

Оригинальная статья

УДК 81.33

<http://doi.org/10.32603/2412-8562-2026-12-3-189-203>

Разработка инструмента автоматизации перевода английских звукоизобразительных единиц на русский язык

Егор Олегович Голованов¹✉, Елена Анатольевна Шамина²

¹Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

²Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

¹✉geo.hotline@gmail.com, <https://orcid.org/0009-0000-0210-3490>

²e.shamina@spbu.ru, <https://orcid.org/0000-0001-5274-1001>

Введение. Статья посвящена разработке инструмента автоматизации перевода художественных текстов, предназначенного для работы с звукоподражательной лексикой. Отличительной особенностью инструмента является его подход к отбору переводческих эквивалентов – алгоритмы предлагают эквиваленты, ориентируясь не на их лексическую близость к оригиналу или частотность, но на звуковую изобразительность. Вследствие стремительного развития средств автоматизированного и машинного перевода решение проблемы перевода элементов звуковой изобразительности представляется в настоящий момент особенно важным.

Методология и источники. Теоретическим фундаментом исследования являются постулаты фоносемантики, изложенные в работах С. В. Воронина, а также описания звукоподражаний для русского языка, приведенные в трудах С. С. Шляховой. Статья рассматривает автоматизированный перевод звукоизобразительных единиц на примере английских глаголов движения, перечисленных в работе Н. К. Хеймонен. Перечни русских эквивалентов для данных глаголов отбирались авторами вручную с использованием ряда англо-русских словарей.

Результаты и обсуждение. Результатом исследования является функционирующий инструмент, представленный в виде веб-приложения. Инструмент выполняет перевод звукоизобразительных элементов текста с английского языка на русский посредством ранжирования их переводческих эквивалентов по звукоизобразительному признаку. Под звукоизобразительным признаком понимается сходство оригинала и его перевода с звукоизобразительной точки зрения. Вычисление этого сходства стало возможным благодаря использованию уникального подхода – сравнения звукоподражательных формул оригинального слова и его переводов с использованием метрики расстояния Левенштейна. Помимо ранжирования эквивалентов статья описывает структуру базы данных, а также организацию работы веб-приложения.

Заключение. Разработанный инструмент успешно справляется с поставленной задачей, а уникальные методы отбора переводческих эквивалентов которые он использует, успешно восполняют отсутствие на рынке готовых решений для перевода звукоизобразительной лексики, обеспечивая потенциал данной технологии.

Ключевые слова: звукоизобразительность, звукоподражание, художественный перевод, глаголы движения, Python

© Голованов Е. О., Шамина Е. А., 2026



Контент доступен по лицензии Creative Commons Attribution 4.0 License.

This work is licensed under a Creative Commons Attribution 4.0 License.

Для цитирования: Голованов Е. О., Шамина Е. А. Разработка инструмента автоматизации перевода английских звукоизобразительных единиц на русский язык // ДИСКУРС. 2026. Т. 12, № 3. С. 189–203. DOI: 10.32603/2412-8562-2026-12-3-189-203

Original paper

Development of an Automation Tool for Translating English Phonosemantic Lexis into Russian

Egor O. Golovanov¹, **Elena A. Shamina²**

¹*Saint Petersburg Electrotechnical University, St Petersburg, Russia*

²*Saint Petersburg State University, St Petersburg, Russia*

¹✉ geo.hotline@gmail.com, <https://orcid.org/0009-0000-0210-3490>

² e.shamina@spbu.ru, <https://orcid.org/0000-0001-5274-1001>

Introduction. This article deals with development of an automation tool for translating fiction texts. The tool is designed to work with sound-symbolic vocabulary. A distinctive feature of the tool is its approach to selecting translation equivalents: the algorithms propose equivalents based on their sound symbolic features. Owing to the rapid development of automated and machine translation technologies, it is the time to address the problem of translating iconic vocabulary.

Methodology and sources. The theoretical foundation of the study is based on the principles of phonosemantics formulated in the works of S.V. Voronin, as well as descriptions of onomatopoeia presented in the studies of S.S. Shlyakhova. The article examines automated translation of sound-symbolic vocabulary through the use of motion verbs listed in the work of N. K. Kheimoinen.

Results and discussion. The result of the study is a functioning tool in the form of a web application. The tool translates sound-symbolic elements from English into Russian by ranking their translation equivalents according to the degree of similarity between the original word and its translation from the perspective of sound symbolism. The calculation of this similarity becomes possible due to a unique approach – comparing the onomatopoeic formulas of the original word and its translations using the Levenshtein distance.

Conclusion. The developed tool successfully accomplishes its task. Moreover, the unique methods it employs for selecting translation equivalents demonstrate the strong potential of this technology.

Keywords: phonosemantics, onomatopoeia, fiction translation, motion verbs, Python

For citation: Golovanov, E.O. and Shamina, E.A. (2026), "Development of an Automation Tool for Translating English Phonosemantic Lexis into Russian", *DISCOURSE*, vol. 12, no. 3, pp. 189–203. DOI: 10.32603/2412-8562-2026-12-3-189-203

Введение. В последние годы научное сообщество наблюдает стремительное развитие нейросетевых технологий, которому сопутствует совершенствование машинного и в том числе автоматизированного перевода. Если использование машинного перевода при работе с художественными текстами до сих пор вызывает оживленную дискуссию среди специалистов по переводу, то автоматизированный перевод, в котором основную работу выполняет человек, используя при этом узкоспециализированные инструменты, спроектированные и разработанные для облегчения труда переводчика, является не только нормой, но и рекомендуемым подходом к переводу художественных текстов и не только. Существующие на дан-

ный момент инструменты автоматизированного перевода не являются идеальными. Так, например, многие из них (электронные словари и глоссарии, память переводов), предлагая эквиваленты на языке перевода, ориентируются преимущественно на частотность данных эквивалентов, их лексическое соответствие оригиналам и, в случае с продвинутыми инструментами, на близлежащий контекст. В то же время художественный текст, будучи многогранным и нетривиальным феноменом, обладает такими аспектами, которые не учитываются существующими на данный момент инструментами автоматизированного перевода. К ним следует отнести среди прочих литературные и культурные аллюзии [1], метафоры, основанные на мультимодальных синестетических переносах [2], и фоносемантические свойства вербальных составляющих текста, которые могут свидетельствовать об общей экспрессивности произведения или являться показателями авторского идиостиля [3]. Создание же качественного и адекватного литературного произведения требует от переводчика взаимодействия со всеми особенностями, которые заложил автор оригинала в произведение [4]. Об одной из таких особенностей, а именно о звуковой изобразительности художественного текста, и пойдет речь в настоящей статье. Так как перевод звукоизобразительных элементов текста не всегда является простой и понятной задачей даже для высококвалифицированного литературного переводчика (проблемы фоносемантики крайне редко поднимаются в переводческом дискурсе – ср. [5]), авторы статьи предприняли попытку разработки инструмента автоматизированного перевода, целью которого является помощь переводчику именно с подбором корректных с точки зрения звуковой изобразительности переводческих эквивалентов.

