

УДК 81'33 Прикладная лингвистика
DOI: 10.26907/2782-4756-2024-77-3-19-25

КОЛИЧЕСТВЕННАЯ ХАРАКТЕРИСТИКА ПРОЦЕССА ИНТЕРНАЦИОНАЛИЗАЦИИ РУССКОЙ ЛЕКСИКИ

© Тимур Галеев, Владимир Бочкарев, Мэняо Ван

QUANTITATIVE CHARACTERISTICS OF RUSSIAN VOCABULARY INTERNATIONALIZATION

Timur Galeev, Vladimir Bochkarev, Mengyao Wang

The article is devoted to the competition of borrowed and native vocabulary in the Russian language in the 18th–20th centuries. The relevance of the study: due to the large number of non-native speakers, the etymological picture of Russian vocabulary is somewhat shifted towards the dominance of internationalisms, characterized by their universality and conventionality in most languages. This fact predetermines a spontaneous transformation of the typical lexical tools of scientific, official business and journalistic speech, and calls for the need to change the lexicographic norms.

Our analysis has shown that internationalisms, denoting already existing concepts, can displace native words, especially in literary speech. The article identifies linguistic patterns in lexical combinability of native and borrowed words, and determines stylistic differences between them. When working with text corpora, we have revealed discrepancies in syntactic combinability between synonyms, which allows us to state the stylistic demarcation of borrowings and native Russian words. Etymological synharmonism and connotation show the way words of different genesis form context. We studied the change in the frequency of native and borrowed words' usage in diachrony using text corpora.

As a result, we have come to the following conclusions: borrowed words have a clearer semantic orientation in the Russian language; they are characterized by a lower frequency and narrow combinability. We have analyzed the cases of native lexemes' displacement by foreign-language borrowings from the point of view of their frequency growth dynamics in the corpus. We have determined the average duration of the period, during which a foreign word exceeds its analog frequency of the Slavic origin in book speech – 40-60 years. The results obtained are important for studying the deep changes in society, since we are talking about internationalisms that denote already existing social phenomena or objects.

Keywords: loanwords, internationalisms, native vocabulary, etymological picture, corpus analysis, semantics, stylistics

Статья посвящена конкуренции заимствованной и исконной лексики в русском языке на протяжении XVIII–XX вв. Актуальность исследования: ввиду большого количества говорящих и пишущих инофонов размывается этимологическая однородность картины русской лексики в сторону засилья интернационализмов в силу их универсальности и конвенциональности. Это предопределяет стихийное преобразование типичного лексического инструментария научной, официально-деловой и публицистической речи, а также необходимость смены норм в лексикографии.

Анализ показывает, что интернационализмы, обозначающие уже существующие понятия, могут вытеснять исконные слова, особенно в книжной речи. Авторы выявляют закономерности в сочетаемости исконных и заимствованных слов, а также определяют стилистические различия между ними. Работа с корпусами текстов также выявляет расхождения в синтаксической сочетаемости между синонимами, что позволяет говорить о стилистическом размежевании заимствований и исконно русских слов. Этимологический сингармонизм и коннотация показывают, как слова различного генеза формируют контекст. Исследуется изменение частотности употребления исконных и заимствованных слов в диахронии с использованием корпусов текстов.

В результате было выявлено, что заимствованные слова имеют более чёткую семантическую структуру в русском языке, но характеризуются более низкой частотностью и узкой сочетаемостью. С учётом хронологии вхождения иноязычных слов в русский язык и их частотности в книжной речи была определена средняя продолжительность периода, во время которого иностранное слово превосходит по частотности свой аналог славянского происхождения в книжной речи – от 40 до 60 лет. Полученные результаты имеют важное значение для изучения глубинных изменений

общества, которые влияют на характер и результат конкуренции интернационализмов, обозначающих существующие общественные явления, с их аналогами исконно русского происхождения.

Ключевые слова: заимствованные слова, интернационализмы, исконная лексика, этимологическая картина, корпусной анализ, семантика, стилистика

Для цитирования: Галеев Т., Бочкарев В., Ван Мэняо. Количественная характеристика процесса интернационализации русской лексики // Филология и культура. Philology and Culture. 2024. № 3 (77). С. 19–25. DOI: 10.26907/2782-4756-2024-77-3-19-25

Многовековой социальный опыт наравне с инновационными идеями непрерывно передаются от народа к народу, из общества в общество, оставляя на своём пути приметный этимологический след в самых разных языках. Обширный пласт слов русского языка восходит к этимонам из распространённых языков древнего мира (древнегреческий, латинский, арабский) благодаря языкам-посредникам европейских и тюркских народов. Многочисленные латинизмы и грецизмы, а в последнее время ещё и англицизмы, не только обогащают русский язык, но и играют важную роль в книжной речи. При этом часть заимствованных слов иноязычного происхождения по значению и употреблению может иметь синонимы, определяемые носителями языка как исконная лексика. Одной из форм языковой вариативности принято считать равнозначное использование исконного и заимствованного слова для обозначения одного понятия.

