746X-2024-51-1-189-200

## Foreign Languages Lexical Exerciser for BMSTU Students in Technical Specialities

Iuliia I. Butenko, Kirill A. Soloshenko

Bauman Moscow State Technical University,  
2/18 Rubtsovskaya emb. Moscow, 105005, Russia  
E-mail: iubutenko@bmstu.ru

**Abstract.** The article presents a lexical exerciser based on the linguodidactic resource "English in Digital Age" to assist students of technical specialities of Bauman Moscow State Technical University in mastering vocabulary in preparation for practical classes in English and English in the professional sphere during the period of speciality training. The factors that determine the relevance of creating a lexical exerciser for mastering special

vocabulary are analysed. The methods of lexicographic description are used in the implementation of the lexical exerciser for students of technical specialities, modelling of the dictionary article structure is carried out, the lexicographic software product in the Python programming language is obtained. The method of text processing of textbooks for extracting the basic vocabulary of each module is proposed, which provides for the analysis of the compositional structure of the whole text of the textbook, as well as the possibility of extracting multi-component terms. The structure of the vocabulary article of the lexical trainer is shown, in which the main grammatical features of a word, transcription and its translation equivalents are reflected in a compact form. The interface of the lexical exerciser and the main types of tasks intended for mastering the lexical material for all and separate modules of the discipline "Foreign Language" are demonstrated. The possibilities of using different linguistic resources in the aspect of vocabulary acquisition with reference to specific teaching material are analysed. The electronic learning dictionary is designed for students of technical specialities of Bauman Moscow State Technical University studying the discipline "Foreign Language".

**Keywords:** electronic dictionary, lexical exerciser, dictionary article, dictionary article structure, lexical exerciser interface

**For citation:** Butenko Iu.I., Soloshenko K.A. 2024. Foreign Languages Lexical Exerciser for BMSTU Students in Technical Specialities. Economics. Information technologies. 51(1): 189–200. DOI 10.52575/2712-746X-2024-51-1-189-200

---

## Введение

Стремительное развитие информационных технологий оказало значительное влияние на лексикографию, которой требовались десятки лет на организацию сбора, обработки, хранения и представления лексического материала. Использование современных технологий позволило сократить время на внесение изменений в словарь от нескольких лет до нескольких минут, а также значительно увеличить объемы содержащейся в словарях информации [Палкова, 2015; Сидорова, Козлова, 2021]. Так, объем словника электронного словаря Мультитран на сегодняшний день составляет приблизительно 12 млн слов, Prompt – 11 млн слов и АБВУ LingvoOnline – 18 млн слов; количество рабочих языков: Мультитран – 12, Google – 102, Prompt – 13, АБВУ LingvoOnline – 28.

Несмотря на значительные успехи электронной лексикографии, использование многочисленных электронных ресурсов часто представляет значительные трудности не только для студентов и начинающих переводчиков, но и опытных специалистов, особенно при переводе специальной научно-технической литературы. Решение проблемы может состоять в возможности использования специализированных словарей, однако их количество также стремительно возрастает: Prompt – 16 словарей, АБВУ LingvoOnline – около 250 словарей, Мультитран – около 2500 словарей. Таким образом, поиск подходящего переводного эквивалента в рамках узкой специализации даже при помощи специализированных словарей занимает значительное время, а отдельные узкоспециализированные переводные словари представляют собой ограниченный класс лексикографических ресурсов [Мезит, 2019]. Особую сложность представляет изучение специальной лексики в группах с разным уровнем владения иностранным языком.

В связи с вышесказанным актуальным является создание лексического тренажера по изучению специальной терминологии, содержащего все лексические единицы и задания на их запоминание, которые должны быть освоены в рамках дисциплины «Иностранный язык (английский)» для студентов технических специальностей МГТУ им. Н.Э. Баумана. При этом проблематика создания учебных словарей и лексических тренажеров довольно широко освещена в научной литературе и имеется значительный спрос на такие ресурсы [Ятаева, 2016; Орлова, 2019], актуальными является создание методов и программных средств автоматической и / или автоматизированной генерации предметных словарей и лексических тренажеров.

Актуальность создания лексического тренажера по освоению специальной лексики обусловлена следующими факторами:

- длительный срок изучения иностранного языка в МГТУ им. Н.Э. Баумана: бакалавриат – 6 семестров, магистратура – 2 семестра, специалитет – 7 семестров. При этом, начиная с 5 семестра бакалавриата и специалитета, а также 1 семестра магистратуры, изучается английский язык в профессиональной сфере.