Методология и источники. Фоносемантика – это лингвистическая дисциплина на стыке фонетики и семантики, постулаты которой были впервые оформлены в 1982 г. в работе выдающегося советского и российского лингвиста С. В. Воронина «Основы фоносемантики». Основоположник фоносемантики характеризует ее следующим образом: «*фоносемантика* – это дисциплина, имеющая своим предметом звукоизобразительную (т. е. звукоподражательную и звукосимволическую) систему языка, изучаемую с позиций пространственных и временных» [6, с. 4]. Он же указывает на ее теоретическое значение и на прагматические (эвристические) возможности [6, с. 192–197].

Благодаря главенствовавшему в лингвистике XX в. направлению структурализма, звуковую изобразительность долгое время относили к языковым «маргиналиям», из-за чего исследования этого явления практически не проводились [6, с. 13]. Тем не менее уже во второй половине XX в. исследователи доказали межъязыковой характер звуковой изобразительности, закрепив таким образом ее статус в качестве языковой универсалии (ср. [7]). В конце века проходит первый международный симпозиум по иконичности в языке и литературе, который с тех пор проводится каждые два года, а результаты фоносемантических исследований публикуются в самых престижных научных изданиях. Таким образом, можно сделать вывод о том, что результатом «маргинализации» звукоизобразительности на протяжении большей части XX в. является тот факт, что исследования иконичности выделились в самостоятельную дисциплину крайне поздно, вследствие чего фоносемантика до сих пор является активно развивающимся лингвистическим направлением.

С. В. Воронин определяет звуковую изобразительность следующим образом: «*звукоизобразительность* есть свойство слова, заключающееся в наличии необходимой, повторя-

ющей и относительно устойчивой произвольной связи между фонемами слова ... и признаком объекта-денотата» [6, с. 166], т. е. постулируется произвольная связь между звучанием и значением слова. Среди исследователей-фоносемантистов принято выделять два основных вида звукоизобразительности: звукоподражание и звукоимовизм. Характеризуя их, С. В. Воронин указывает, что *звукоподражания* (или ономотопы) – это слова, которые подражают звукам окружающей действительности (*выть, звенеть, гроыхать* и т. п.), *звукоимовизмы* же отражают признаки объектов в любой из сенсорных модальностей человека, кроме слуховой: зрение, обоняние, вкус, осязание, ощущение (*овал, лакать, шершавый, фуфло* и т. п.) [6, с. 88]. М. А. Флакман, уточняя классификацию С. В. Воронина, находит четыре способа иконического соотношения между означающим и означаемым, сохраняя базовое деление иконических вербальных единиц на звукоподражания и звукоимовизмы [8].

Определив сущность звукоизобразительности и установив ее основные виды, необходимо выяснить, какими именно единицами должен оперировать инструмент, осуществляющий автоматизированный перевод таких важных, но неочевидных на первый взгляд элементов художественного текста. В данном вопросе мнения исследователей фоносемантики определенным образом расходятся. Так, в ранних работах С. В. Воронина можно встретить определение *фонемотипа*, который подразумевает ряд фонем, обладающих общим или схожим звукоизобразительным значением [6, с. 124–125]. А. П. Журавлев в работе «Фонетическое значение» утверждает, что носителем звукоизобразительного значения является не только звук, но и графема, и вводит термин *звукобуква*, т. е. «написанный звук» [9, с. 18], который также иногда называют *графоном*. С. С. Шляхова предлагает считать носителем звукоизобразительного значения «морфему определенного акустико-артикуляторного значения» [10], а зарубежные исследователи иконичности нередко называют носителями звукоизобразительного значения буквально любые произносительные сущности: фонему, слог, дифференциальный признак или тон, которые выходят за рамки конститутивной функции не несущих никакого лексического значения единиц и начинают выражать какие-либо оттенки смысла [11, с. 228]. В настоящем исследовании носителем звукоизобразительного значения признается фонотип, единица, выделенная в более поздних трудах С. В. Воронина, пришедшая на смену фонемотипу. *Фонотип* понимается как группа звуков, объединенная на основе того или иного признака и неразрывно связанная с каким-либо звукоизобразительным значением [12, с. 102–103]. Ярким примером фонотипа можно признать, например, группу глухих фрикативных согласных, которая представлена в английском и русском языках звуками [f], [ʃ], [s], [x] и обычно ассоциируется с шумовой характеристикой объекта или процесса (*shuffle, hiss, шуршать, свистеть, хохотать* и т. п.).

В силу многообразия возможных фонетических значений (см. ранее), при разработке инструмента автоматизированного перевода было решено сконцентрировать усилия на единицах одной конкретной лексико-семантической группы – английских *глаголах движения* и их переводах на русский язык. Выбор этой группы во многом обусловлен высокой частотностью входящих в нее лексем: невозможно представить художественный текст, в котором глаголы движения встречаются редко. Кроме того, наблюдения показывают, что глаголы движения, которые должны быть классифицированы как звукоимовизмы из-за концептуальной принадлежности денотата, по своей природе являются звукоподражательными, так как боль-

шинство движений, совершаемых в водной или воздушной среде, сопровождаются звучанием (ср. *шлепать, стучать, топнуть, scratch, stomp, flop*), акустические особенности которого зависят от характера движения [13]. Можно предположить, что «двойственный» звукоизобразительный характер глаголов движения предопределяет их экспрессивность и, как следствие, востребованность в литературе, что делает эту лексико-семантическую группу ценным референтным материалом для сравнительного переводческого анализа.

Источником лексического материала послужили списки английских глаголов движения, представленные в работах Л. Ф. Лихомановой [14] и Н. К. Хеймонен [15], а также их русскоязычные эквиваленты, зарегистрированные в словаре В. К. Мюллера [16] и англо-русском веб-словаре Мультигран [17].