Существует обширная лексическая категория, охватывающая распространённые слова, которые, приходя в тот или иной язык, начинают конкурировать по частотности и лексическому объёму с существующими в языке словами исконного происхождения. Лингвисты называют такую категорию интернациональными словами (интернационализмами): *ресурс / запас, негативный / отрицательный, компенсировать / возмещать, публиковать / издавать*. Со временем такая конкуренция стремится к разрешению в семантическом (*посещение / визит, секрет / тайна, стабильный / устойчивый, идея / мысль / понятие*) или стилистическом плане (*рекомендовать / советовать, инфекционный / заразный, термометр / градусник*). Представляется, что наличие интернационализмов и способ их конкуренции с исконными словами свидетельствуют и о появлении новых явлений в обществе, и о глубинных изменениях самого общества, особенно если речь идёт об интернационализмах, обозначающих уже существующие общественные явления или предметы (*гомосексуалист / мужеложец, суицид / самоубийство*) (см. рис. 1).

Актуальность изучения вариативности определяется также тем, что русский язык, уже выйдя на уровень языка международного общения (хотя

и относительно компактного – СНГ и страны Восточной Европы), со временем набирает большую популярность в государствах Юго-Восточной Азии и Латинской Америки, становясь по-настоящему глобальным языком. В связи с этим ввиду большого количества говорящих и пишущих инофонов несколько смещается этимологическая картина русской лексики в сторону засилья интернационализмов благодаря их универсальности и конвенциональности в большинстве языков мира. Это предопределяет стихийное преобразование типичного лексического инструментария научной, официально-деловой и публицистической речи, а также необходимость смены норм в словарях, учебниках русского языка.

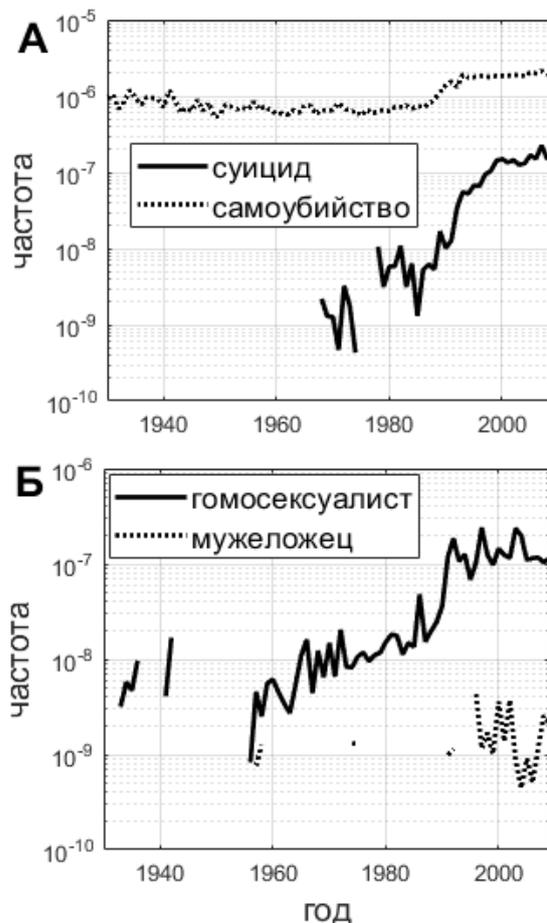


Рис. 1. Корреляция частотности в диахронии слов *суицид* и *самоубийство* (А); *гомосексуалист* и *мужеложец* (Б).

С целью изучения конкуренции исконных лексем с интернационализмами мы составим выборку самых распространённых пар «заимствованное / исконное слово» по словарям и справочникам [1], [2], [3]. Далее, применяя количественный метод на материале корпуса текстов «Гугл Книги» [4], мы сравним частотность иноязычных лексем и их русских эквивалентов в разные исторические периоды, а также, пользуясь семантической разметкой ГИКРЯ (Генеральный Интернет-Корпус Русского Языка) [5], проанализируем степень близости значений слов в каждой синонимической паре, а также различия в комбинаторных свойствах каждого её члена.