- большое разнообразие узких предметных областей. В 2022/2023 гг. в МГТУ им. Н.Э. Баумана насчитывается 479 образовательных программ, подавляющее число которых имеют свою уникальную узкопрофессиональную терминологию, что в свою очередь свидетельствует об использовании большого числа разных учебных пособий по иностранному языку.

- разработка пособий преподавателями университета. Учебные пособия пишутся преподавателями иностранного языка в МГТУ им. Н.Э. Баумана, т. е. требуется затрачивание значительных временных и человеческих ресурсов для создания комплектов учебных и практических материалов.

- использование нового учебного пособия каждый семестр препятствует возможности повторения пройденного ранее лексического материала с целью его повторения перед итоговыми аттестациями.

- сложность использования современных словарей и терминологических баз данных из-за их универсальности, с одной стороны, и перегруженности избыточной для студентов информацией, с другой стороны, а также отсутствие инструментов для запоминания специальной лексики.

Таким образом, цель создания лексического тренажера – оказание содействия студентам инженерных специальностей МГТУ им. Н.Э. Баумана в освоении специальной лексики при подготовке к практическим занятиям и итоговым аттестациям по английскому языку.

### **Метод обработки текстов учебных пособий для создания лексического тренажера**

Основные задачи разработки лексического тренажера для студентов инженерных специальностей МГТУ им. Н.Э. Баумана включают:

- сбор языкового материала, необходимого для освоения учебной программы дисциплины «Иностранный язык»;
- форматирование и разметка полученного языкового материала в виде, пригодном для машинной обработки;
- разработка алгоритмов лексического тренажера, соответствующих лингводидактическим требованиям лексического тренажера;
- создание статического сайта / приложения для размещения языкового материала и создание лексического тренажера.

Из перечисленных задач наиболее рутинной и требующей значительных временных и человеческих затрат является сбор и разметка языкового материала. При этом для решения поставленной задачи целесообразно использовать доступные подходы и программные средства по обработке специальной лексики, подробный анализ которых представлен в работе [Astrakhansev et al, 2015]. Однако использование стандартных средств обработки специальной лексики затруднено тем фактом, что учебные пособия по дисциплине «Иностранный язык» написаны полностью на английском языке, что приводит к необходимости разработки метода извлечения лексики из учебного пособия, а также подбора переводных эквивалентов на русском языке и последующей генерацией тренажера по изучению извлеченной лексики.

Метод обработки учебных пособий для создания лексического тренажера по изучению специальной терминологии представлен на рис. 1.

Рассмотрим использование метода обработки текстов учебных пособий на основе учебного пособия «English in the Digital Age» [Калугина, Лосева, 2018], который используется на 1 и 2 курсах инженерных специальностей на кафедре ЛЗ «Английский для машиностроительных специальностей» МГТУ им. Н.Э. Баумана, состоящего из 4 частей. Каждое учебное пособие состоит из трех модулей, содержащих современные аутентичные тексты,

целью которых является развитие коммуникативной, лингвострановедческой и социокультурной компетенций. Издание также содержит ссылки на аудио- и видеоматериалы из открытых источников англоязычной информационно-образовательной среды и разнообразные лексико-грамматические упражнения. На изучение каждого модуля отводится 5 семинарских занятий, где 1–4 занимаются по учебному пособию, а на 5 занятия проводят рубежный контроль и защиту презентации по теме модуля.



Рис. 1. Метод обработки учебных пособий для создания лексического тренажера по изучению специальной терминологии

Fig. 1. Method of processing teaching aids for creating a lexical simulator for learning specialised terminology

Тексты и упражнения второй части учебного пособия «English in the Digital Age» знакомят студентов с историей развития современных средств коммуникации и компьютерных технологий, с новейшими разработками в области связи, компьютерных технологий и космических исследований.

На первом этапе проводится морфологическая разметка текста учебного пособия, которая в последующем является основой для извлечения и разметки многокомпонентных терминов в тексте, генерации тестовых заданий и отражения грамматической информации о лексической единице в словаре тренажера.