Рассматривая звукоизобразительную природу глаголов движения, необходимо обратиться к универсальной классификации ономастов, разработанной С. В. Ворониным. Данная классификация разделяет звукоподражания на несколько крупных классов: инстанты, континуанты, фреквентативы, инстанты-континуанты, фреквентативы-инстанты, фреквентативы-континуанты, фреквентативы-инстанты-континуанты [6, с. 39–70]. *Инстанты* представляют собой звуки, которые воспринимаются человеческим ухом как единичный удар. Они передают стуки, шелканья и прочие отрывистые звуки (*clap, tick, стук, топ*). *Континуантам* характерно длительное, слитное звучание. Чаще всего они связаны с гулом, пищанием, протяжными шумовыми и тоновыми звуками (*meow, hiss, мычать, шипеть*). *Фреквентативы* являют собой диссонанс, т. е. последовательность быстрых ударов, которые не полностью сливаются в процессе восприятия, и обозначают вибрацию, дрожание, резкие движения (*roar, crash, реветь, крошить*). *Инстанты-континуанты* – это сложные звучания, в которых удару сопутствует тон или шум. Они передают резкое движение, падение и столкновения предметов (*bomb, slap, бамс, шлепнуться*). *Фреквентативы-инстанты* и *фреквентативы-континуанты* – такие же сложные звучания, в которых диссонансу сопутствует тон, шум или удар. Они означают хрустяще-трещащие звуки, шорох, скрежет и лязганье (*crumple, creak, хрустеть, шуршать*). *Фреквентативы-инстанты-континуанты* – комплексные звучания, в которых встречаются все три звучания низшего порядка: диссонанс, удар и тон или шум (*strum, бренчать*).

Каждый из приведенных классов ономастов в английском языке получил формульное представление, разработанное С. В. Ворониным. Формулы представляют собой совокупности фонотипов, элементы которых расположены в определенной последовательности, наилучшим образом отражающей устройство слов того или иного класса звукоподражаний [6]. Аналогичные формулы, но уже для русского языка, приводит в работе «Тень смысла в звуке» С. С. Шляхова [18]. Необходимо отметить тот факт, что формулы звукоподражаний для русского и для английского языка нередко совпадают. В рамках разработки автоматизированного инструмента, освещаемого в настоящей статье, формулы звукоподражаний русского языка были сопоставлены формулам звукоподражаний для английского. Кроме того, с опорой на работу М. А. Флакман «Словарь английской звукоизобразительной лексики в диахроническом освещении» [19, с. 13–17], оказалось возможным определить своеобразные ядра каждой из формул, т. е. фонотипы, играющие наиболее важную роль в проявлении звуковой изобразительности данного слова. Представим формулы звукоподражаний для английского и русского языков, а также их ядра.

1. Инстанты:

– *английский*: $\frac{\text{PLOS}(+\text{SON}^{\text{lat/nas/den}})}{\text{AFFR}} + \text{VOC} + \text{PLOS}$;

– *русский*: $\frac{\text{PLOS}}{\text{AFFR}} + \text{VOC} + \frac{\text{PLOS}}{\text{AFFR}}$.

Ядро: звукоподражание должно завершаться гласным фонотипом, за которым следует фонотип смычного или аффрикаты, кроме того, в английском языке необходимо, чтобы гласный фонотип обладал свойством краткости.

2. Континуанты:

а) тоновые континуанты:

– *английский*: $(\text{CONS}(+\text{SON}^{\text{lat/lab}})+)\text{VOC}(+\text{PLOS})$;

– *русский*: $(\text{PLOS}+)\text{VOC}(+\text{PLOS})$.

Ядро: необходимо наличие гласного фонотипа, кроме того, в английском языке необходимо, чтобы гласный фонотип был долгим, в обоих языках приветствуется редупликация.

б) шумовые континуанты:

– *английский*: $(\text{FRIC} + (\text{SON}^{\text{lat/lab}}) + \text{VOC}+) \frac{\text{FRIC}}{\text{SIB}}$;

– *русский*: $(\text{PLOS} + \text{VOC}+) \text{FRIC}^{\text{wk}} (+\text{VOC})$, $(\text{PLOS}+) \text{FRIC}^{\text{str}} (+\text{VOC} + \text{FRIC}^{\text{str}})$.

Ядро: наиболее важен фонотип фрикативного согласного.

3. Фреквентативы:

– *английский*: $\frac{\text{PLOS}}{\text{AFFR}} + \text{VOC} + \text{R}$;

– *русский*: $\text{PLOS}(+\text{VOC}) + \text{R}$.

Ядро: наиболее важен фонотип вибранта, который обыкновенно располагается в конце ономатопа.

4. Инстанты-континуанты;

– *английский*: $\frac{\text{PLOS}(+\text{SON}^{\text{lat/lab}})}{\text{AFFR}} + \text{VOC} + \frac{\text{SON}(+\text{PLOS})}{\text{FRIC}}$,

$\frac{\text{PLOS}(+\text{SON}^{\text{lat/lab}})}{\text{AFFR}} + \text{VOC} + \frac{\text{SON}(+\text{PLOS})}{\text{FRIC}}$, $\text{FRIC}(+\text{SON}^{\text{lat/lab}}) + (\text{VOC}+) \text{PLOS}$,

$\text{FRIC}(+\text{SON}^{\text{lat/lab}}) + \text{VOC} + \text{SON}^{\text{nas}}(+\text{PLOS})$;

– *русский*: $\text{PLOS}(+\text{SON}) + \text{VOC} + \text{SON}$, $\text{PLOS}(+\text{VOC}) + \text{FRIC}$,

$\frac{\text{FRIC}^{\text{wk}} + \text{SON}^{\text{lab/lat}}}{\text{AFFR} + \text{FRIC}^{\text{wk}}} + \text{VOC}(+\text{FRIC}^{\text{wk}}) + \text{PLOS}^{\text{wk}}$, $(\text{FRIC}^{\text{str}}+) \text{FRIC}^{\text{str}} + \text{VOC} + \text{PLOS}^{\text{wk}}$,

$\frac{\text{FRIC}^{\text{wk}}}{\text{AFFR}} + \text{SON}$, $\text{FRIC}^{\text{str}} + \text{PLOS}^{\text{str}} + \text{VOC} + \text{SON}^{\text{lab}}$.