К применяемым методам исследования необходимо также отнести семантическую интерпретацию языковых единиц, классификацию, сопоставление синтаксических моделей, статистическую обработку материалов.

Основным материалом данного исследования будут данные корпусов «Гугл Книги Ngram» [4] и «Открытый корпус русского языка» [5].

Корпус «Гугл Книги Ngram» содержит 67 миллиардов словоупотреблений на материале 600 тыс. размеченных текстов книг на русском языке из 40 университетских библиотек разных стран мира. «Открытый корпус русского языка» (OpenCorpora) содержит тексты различных жанров и стилей, что позволяет использовать его для проведения лексикологических исследований. Его преимуществами являются объём корпуса (более 1 миллиарда словоупотреблений); возможность поиска по ключевым словам, частям речи, грамматическим признакам и другим параметрам; наличие разметки, которая позволяет анализировать семантические, синтаксические и другие особенности текстов.

По графикам, построенным уже на основе тысяч словоупотреблений, мы можем со значительной уверенностью судить о степени освоенности иноязычного слова. К тому же, именно корпусные данные могут помочь исследователям с примерной датировкой вхождения слова, в том числе заимствованного, в русский язык.

К применяемым методам исследования необходимо также отнести морфологическую классификацию языковых единиц, их семантическую интерпретацию, сопоставление синтаксических моделей, в которых функционируют упомянутые заимствования.

Исследование пар конкурирующих лексем как элемента языковой избыточности представляет стабильный интерес в русистике, а также корпусной и теоретической лингвистике. На тему проникновения в различные языки интерна-

циональных слов написано немало работ, в которых поднимались вопросы эволюции языка [6] и психологии личности [7] в условиях глобализации, возможных семантических преобразований в языке-реципиенте [8], стилистической окраски заимствований [9], их коннотации [10]. Дать исчерпывающий обзор литературы в области исследования языковых контактов и эволюции языка не представляется возможным, поэтому мы остановимся на нескольких наиболее интересных с нашей точки зрения работах.

Х. Пфандль, изучая влияние интернационализмов на немецкий и славянские языки, указывает на зависимость «успеха» адаптации заимствования в языке от следующих факторов: краткость интернационализма; прозрачность его внутренней формы для носителей принимающего языка; степень освоенности языковым коллективом того явления, которое обозначается заимствованием; экстралингвистические факторы, в частности – взаимодействие государства и СМИ [11].

В Юго-Восточной Азии, где китайский язык выступает в качестве регионального языка международного общения, при изучении влияния китайизмов на малайский язык было использовано несколько китайских лингвистических корпусов для сбора лексических данных, морфологической и семантической разметки и аннотирования текстов [12].

Мы отобрали 125 наиболее частотных заимствований из древнегреческого, латинского и арабского языков, из них 73 существительных, 29 глаголов, 23 прилагательных и наречия.

В паре *дискуссия / спор* с заимствованным словом наиболее часто сочетаются прилагательные *острый, широкий и оживлённый*, а с русизмом – *горячий*. Необходимо отметить, что слова почти не имеют точек соприкосновения в плане сочетаемости. Представляется, что это является иллюстрацией разницы в семантике между заимствованиями и исконными словами, по крайней мере, в книжной речи.

В паре *бизнесмен (молодой, преуспевающий, талантливый)* и *предприниматель (частный, обычный, простой)* сочетаемость отражает коннотацию и несёт аксиологическую нагрузку, то есть выражает позицию автора. Следует отметить, что слово *предприниматель* в корпусе «Гугл Книги» значительно более частотно за счёт обязательного использования его в юридических текстах. *Бизнесмен* же, несмотря на сходный объём понятия, не применяется в юридических и экономических текстах, оставаясь только в языке СМИ.

В паре *спиртной / алкогольный* заимствованное из арабского языка прилагательное гораздо шире комбинаторно и, помимо существительных *напитки, изделия, пары* (бывают и *спиртные, и алкогольные*), сочетается также с существительными *психоз, галлюцинация, эксцесс, раствор* (только *алкогольные*). Это редкий случай для русского языка, когда заимствованное слово демонстрирует более высокую частотность и широкую сочетаемость. Следует отметить, что дополнительную лексическую сочетаемость прилагательное *алкогольный* приобретает именно в медицинской сфере.