На втором этапе проводится структурная разметка текста учебного пособия. Особенности организации научно-учебных текстов подробно в работах [Тюрина, 2005; Тюрина, 2007], а обобщенная модель научно-учебного текста представлена в [Бутенко, 2021]. Ее ис-

пользование позволит определить структурные элементы, которые не содержат специальной лексики (формулировки заданий, грамматические комментарии, задания на перевод на английский язык), а также структурные элементы, содержащие основную лексику модуля. Таким образом, можно будет сформировать перечень всей лексики, используемой в пособии, и основной лексики для изучения в каждом модуле учебного пособия. Основной языковой материал учебном пособии «English in the Digital Age» представляет собой лексику из разделов «Essential vocabulary», «Vocabulary related to...», «Useful notes», «Module's-Topic Related Vocabulary».

На третьем этапе осуществляется разметка многокомпонентных терминов как одной лексической единицы при создании заданий тренажера, а затем формирование списка основной лексики модуля [Бутенко и др., 2022]. В результате обработки текста учебного пособия по каждому из 12 модулей выделено: 1 модуль – 67, 2 модуль – 66, 3 модуль – 51, 4 модуль – 39, 5 модуль – 141, 6 модуль – 124, 7 модуль – 83, 8 модуль – 102, 9 модуль – 38, 10 модуль – 70, 11 модуль – 70, модуль 12 – 49 единиц основной лексики, владение которой студент должен продемонстрировать на семинарских занятиях и промежуточных аттестациях.

На следующем этапе к каждой лексической единице модуля на основе параллельного корпуса научно-технических текстов, разрабатываемого в МГТУ им. Н.Э. Баумана [Бутенко и др., 2022], подбирается 1–3 переводных эквивалента. Стоит отметить, что широкий спектр исследований в области как создания [Bolshina, Loukhashevich, 2020; Tikhomirov et al., 2020], так и использования корпусов текстов в образовательных целях свидетельствует об эффективности и применимости таких ресурсов для решения большого класса образовательных задач [Кокарева, 2013а], в том числе по разработке заданий [Слободян, 2010] и изучению лексики [Кокарева, 2013б].

При подборе переводного эквивалента на основе параллельного корпуса используется принцип соответствия предметной области учебного пособия и соответствующего подкорпуса, а также частотность использования переводного эквивалента в подкорпусе. Так, например, на основе словаря МультиТран к лексической единице *skin*, которая используется в шестом модуле, посвященном космонавтике, даны следующие переводные эквиваленты: в общем словаре: *мех; кожура; кожица; оболочка; обшивка; корка; плена; наружный слой; шкурка; кляча; жулик; скряга; кожа; шкура; снять кожуру; покрыть кожей; саба; снять кожу; снять шкуру; шелуха*. В словаре «Техника» ресурса МультиТран – *окалина; покрытие; поверхностный слой; кожа; обшивочный лист; очищать от изоляции*, а в словаре «Космонавтика» вообще отсутствует. Уточнение запроса путем добавления слова *spacescraft skin* приводит к тому, что нужно смотреть уже две словарные статьи. Безусловно подбор переводного эквивалента может и очевиден для специалиста в данной предметной области и не представляет сложностей – *обшивка*, но студентам как минимум требуется время на его подбор или они вообще выбирают неправильный перевод. На основе данных их параллельного корпуса научно-технических текстов, тематическое деление которого основано на направлениях подготовки МГТУ им. Н.Э. Баумана, у лексической единицы *skin* наиболее частотными являются переводные эквиваленты по космонавтике *обшивка, обшивочный лист*, а биомедицинским технологиям – *кожа, кожный покров*.

Затем на основе подкорпуса осуществляется подбор примеров употребления лексической единицы в контексте, а также его перевод. Приоритетными контекстами являются те, которые содержат не только основную лексику модуля, но и специальную терминологию изучаемой предметной области, так как эти контексты используются для автоматической генерации заданий тренажера. В противном случае могут быть сформированы контексты, куда по смыслу может подходить практически любое слово, например:

*The phenomenon of <<b>...<d>> has always been very exciting, both for its fundamental scientific interest and because of its many applications.*

*Феномен <<b>...<d>> всегда был очень захватывающим как из-за его фундаментального научного интереса, так и из-за его многочисленных применений.*

В лексическом тренажере представлено несколько видов заданий для овладения специальной лексикой: карточки, выбор верного переводного эквивалента и заполнение пропусков. Генерация заданий реализуется на основе следующих правил: варианты ответов подбираются из одного модуля, обладают одинаковыми морфологическими, грамматическими и синтаксическими характеристиками. Например, при формировании задания на выбор правильного переводного эквивалента, все варианты ответов будут относиться к одной части речи, числа, использоваться в рамках одной предметной области, иметь одинаковую формальную структуру, т. е. если правильный ответ является словосочетанием, то и другие варианты ответов также будут словосочетаниями и т. п.