Ядро: наибольший вес имеют смычные, сонорные и глухие щелевые фонотипы.

5. Фреквентативы-инстанты;

– *английский*: (CONS+)R + VOC(+SON^{nas}) + PLOS ;

– *русский*: $\frac{\text{FRIC}}{\text{PLOS}} + \text{R} + \text{VOC}(+\text{FRIC}) + \text{PLOS}$.

Ядро: данному классу звукоподражаний присущи фонотипы смычных и фонотипы вибрантов.

6. Фреквентативы-континуанты;

– *английский*: (CONS +)R + VŌC(+CONS), FRIC(+SON^{lat/lab}) + VOC + R ,
R + VOC + FRIC ;

– *русский*: PLOS + VOC + R , FRIC^{wk} + VOC + R(+PLOS) , FRIC^{str} + VOC + R .

Ядро: наиболее важным является фонотип вибранта, а также фонотип фрикативного согласного (шумовой континуант), кроме того, для английского языка важен фонотип долгого гласного (тоновый континуант).

7. Фреквентативы-инстанты-континуанты.

– *английский*: PLOS + R + VOC + $\frac{\text{FRIC}}{\text{SON}^{\text{nas}}}$ (+PLOS) ;

– *русский*: PLOS + R + VOC + SON^{lab/nas} , $\frac{\text{PLOS}}{\text{FRIC}^{\text{wk}}} + \text{R} + \text{VOC} + \frac{\text{FRIC}^{\text{wk}}}{\text{AFFR}}$,

FRIC^{wk} + VOC + R + PLOS^{wk} .

Ядро: вибрант, сонорный, шумный щелевой, смычный.

Перечислив формулы звукоподражаний для русского и английского языков, представляется необходимым уточнить их чтение. Та или иная формула являет собой лишённую формообразующих суффиксов значимую часть звукоподражания. Последовательность звуков, образующих акустическую оболочку звукоподражания, строится при помощи символа «+». Круглые скобки («()») указывают на факультативность находящихся внутри них фонотипов. В общей последовательности такие фонотипы могут быть пропущены без ущерба для звуковой изобразительности. Символ обелюса («÷») сообщает, что в данной части формулы может быть использован любой из двух приведенных фонотипов (или любая из двух последовательностей фонотипов). Символы верхнего индекса («son^{lab}») указывают на ту или иную особенность фонотипа, например, *латеральный*, *лабиальный* или *назальный* сонорный согласный).

После детального рассмотрения теоретических оснований исследования, представляется возможным перейти к описанию этапов разработки инструмента автоматизации перевода. (краткое изложение основных этапов исследования представлено в работе [20]).

Результаты и обсуждение. Прежде всего, каждый английский глагол движения (из списка Н. К. Хеймонен) снабжается фонематической транскрипцией, словарным толкованием, исчерпывающим набором переводческих эквивалентов, отобранных вручную экспертом-лингвистом в рамках имеющихся словарных толкований, а также фонематическими транскрипциями для каждого из отобранных переводческих эквивалентов. Итоговые наборы данных (английский глагол, его транскрипция, толкование, эквиваленты на русском языке, а также транскрипции эквивалентов) вносятся в базу данных, откуда инструмент по запросу пользователя получает всю необходимую для своих вычислений информацию.

Для хранения вышеперечисленных данных была избрана СУБД (система управления базами данных) *SQLite*. Данная СУБД была выбрана благодаря тому факту, что *SQLite* является легко развертываемой и легко встраиваемой однофайловой системой, которая не требует отдельно запущенного процесса для получения или записи данных, и все операции производятся посредством простых функций, выполнение которых может быть инициализировано в рамках *Python*-скрипта. Структура базы данных включает четыре таблицы: *words*, *translations*, *relations*, а также *classes*.

Таблица *words* предназначена для хранения основной информации по тому или иному английскому слову или словосочетанию и содержит следующие поля: *id* (int, primary key), *word* (varchar(255)), *class* (int, foreign key), *meaning* (varchar(255)), *transcription* (varchar(255)). Наиболее важным является поле *id*, которое хранит уникальный идентификатор лексической единицы. С помощью данного идентификатора инструмент получает список переводческих эквивалентов и их транскрипций из таблицы *translations* для той или иной лексической единицы английского языка. Не менее важным является поле *class*, которое хранит уникальный идентификатор класса из таблицы *classes*. С помощью данного идентификатора инструмент определяет, к какому именно классу звукоподражаний принадлежит английское слово или словосочетание, и какие формулы из ранее приведенных к нему применимы.

Таблица *translations* предназначена для хранения переводческих эквивалентов на русском языке, а также их фонематических транскрипций. Таблица обладает следующими полями: *id* (int, primary key), *translation* (varchar(255)), *transcription* (varchar(255)). Наиболее важным является поле *id*, которое хранит уникальный идентификатор переводческого эквивалента. Данный идентификатор обеспечивает двустороннюю связь между эквивалентом и оригинальным словом.

Таблица *relations* связывает слова или словосочетания из таблицы *words* с их переводами из таблицы *translations* и являет собой таблицу связей. Таблица содержит следующие поля: *word_id* (int, foreign key) и *translation_id* (int, foreign key). Значения в *word_id* указывают на строки таблицы *words*, а значения в *translations_id* – на строки таблицы *translations*. Таким образом формируется двухсторонняя связь. Лексические единицы на английском языке и их переводы не могут существовать в рамках одной таблицы и нуждаются в таблице-посреднике (*relations*), потому что, как показывает практический опыт, переводческие эквиваленты для разных слов или словосочетаний на английском языке регулярно повторяются. Таким образом, хранение переводческих эквивалентов в отдельной таблице позволяет избегать дублирования данных, что облегчает разработку приложения и экономит память устройства.

Заключительная таблица *classes* предназначена для хранения перечня звукоподражательных классов и обладает следующими полями: *id* (int, primary key), *class_name* (varchar(255)). Поле *id* хранит уникальные идентификаторы классов, обеспечивающие связь таблицы *words* (по полю *class*) непосредственно с таблицей *classes*.

Помимо хранения лексических единиц и их переводов необходимо рассмотреть хранение и использование упомянутых ранее фонотипов. Вследствие того, что набор фонотипов статичен и не занимает большого объема памяти (в отличие от множества английских глаголов и их эквивалентов на русском языке), хранение фонотипов осуществляется посредством обыкновенного *Python*-словаря, который импортируется основным скриптом и не требует при этом отдельного запроса к базе данных (таблица).