### Структура словарной статьи лексического тренажера

Словарная статья учебного ресурса должна не содержать избыточной и неточной информации, но в то же время отражать все аспекты, необходимые для освоения дисциплины [Жеребило, 2009; Мячина, 2018]. Каждая лексическая единица сопровождается транскрипцией и аудиофайлом для прослушивания данной единицы в британском варианте произношения. К полученным словарным статьям обязательно добавлялась грамматическая категория части речи, что позволило составлять соответствующие лингводидактическим нормам задания для тренажера автоматически и без использования способных проводить грамматический парсинг лексемы алгоритмов искусственного интеллекта для обработки естественного языка. Для тех же целей в примерах употребления лексема обязательно выделялась тегами: открывающим «<<b>» и закрывающим «<<d>>». В каждой словарной статье указывался модуль учебного пособия, к которому относится слово.

Словарные статьи обрабатывались и загружались на сервер в виде объекта JavaScript, содержащего объекты JSON вида:

```
{ "word": string,  
  "translation": string,  
  "transcription": string,  
  "pos": string,  
  "examples": [{  
    "eng": string,  
    "ru": string  
  }],  
  "module": number }
```

где поле "word" хранит лексему на английском языке, поле "translation" хранит перевод или переводы, разделенные запятой, поле "transcription" хранит транскрипцию лексемы, поле "pos" хранит часть речи лексемы, поле "examples" хранит массив с одним элементом – объектом JSON, хранящим поля «eng» и «ru», поле "eng" хранит пример употребления лексемы на английском языке, поле "ru" хранит пример употребления лексемы на русском языке, поле "module" хранит номер модуля, к которому относится лексема.

Словарные статьи лексического тренажера содержат грамматический материал о лексической единице и аудиозаписи британского варианта произношения. Пример словарной статьи показан на рис. 2.

При заполнении пропусков используются контексты, внесенные в словарь лексического тренажера, а основная лексика имеет соответствующую разметку и при формировании заданий является правильным ответом, например:

*His invention was a long way away where we are today with our mobile phones, <<b>tablets<d>> and laptops but it was a significant development in the use of calculating devices.*

*Его изобретение было далеко от современных телефонов, <<b>планшетов<d>> и ноутбуков, но оно было значительным прорывом в области счетных устройств.*

Другими вариантами являются *abacus, calculator, display* и *абак, калькулятор, монитор* соответственно на русском языке.

## action bar (phr) - горизонтальное меню

/ˈækʃn bɑː(r)/



When an update is available, an icon in the **action bar** displays some explaining text.

Если доступно обновление, значок в **горизонтальном меню** отображает текст пояснения.

Рис. 2. Словарная статья к терминологической единице «action bar»

Fig. 2. Dictionary entry for the terminological unit "action bar"

### Аспекты программной реализации лексического тренажера

Для разработки лексического тренажера выбраны следующие технологии:

- HTML – стандартизированный язык разметки документов для просмотра web-страниц в браузере. Использовался как основной инструмент верстки web-страниц сайта ЭУС.
- CSS – формальный язык описания внешнего вида документа (web-страницы). Использовался для описания дополнительных свойств элементов web-страниц.
- Bootstrap 4 – свободный набор HTML- и CSS-шаблонов для создания сайтов и web-приложений. Классы Bootstrap были использованы для стилизации большинства элементов сайта, что позволило сократить затраченное на верстку время к минимуму при сохранении качественного внешнего вида web-страниц ресурса.
- JavaScript – мультипарадигменный язык программирования, поддерживающий слабую динамическую типизацию и автоматическое управление памятью. Был использован в качестве основной технологии для разработки поведения сайта и хранения словарных статей.
- Git – распределённая система управления версиями. Использовалась для удаленной разработки.
- JSON – основанный на JavaScript текстовый формат обмена данными. Использовался для хранения словарных статей в удобном не занимающем большого количества оперативной памяти формате.
- Go – компилируемый многопоточный язык программирования. Использовался для разработки программы-парсера для получения аудиозаписей произношений слов с сайта «Oxford learner's dictionary» и Google Translate API.
- Python – высокоуровневый язык программирования общего назначения с динамической строгой типизацией и автоматическим управлением памятью. Использовался как основной инструмент для обработки словарных статей и тегирования предложений-примеров.

Основными принципами проектирования словаря являются простота и доступность, так как минималистичный и интуитивно понятный графический интерфейс ресурса фокусирует внимание студента на изучении языкового материала.

Все страницы сайта лексического тренажера содержат навигационную панель с поисковой строкой и ссылками на главную страницу сайта и лексический тренажер. На главной странице сайта расположены ссылки для поиска языкового материала по первой букве и по модулю учебного пособия (рис. 3).

Для поисковой строки реализован «живой поиск» методом поиска соответствий по средствам строгого регулярного выражения. Ссылки поиска по первой букве и по модулю ведут на соответствующие страницы с прямыми ссылками на искомые словарные статьи. Ссылки из списка всего языкового материала и из поисковой строки являются прямыми ссылками на искомые словарные статьи.

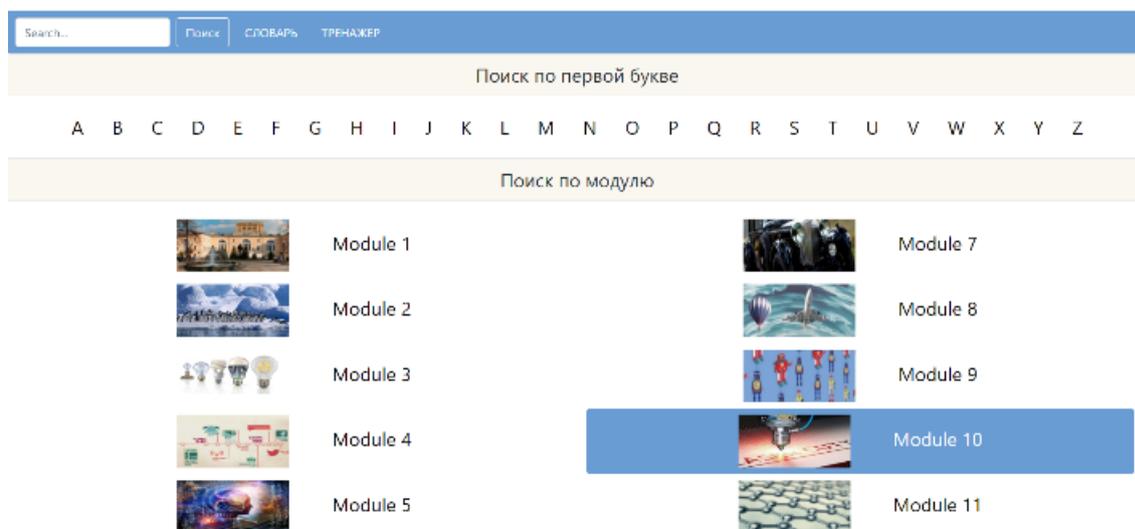


Рис. 3. Интерфейс лексического тренажера  
 Fig. 3. Interface of the lexical simulator

Алгоритмы данного тренажера позволяют автоматически составлять более 100 000 типовых задач на перевод, выбор верного ответа, заполнение пропусков в предложении (Рис. 4). В лексическом тренажере присутствует классическая система карточек для заучивания лексики. Имеется система фильтрации задач по модулям в зависимости от потребностей учащегося.

| Choose correct answ        | Choose correct answ  |
|----------------------------|--|
| улучшать, совершенствовать | MSTU was the first among technical ____ educational institutions that received the status of technical university in 1989. |
| improve                    | higher   |
| invent                     | keen   |
| melt                       | safe   |
| apply                      | edible   |

Рис. 4. Примеры сгенерированных заданий в лексическом тренажере  
 Fig. 4. Examples of generated tasks in the lexical simulator

В таблице представлены результаты сравнения возможностей лексикографических ресурсов для поиска переводных эквивалентов и изучения специальной лексики студентами технических специальностей МГТУ им. Н.Э. Баумана. Основным критерием при отборе лексикографических ресурсов была доступность и частота использования в студенческой среде.