Визуализация словаря фонотипов
Phonotype dictionary visualization

ID	Символы транскрипции	Формульное представление
0	a, i, o, e, u, y	VOC (гласный)
1	b, d, g	PLOS ^{str} (смычный звонкий)
2	p, t, k	PLOS ^{wk} (смычный глухой)
3	v, z, ž, j	FRIC ^{str} (фрикативный звонкий)
4	f, s, š, x	FRIC ^{wk} (фрикативный глухой)
5	c, č	AFFR (аффриката)
6	r	R (вибрант)
7	m	SON ^{lab} (сонорный губной)
8	n	SON ^{nas} (сонорный носовой)
9	l	SON ^{lat} (сонорный боковой)

В основе работы автоматизированного инструмента лежит *Python*-скрипт, который получает на вход слово, словосочетание или предложение на английском языке, обнаруживает в нем лексическую единицу, которая уже существует внутри базы данных, после чего ранжирует ее переводческие эквиваленты, ориентируясь на близость их фонематических транскрипций к одной из звукоподражательных формул оригинального слова. Итогом работы инструмента становится ранжированный список переводческих эквивалентов для каждой обнаруженной лексической единицы, в котором чем выше располагается слово, тем более близким оно является по своим звукоизобразительным свойствам к оригиналу.

Итоговый инструмент автоматизации перевода английских звукоизобразительных единиц представляет собой комплексное веб-приложение, состоящее из нескольких компонентов: клиентская сторона, программный интерфейс (API), *Python*-скрипт фоносемантического анализа, а также база данных *SQLite*. Формат веб-приложения имеет следующие достоинства: простота разработки, четкое разграничение элементов (интерфейс существует отдельно от аналитических скриптов), высокая доступность (инструмент можно развернуть на удаленном сервере, обеспечив к нему доступ из глобальной сети Интернет). Последовательная работа инструмента представлена на приведенных далее рисунках. Так, на рис. 1 изображены основные компоненты, образующие собой инструмент.

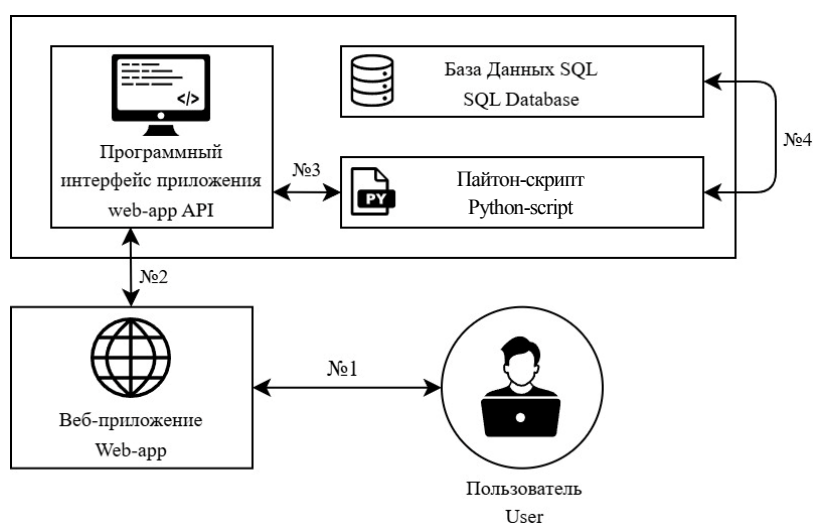


Рис. 1. Схема автоматизированного инструмента
Fig. 1. Automated tool design

Шаг 1 (рис. 1): пользователь получает доступ к инструменту посредством того или иного веб-браузера и вводит в текстовое поле лексическую единицу для последующего анализа. Из-за того, что инструмент имеет встроенный обработчик запросов на естественном языке, в роли лексической единицы может выступать как любой английский глагол или словосочетание, так и целое предложение или даже полноценный текст.

Клиентская сторона веб-приложения написана на языке программирования *JavaScript* с использованием фреймворков *Vue.js* и *QuasarFramework*. Данная комбинация инструментов позволяет создать современные и удобные пользовательские интерфейсы с минимальными трудозатратами.

Шаг 2 (рис. 1): клиентская сторона веб-приложения забирает введенный пользователем текстовый материал, производит его валидацию и отправляет программному интерфейсу веб-приложения в формате *POST*-запроса.

Программный интерфейс веб-приложения представлен простейшим веб-контроллером, написанным на языке программирования *Python* с использованием фреймворка *Flask*. Данная комбинация инструментов обеспечивает простоту разработки и, что особенно важно, позволяет избежать более распространенных решений, выполненных на языке программирования *PHP*. Таким образом, область разработки инструмента требует использования только двух языков программирования (*Python* и *JavaScript*), а не трех.

Шаг 3 (рис. 1): программный интерфейс веб-приложения обрабатывает запрос, полученный от клиента, т. е. проводит повторную валидацию данных, после которой запускает *Python*-скрипт фоносемантического анализа.

Python-скрипт фоносемантического анализа осуществляет непосредственно перевод звукоизобразительных единиц, т. е. производит ранжирование переводческих эквивалентов оригинала в зависимости от их звуковой изобразительности. Шаги, выполняемые инструментом на данной стадии, представлены на рис. 2.