Таким образом, несмотря на наличие больших лингвистических баз данных и словарей, которые содержат в себе значительные объемы информации о каждом ресурсе, для изучения специальной лексики в рамках дисциплины «Иностранный язык» целесообразно использовать ресурсы, привязанные к конкретным учебным пособиям. Более того, в современной лексикографии формируется адресный подход к созданию словарей, то есть делается акцент на потребности рядового пользователя с целью более тонко и дифференцированно реализовать основные принципы лексикографического описания [Самосенкова, Ряполова; 2018; Каримуллина, Каримуллина, 2022]. Таким образом, метод обработки научно-учебных текстов, предложенный в статье, позволит автоматизировать процедуру создания адресных словарей. Кроме того, он может быть также использован для проверки разрабатываемых учебных пособий, контрольных работ и других материалов на наличие основной лексики и количестве ее употреблений в тексте пособия, или же использования незнакомой лексики при создании проверочных работ.

Таблица 1  
 Table 1

Сравнительный анализ возможностей лексикографических ресурсов по изучению специальной лексики  
 Comparative analysis of lexicographic resources for the study of special vocabulary

|          | Ресурс  | Мультигран | Abby Lingvo | Prompt | Словарь Google | Словарь Yandex | Лексический тренажер |
|----------|---|------------|-------------|--------|----------------|----------------|----------------------|
| Критерий |   |            |             |        |                |                |                      |
| 1        | Количество переводных эквивалентов в словаре        | 1-50       | 1-15        | 1-10   | 1-20           | 1-20           | <b>1-3</b>           |
| 2        | Соотнесенность с конкретной предметной областью     | +          | +           | +      | +              | +              | +                    |
| 3        | Привязка к конкретному учебному материалу / пособию | -          | -           | -      | -              | -              | +                    |
| 4        | Использование карточек для изучения лексики         | -          | +           | -      | -              | -              | +                    |
| 5        | Тестовые задания для изучения лексики               | -          | -           | -      | -              | -              | +                    |
| 6        | Наличие транскрипции и грамматической информации    | +          | +           | +      | +              | +              | +                    |
| 7        | Примеры употребления изучаемой единицы в контексте  | -          | +           | -      | +              | -              | +                    |

### Заключение

В статье представлен лексический тренажер по изучению специальной терминологии на основе учебных пособий, используемых на занятиях по иностранному языку на примере лингводидактического ресурса «English in Digital Age» для студентов технических специальностей МГТУ им. Н.Э. Баумана. Основой для разработки лексического тренажера послужил метод обработки учебных пособий, предусматривающий реализацию в несколько этапов: морфологическую, структурную, терминологическую разметки, формирование списков лексических единиц к каждому модулю, подбор переводных эквивалентов и контекстов употребления лексических единиц модуля на основе параллельного корпуса, а также генерацию словарных статей и заданий лексического тренажера на основе шаблонов. Каждая словарная статья лексического тренажера отражает основные грамматические особенности, транскрипцию, переводные эквиваленты и контексты употребления лексической единицы. На основе сравнительного анализа возможностей разных лингвистических ресурсов обосновано, что при изучении иностранного языка студентами технических специальностей МГТУ им. Н.Э. Баумана целесообразно использовать ресурсы, привязанные к конкретным учебным пособиям и содержащим практические задания на усвоение специальной лексики.

### Список литературы

- Бутенко Ю.И., Строганов Ю.В., Сапожков А.М. 2022. Система извлечения многокомпонентных терминов и их переводных эквивалентов из параллельных научно-технических текстов. Научно-техническая информация: Серия 2. Информационные процессы и системы, 9: 12–21. DOI: 10.36535/0548-0027-2022-09-3
- Бутенко Ю.И. 2021. Модель учебно-научного текста для разметки корпуса научно-технических текстов. Экономика. Информатика, 1(48): 123–129. DOI:10.52575/2687-0932-2021-48-1-123-129