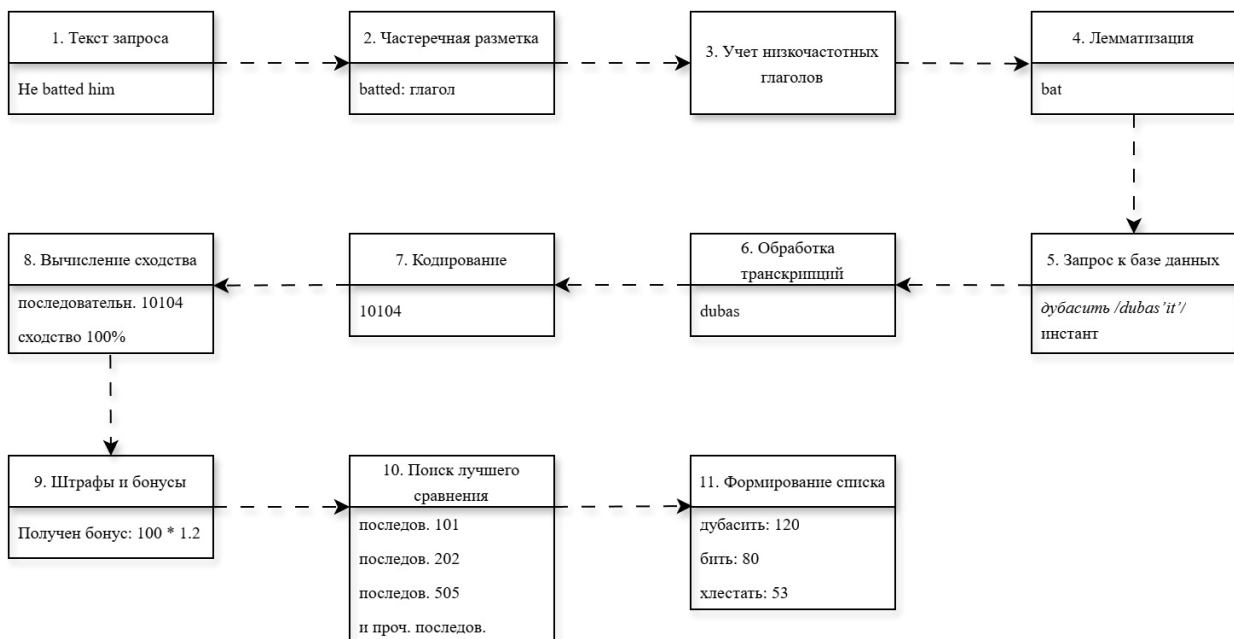


Рис. 2. Шаги скрипта фоносемантического анализа
Fig. 2. Phonosemantic analysis steps

Представляется необходимым рассмотреть, какие именно операции выполняет скрипт *фоносемантического анализа*, чтобы определить наиболее соответствующие по типу звукоизобразительности русскоязычные эквиваленты исходных английских глаголов движения.

Блок 1: входными данными *Python*-скрипта является текстовый ввод пользователя, прошедший все этапы валидации (слово, словосочетание, предложение, текст), в данном случае – *she batted him*.

Блок 2: перед измерением звуковой изобразительности необходимо обнаружить в текстовом вводе пользователя все лексические единицы, с которыми работает инструмент (в настоящей версии это глаголы движения). С использованием тяжеловесной языковой модели, предоставленной библиотекой *spaCy*, производится частеречная разметка текстового ввода. Таким образом скрипт оставляет для последующих операций только необходимые части речи, т. е. глаголы, в данном случае – *batted*.

Блок 3: нередки случаи, когда языковая модель ошибочно распознает глагол в качестве существительного (пример: *barrel* (сущ.) и *to barrel* (гл.)) или другой части речи. Чтобы обойти некорректную частеречную разметку и «заставить» инструмент смотреть на то или иное слово исключительно в качестве глагола, предусмотрена специальная проверка. Так, когда ни один глагол в запросе пользователя найден не был, скрипт проверяет, является ли первое слово в последовательности частицей *to*. Если частица *to* найдена, последующему слову сразу же присваивается глагольный статус.

Блок 4: обнаруженные глаголы подвергаются операции лемматизации, которая осуществляется с использованием библиотеки *spaCy*. Результатом данной операции становятся леммы (словарные формы) обнаруженных глаголов. Во избежание повторения словарных форм результаты лемматизации записываются внутрь множества (структура данных «множество» хранит только уникальные элементы).

Блок 5: данный блок соответствует **Шагу 4** работы инструмента (рис. 1). Скрипт *фоносемантического анализа* обращается к базе данных, осуществляя поиск по таблице *words* тех слов, которые совпадают с леммами глаголов внутри множества. Таким образом скрипт отделяет глаголы движения от словарных единиц прочих лексико-семантических групп. Далее, инструмент, используя *id* обнаруженных в базе глаголов движения, получает список их переводческих эквивалентов из таблицы *translations*. По завершении данного блока скрипт располагает следующей информацией о каждом глаголе движения: слово, транскрипция, класс звукоподражания, толкование, список переводческих эквивалентов и их транскрипций.

Блок 6: производится предобработка фонематических транскрипций переводческих эквивалентов – удаляются формообразующие суффиксы и символы палатализации. Данная операция оправдана тем фактом, что небольшое число формообразующих суффиксов является общим для большого числа слов. Кроме того, формообразующие суффиксы не способствуют возникновению звуковой изобразительности (по крайней мере, того типа, что является предметом рассмотрения в настоящем исследовании). Палатализация же не учитывается в имеющихся формулах звукоподражаний (ср. [6; 16]), из-за чего в настоящей версии инструмента ее приходится игнорировать. Таким образом, удаление данных элементов сродни очистке транскрипций от шумов.

Блок 7: преобработанные транскрипции переводческих эквивалентов кодируются по словарю фонотипов (табл. 1) и превращаются в простые числовые последовательности (/dubas/ (дубасить) → 10104).

Блок 8: вычисление сходства между числовой последовательностью эквивалента и формулами того класса звукоподражаний, к которому принадлежит оригинальное слово, с использованием расстояния Левенштейна. Рассмотрим работу данного блока на конкретном примере. Последовательность 10104 (дубасить), формула $\frac{PLOS}{AFFR} + VOC + \frac{PLOS}{AFFR}$, вариации формулы, закодированные с использованием словаря фонотипов: 101, 202, 505, 102, 105, 201, 205, 501, 502. Вариация формулы 101 целиком входит в числовую последовательность эквивалента 10104, следовательно, сходство равно 100 %. Вариация 102 не полностью входит в числовую последовательность 10104, необходимо сделать одну замену, следовательно, сходство равно 67 %.

Блок 9: применение системы штрафов и бонусов. Если *ядро* формулы присутствует в числовой последовательности эквивалента, сходство получает двадцатипроцентный бонус. Если *ядро* отсутствует, то сходство получает двадцатипроцентный штраф. Ядро формулы $\frac{PLOS}{AFFR} + VOC + \frac{PLOS}{AFFR}$ – это 0 (гласный фонотип) и 1 (звонкий смычный) или 2 (глухой смычный) или 5 (аффриката). Последовательность 10104 эквивалента дубасить включает 0 и включает 1, значит, вариация 101 со сходством 100 % получает бонус и становится 120 %. Вариация 102 со сходством 67 % тоже получает бонус, т. к. принадлежит к этой же формуле, и становится 80,4 % (67*1,2). Размер штрафа/бонуса можно изменить вручную для каждой из формул.

Блок 10: поиск вариации, сравнение числовой последовательности с которой показывает наилучшее сходство. 120 % (вариация 101) больше, чем 80,4 % (вариация 102), значит эквивалент дубасить имеет 120 баллов звуковой изобразительности (конвертация процентов в баллы здесь условная и не играет важной роли).

Блок 11: составление списка переводческих эквивалентов, их звукоизобразительных баллов и его сортировка. Так, например, дубасить имеет 120 баллов, ударять – 80, а хлестать – 53.

Получившейся список передается программному интерфейсу веб-приложения, который отправляет его клиентской стороне, после чего в окне веб-браузера можно увидеть каждый из обнаруженных глаголов движения, их транскрипции и толкования, в рамках которых отбирались переводы, а также таблицы со списками эквивалентов, отсортированных по убыванию схожести звукоизобразительности этих эквивалентов и английского глагола движения, использованного в оригинальном тексте. Окончательное решение о выборе того или иного русскоязычного эквивалента остается за переводчиком.

Заключение. Представленный автоматизированный инструмент – помощник переводчика – использует уникальные методы отбора переводческих эквивалентов, а предлагаемые им варианты перевода обладают достаточной степенью звуковой изобразительности. Разработанный инструмент способен впоследствии стать полезным пополнением в арсенале любого литературного переводчика, так как умеет визуализировать звуковую изобразитель-

ность лексических вариантов перевода, чем помогает специалисту выбрать из предложенного списка именно тот эквивалент, который наилучшим образом соответствует потребностям перевода и особенностям текста.

В дальнейшем планируется продолжить разработку инструмента в нескольких направлениях, в том числе: расширить область работы инструмента на другие группы звукоизобразительной английской лексики; автоматизировать первичный отбор переводческих эквивалентов по их лексической близости к оригиналу; автоматизировать транскрибирование отобранных переводческих эквивалентов; включить в используемый алгоритм учет ближайшего контекста иконической лексической единицы, подлежащей переводу; оптимизировать формулы классов русскоязычных звукоподражаний для наиболее точного сопоставления с англоязычными моделями звукоподражательных слов.

СПИСОК ЛИТЕРАТУРЫ

1. Силюнская Н. П. Аллюзии в художественном тексте как проблема перевода (на материале романов А. Брэдли и их переводов на русский язык) // Тез. докл. XLVIII Междунар. филол. науч. конф., СПб., 18–19 марта 2019 г. / СПбГУ. СПб., 2019. URL: <https://conference-spbu.ru/en/conference/40/reports/11196/> (дата обращения: 10.01.2026).
2. Шамина Е. А. К вопросу о восприятии синестетических метафор в литературном тексте: экспериментальное исследование // Искусство звука и света: Галеевские чтения – 2023: материалы XX Междун. науч.-практ. конф. «Прометей-2023», Казань, 30 июня – 4 июля 2023 г. / Казань: Изд-во АНРТ, 2023. С. 347–351.
3. Беседина Е. И. К вопросу о функционировании звукоизобразительных глаголов движения (на материале романа Брайана Джейкса «Воин Редволла») // Актуальные проблемы языкознания. 2019. Т. 1. С. 239–243.
4. Комиссаров В. Н. Теория перевода (лингвистические аспекты): учебник. М.: Высш. шк., 1990.
5. Shamina E. A., Besedina E. I., Veselova M. V. By Any Other Name Would It Smell Just as Sweet? Or How to Learn to Translate Iconic Lexis // *The Magic of Innovation: Outlook for Linguistics and Language Teaching in the Present Context*. 2024. Vol. 6, № 2. P. 635–641. DOI: 10.24833/2949-6357.2024.ShBV.2
6. Воронин С. В. Основы фоносемантики. 2-е изд., стереотип. М.: ЛЕНАНД, 2006.
7. Magnus M. What's in a word? Studies in Phonosemantics: Dr. Sci. (Philosophy) Thesis / NTNU. Trondheim, 2020.
8. Флакман М. А. Классификация звукоизобразительной лексики по способу имитации // LIII Междунар. филол. науч. конф. им. Л. А. Вербицкой: тез. докл. СПб., 25–31 марта 2025 г. / СПбГУ. СПб., 2025. С. 994–995.
9. Журавлев А. П. Фонетическое значение. Л.: ЛГУ, 1974.
10. Шляхова С. С. О состоянии фоносемантики в России. Статья третья. Российские фоносемантические школы и центры // Вопросы психолингвистики. 2020. № 2 (44). С. 124–145. DOI: 10.30982/2077-5911-2020-44-2-124-145
11. Nuckolls J. B. The case for sound symbolism // *Annual Review of Anthropology*. 1999. Vol. 28. P. 225–252. DOI: <https://doi.org/10.1146/annurev.anthro.28.1.225>
12. Voronin S. V. Iconicity. Glottogenesis. Semiosis: sundry papers. SPb.: St. Petersburg State Univ., 2005.
13. Шамина Е. А. Звукоподражательная валентность глаголов движения (на материале английского языка) // 50-я Междунар. филол. науч. конф. им. Л. А. Вербицкой: тез. докл., СПб., 15–23 марта 2022 г. / СПбГУ. СПб.: Изд-во СПбГУ, 2022. С. 306–307.

14. Лихоманова Л. Ф. Семантическая филиация английских звукоизобразительных глаголов движения: автореф. дис. ... канд. филол. наук / ЛГУ. Ленинград, 1986.
15. Хеймонен Н. К. Звукоподражательная валентность английских глаголов движения: дипломная работа бакалавра / СПбГУ. СПб., 2019.
16. Мюллер В. К. Самый полный англо-русский словарь с современной транскрипцией. М.: АСТ, 2020.
17. Мультитран. URL: <https://www.multitran.com/> (дата обращения: 07.05.2026).
18. Шляхова С. С. Тень смысла в звуке. Введение в русскую фоносемантику: учеб. пособие. Пермь: Изд-во ПГПУ, 2003.
19. Флакман М. А. Словарь английской звукоизобразительной лексики в диахроническом освещении. СПб.: Изд-во РХГА, 2016.
20. Шамина Е. А., Голованов Е. О. Звукоизобразительные глаголы движения в художественной литературе: виртуальный помощник переводчика // Современные исследования звучащей речи 2025: сб. науч. ст. Минск: Изд-во БГУИЯ, 2024. С. 58–63.