- Жеребило Т.В. 2009. Типы лексикографической информации в учебном словаре лингвистических терминов. Вестник Мордовского университета, 2(19): 112–121.
- Калугина Л.В., Лосева О.М. 2018. Английский язык в эпоху цифровых технологий. Книга 1 = English in the Digital Age : мультимедийное учебное пособие / Под ред. Гуровой Г.Г. М.: Изд-во МГТУ им. Н. Э. Баумана, 108.
- Каримуллина Г.Н., Каримуллина Р.Н. 2022. Медицинская терминология в двуязычной учебной лингвографии. Филология и культура, 2(68): 18–23. DOI: 10.26907/2074-0239-2022-68-2-18-23
- Кокорева А.А. 2013а. Корпус параллельных текстов в обучении иностранному языку. Вестник Тюменского государственного университета. Серия Гуманитарные науки. Педагогика и психология, 2(118): 57–62.
- Кокорева А.А. 2013б. Методические условия обучения студентов профессионально-ориентированной лексики на основе корпуса параллельных текстов. Вестник Тамбовского государственного университета. Серия Гуманитарные науки, 1(117): 142–146.
- Мезит А.Э. 2019. Концепция «Словаря специальной лексики русской гидроэнергетической отрасли». Вопросы лексикографии, 16: 138–152. DOI: 10.17223/22274200/16/8.
- Мячина В.В. 2018. Объём и структура лексико-семантического поля туризм в содержании обучения русскому языку иностранных студентов. Мир науки, культуры, образования, 1(68): 86–87.
- Орлова Е.В. 2019. Электронный учебный словарь коллокаций для специалистов МЧС России как средство развития учебной иноязычной лексической компетенции. Пожарная и аварийная безопасность, 3(14): 32–35.
- Палкова А.В. 2015. Основные понятия электронной лексикографии. Вестник Тверского государственного университета. Серия «Филология», 4: 88–93.
- Самосенкова Т.В., Ряполова М.А. 2018. Инновационное информативное обеспечение личности посредством создания личных словарей. Filologos, 3(38): 46–53. DOI: 10.24888/2079-2638-2018-38-3-46-53
- Сидорова О.Г., Козлова Т.В. 2021. Первые отечественные учебные словари английского языка. Педагогическое образование в России, 1: 8–16. DOI 10.26170/2079-8717\_2021\_01\_01.
- Слободян Е.А. 2010. Обучение методике составления упражнений с помощью Национального корпуса русского языка. Гуманитарные исследования в Восточной Сибири и на Дальнем Востоке, 1: 71–73.
- Тюрина Л.Г. 2005. Состав и структура учебной книги как педагогической системы. Известия высших учебных заведений. Проблемы полиграфии и издательского дела, 4: 78–88.
- Тюрина Л.Г. 2007. Особенности текста учебной книги. Известия высших учебных заведений. Проблемы полиграфии и издательского дела, 3: 70–73.
- Ятаева Е.В. 2016. Электронный учебный словарь как средство развития учебной иноязычно-лексической компетенции. Вестник Челябинского государственного педагогического университета, 10: 135–140.
- Astrakhantsev N.A., Fedorenko D.G., Turdakov D.Yu. 2015. Methods for automatic term recognition in domain-specific text collections: a survey. Programming and Computer Software, 41: 336–349.
- Bolshina, A., Loukashevitch N. 2020. Automatic labelling of genre-specific collections for word sense disambiguation in Russian. Lecture Notes in Computer Science, 12412: 215–227. DOI 10.1007/978-3-030-59535-7\_15
- Tikhomirov M., Loukachevitch N., Sirotina A, Dobrov B. 2020. Using BERT and augmentation in named entity recognition for cybersecurity domain. Lecture Notes in Computer Science, 12089: 16–24. DOI 10.1007/978-3-030-51310-8\_2

## References

- Butenko Yu.I. 2021. Model of educational texts for markup in a corpus of scientific and technical texts. Economics. Information technologies, 1(48): 123–129. (In Russian). DOI:10.52575/2687-0932-2021-48-1-123-129
- Butenko Yu.I., Stroganov Yu.V., Sapozhkov A.M. 2022. System for extracting multicomponent terms and their translated equivalents from parallel scientific and technical texts. Scientific and Technical Information: Series 2. Information processes and systems, 9: 12–21. (In Russian). DOI: 10.36535/0548-0027-2022-09-3