Информация об авторах

Голованов Егор Олегович – аспирант кафедры иностранных языков Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В. И. Ульянова (Ленина), ул. Профессора Попова, д. 5Ф, Санкт-Петербург, 197022, Россия. Автор трех научных публикаций. Сфера научных интересов: компьютерная лингвистика, фоносемантика, перевод.

Шамина Елена Анатольевна – кандидат филологических наук (1988), доцент кафедры фонетики и методики преподавания иностранных языков филологического факультета Санкт-Петербургского государственного университета, Университетская наб., д. 7–9, Санкт-Петербург, 199034, Россия. Автор более 70 научных публикаций. Сфера научных интересов: фонетика, методика преподавания иностранных языков, фоносемантика.

*О конфликте интересов, связанном с данной публикацией, не сообщалось.
Поступила 27.01.2026; принята после рецензирования 01.04.2026; опубликована онлайн 23.06.2026.*

REFERENCES

1. Silinskaya, N.P. (2019), "Allusions in fiction as a translation problem (based on A. Bradley's novels and their translations into Russian)", *XLVIII Int. Sci. Philological conf.*, SPb., RUS, 18–19 March 2019, available at: <https://conference-spbu.ru/en/conference/40/reports/11196/> (accessed 10.01.2026).
2. Shamina, E.A. (2023), "On the issue of perception of synesthetic metaphors in a literary text: an experimental study", *The Art of Sound and Light: Gale Readings – 2023, Proc. of the XX Int. Sci. and Practical Conf.*, (Prometheus-2023), Kazan, RUS, 30 June – 4 July 2023, pp. 347–351.
3. Besedina, E.I. (2019), "On the issue of iconic verb functioning in James Brian Jacques's novel "Redwall"", *Current Issues of Linguistics – CIL*, vol. 1, pp. 239–243.
4. Komissarov, V.N. (1990), *Teoriya perevoda (lingvisticheskie aspekty)* [Translation theory (linguistic aspects)], Vysshaya shkola, Moscow, USSR.
5. Shamina, E.A., Besedina, E.I. and Veselova, M.V. (2024), "By Any Other Name Would It Smell Just as Sweet? Or How to Learn to Translate Iconic Lexis", *The Magic of Innovation: Outlook for Linguistics and Language Teaching in the Present Context*, vol. 6, no. 2, pp. 635–641. DOI: 10.24833/2949-6357.2024.ShBV.2
6. Voronin, S.V. (2006), *Osnovy fonosemantiki* [Fundamentals of phonosemantics], LENAND, Moscow, RUS.
7. Magnus, M. (2020), "What's in a word? Studies in Phonosemantics", Dr. Sci. (Philosophy) Thesis, NTNU, Trondheim, NOR.
8. Flaksman, M.A. (2025), "Classification of iconic vocabulary based on the method of imitation", *LIII Int. Sci. Philological Conf.*, SPb., RUS, 25–31 March 2025, pp. 994–995.

9. Zhuravlev, A.P. (1974), *Foneticheskoe znachenie* [Phonetic meaning], Leningrad State Univ., Leningrad, USSR.
10. Shlyakhova, S.S. (2020), "On phonosemantics in Russia. Article #3. Russian phonosemantic science schools and centres", *J. of Psycholinguistics*, no. 2 (44), pp. 124–145. DOI: 10.30982/2077-5911-2020-44-2-124-145
11. Nuckolls, J. B. (1999), "The case for sound symbolism", *Annual Review of Anthropology*, vol. 28, pp. 225–252. DOI: <https://doi.org/10.1146/annurev.anthro.28.1.225>
12. Voronin, S.V. (2005), *Iconicity. Glottogenesis. Semiosis: sundry papers*, SPbU, SPb., RUS.
13. Shamina, E.A. (2022), "Onomatopoeic valency of motion verbs in English", *50th Int. Sci. Philological Conf.*, SPb., RUS, 15–23 March 2022, pp. 306–307.
14. Likhomanova, L.F. (1986), "Semantic filiation of English onomatopoeic verbs of motion", Abstract of Can. Sci. (Philology) dissertation, Leningrad State Univ., Leningrad, USSR.
15. Kheimonen, N.K. (2019), "Onomatopoeic valency of English verbs of motion", Bachelor's Thesis, SPbSU, SPb., RUS.
16. Myuller, V.K. (2020), *Samyi polnyi anglo-russkii slovar' s sovremennoi transkriptsiei* [The most complete English-Russian dictionary with modern transcription], AST, Moscow, RUS.
17. *Multitran*, available at: <https://www.multitran.com/> (accessed 07.05.2026).
18. Shlyakhova, S.S. (2003), *Ten' smysla v zvuke. Vvedenie v russkuyu fonosemantiku* [The Shadow of Meaning in Sound: An Introduction to Russian Phonosemantics], PGPU, Perm, RUS.
19. Flaksman, M.A. (2016), *Slovar' angliiskoi zvukoizobrazitel'noi leksiki v diakhronicheskom oshchchenii* [Dictionary of English sound-symbolic vocabulary in diachrony], Izd-vo RHGA, SPb., RUS.
20. Shamina, E.A. and Golovanov, E.O. (2025), "Iconic verbs of motion in fiction writing: the translator's virtual assistant", *Sovremennye issledovaniya zvuchashchej rechi* [Contemporary Studies of Oral Speech], Belorussia State Univ. of Foreign Languages, Minsk, BLR, pp. 58–63.

Information about the authors.

Egor O. Golovanov – Postgraduate Student at the Department of Foreign Languages, Saint Petersburg Electrotechnical University, 5F Professor Popov str., St Petersburg 197022, Russia. The author of 3 scientific publications. Area of expertise: computational linguistics, phonosemantics, translation.

Elena A. Shamina – Can. Sci. (Philology, 1988), Associate Professor at the Department of Phonetics and Foreign Language Teaching Methodology, Saint Petersburg State University, 7/9 University emb., St Petersburg 199034, Russia. The author of more than 70 scientific publications. Area of expertise: phonetics, foreign language teaching methodology, phonosemantics.

*No conflicts of interest related to this publication were reported.
Received 27.01.2026; adopted after review 01.04.2026; published online 23.06.2026.*