- Zherebilo, T. V. 2009. Types of lexicographic information in the educational dictionary of linguistic terms. *Vestnik Mordovskogo universiteta*, 2(19): 112–121. (In Russian.)
- Kalugina L. V., Loseva O.M. 2018. English in the Digital Age. Book 1 = English in the Digital Age: multimedia textbook / Edited by Gurova G. G. Moscow: Izd. G. - M.: Izd-vo MGU im. N.E. Bauman, 108.
- Karimullina G.N., Karimullina R.N. 2022. Medical terminology in bilingual academic linguography. *Philology and Culture*, 2(68): 18–23. (In Russian.) DOI: 10.26907/2074-0239-2022-68-2-18-23
- Kokoreva A.A. 2013a. Parallel corpus in foreign language teaching. *Tyumen State University Herald. Humanities Research. Humanitates*, 2(118): 57–62.
- Kokoreva A.A. 2013b. Methodological conditions for of professional lexis skills via corpus of parallel texts. *Tambov University Reviews. Series Humanitie*, 1(117): 142–146. (In Russian.)
- Mezit A.E. 2019. The concept of the Russian hydropower industry specific dictionary. *Russian Journal of Lexicography*, 16: 138–152. (In Russian.) DOI: 10.17223/22274200/16/8
- Myachina V.V. 2018. Volume and structure of lexical-semantic field “Tourism” in the contents of teaching Russian language to foreign students. *The world of science, culture and education*, 1(68): 86–87. (In Russian.)
- Orlova E. V. 2019. Electronic learner’s dictionary of collocations for specialists of emercom of Russia as a means of developing lexical competence in English. *Fire and Emergency Safety*, 3(14): 32–35.
- Palkova A.V. 2015. Basic definitions of electronic lexicography. *Vestnik TvSU. Series Philology*, 4: 88–93. (In Russian.)
- Samosenkova T.V., Ryapolova M.A. 2018. Informative providing of the individual by creating targeted dictionaries. *Filologos*, 3(38): 46–53. (In Russian.) DOI: 10.24888/2079-2638-2018-38-3-46-53
- Sidorova, O. G., Kozlova T. V. 2021. The first Russian-English dictionaries compiled in Russia. *Pedagogical Education in Russia*, 1: 8–16. (In Russian.) DOI 10.26170/2079-8717\_2021\_01\_01.
- Slobodyan E.A. 2010. Learning methodology of composing exercises with the help of the Russian national corpus. *Humanities Research in the Russian Far East*, 1: 71–73.
- Tyurina L.G. 2005. Composition and structure of the educational book as a pedagogical system. *Izvestia of higher educational institutions. Problems of polygraphy and publishing*, 4: 78–88.
- Tyurina L.G. 2007. Features of the text of an educational book. *Izvestia of higher educational institutions. Problems of polygraphy and publishing*, 3: 70–73.
- Yataeva E.V. 2016. Learner’s electronic dictionary as a means of foreign language learning lexical compeyence development. *The Herald of South-Ural state Humanities-Pedagogical University*, 10: 135–140.
- Astrakhantsev N.A., Fedorenko D.G., Turdakov D.Yu. 2015. Methods for automatic term recognition in domain-specific text collections: a survey. *Programming and Computer Software*, 41: 336–349.
- Bolshina, A., Loukashevitch N. 2020. Automatic labelling of genre-specific collections for word sense disambiguation in Russian. *Lecture Notes in Computer Science*, 12412: 215–227. DOI 10.1007/978-3-030-59535-7\_15
- Tikhomirov M., Loukachevitch N., Sirotina A, Dobrov B. 2020. Using BERT and augmentation in named entity recognition for cybersecurity domain. *Lecture Notes in Computer Science*, 12089: 16–24. DOI 10.1007/978-3-030-51310-8\_2

**Конфликт интересов:** о потенциальном конфликте интересов не сообщалось.

**Conflict of interest:** no potential conflict of interest related to this article was reported.

Поступила в редакцию 02.11.2023

Received November 02, 2023

Поступила после рецензирования 03.03.2024

Revised March 03, 2024

Принята к публикации 04.03.2024

Accepted March 04, 2024



## ИНФОРМАЦИЯ ОБ АВТОРАХ

**Бутенко Юлия Ивановна**, кандидат технических наук, доцент кафедры «Теоретическая информатика и компьютерные технологии», Московский государственный технический университет им. Н.Э. Баумана, г. Москва, Россия

**Солошенко Кирилл Александрович**, магистрант факультета «Лингвистика», Московский государственный технический университет им. Н.Э. Баумана, г. Москва, Россия

## INFORMATION ABOUT THE AUTHORS

**Iuliia I. Butenko**, Candidate of Technical Sciences, Associate Professor of the Department Theoretical Informatics and Computer Technologies. Bauman Moscow State Technical University, Moscow, Russia

**Kirill A. Soloshenko**, Master of the Department of Linguistics. Bauman Moscow State Technical University, Moscow, Russia