На основе данных о частотности и сочетаемости прилагательных *положительный* и *позитивный* мы можем заметить, что с обоими прилагательными наиболее часто сочетаются следующие существительные: *ответ, опыт, результат, эффект*. Стоит отметить, что существительные имеют схожую сочетаемость, однако её же можно считать устойчивой и исторически сложившейся. Например, можно сказать *положительный заряд* (*позитивный заряд* почти не употребляется), но при этом *положительная оценка* (*позитивная* – нельзя). Следовательно, лексема славянского происхождения многократно превосходит по частотности интернационализм и, согласно лексикографии, при равной сочетаемости обладает большим количеством лексических значений. Анализ комбинаторных свойств лексемы позволил определить стилистическую маркировку «науч.» для заимствований и их славянских аналогов.

Рассмотрим пару *аналогичный / подобный*. Анализ заимствованного прилагательного *аналогичный* в корпусе «Гугл Книги Ngram» показывает его сочетаемость с существительными *показатель, результат, процесс, эффект, случай* и др. С исконным словом *подобный* часто сочетаются существительные *случай, вещи, труд, вопрос, взгляд, порядок* и т. д. Разная сочетаемость отражает стилистическое размежевание синонимов: заимствование используется в сочетании с более книжными словами (*аналогичный результат*), а с русизмом – слова стилистически нейтральные (*подобный случай*). Стоит отметить, что у словосочетания *подобный случай* более высокая частотность.

Конкуренция среди наречий (*абсолютно / совершенно, абсолютно / совсем*) также заслуживает внимания исследователей. По сравнению с заимствованным словом *абсолютно* исконно-русские лексемы *совсем* и *совершенно* обычно сочетаются с наречиями (*совсем недавно, совсем рядом, совершенно неожиданно, совершенно невозможно*). Заимствованное же слово *абсолютно*

в основном стоит в препозиции к прилагательным: *абсолютно сухой, чёрный*. Таким образом, мы наблюдаем расхождение синонимов в плане синтаксической сочетаемости.

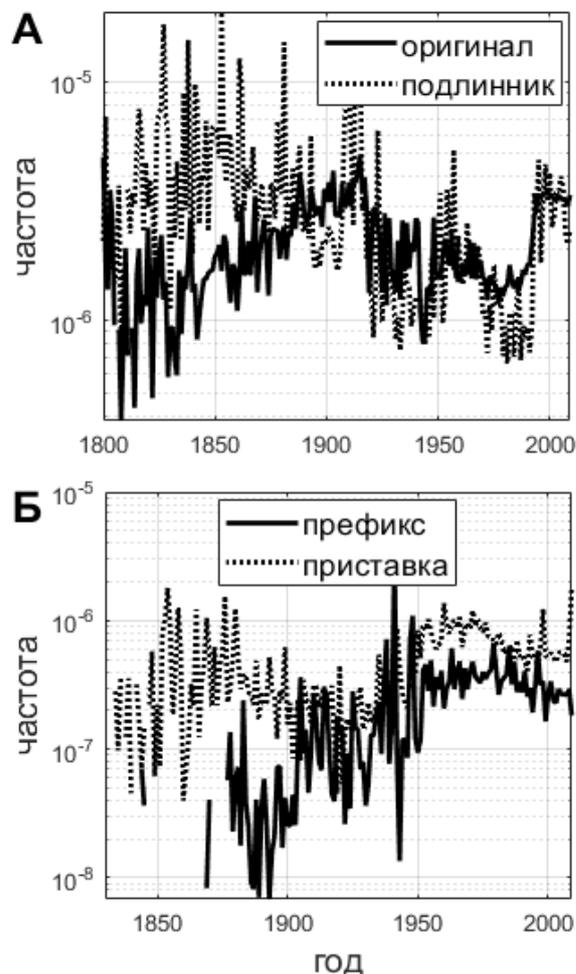


Рис. 2. Корреляция частотности в диахронии слов *оригинал* и *подлинник* (А); *префикс* и *приставка* (Б).

На основании анализа частотности и сочетаемости этих и многих других лексем мы пришли к выводу, что исконные лексемы значительно более распространены за счёт большего количества значений и широкой сочетаемости. Приведём для примера несколько графиков частотности конкурирующих лексем, демонстрирующих различные тенденции сосуществования когнат и исконных слов в языке. Наиболее распространёнными являются долговременная конкуренция (рис. 2) и резкое вытеснение славянизма иноязычным словом, завершившееся (рис. 3, А) и продолжающееся (рис. 4, Б).

Исследование комбинаторных свойств лексических единиц позволило выявить определённые закономерности:

1) этимологический сингармонизм, предполагающий тенденцию к сочетанию заимствованных и исконных слов друг с другом (*абсолют-*

ный ... + интернационализм, напр. *абсолютный максимум*, *совершенный* ... + русизм, напр. *совершенный образец*;

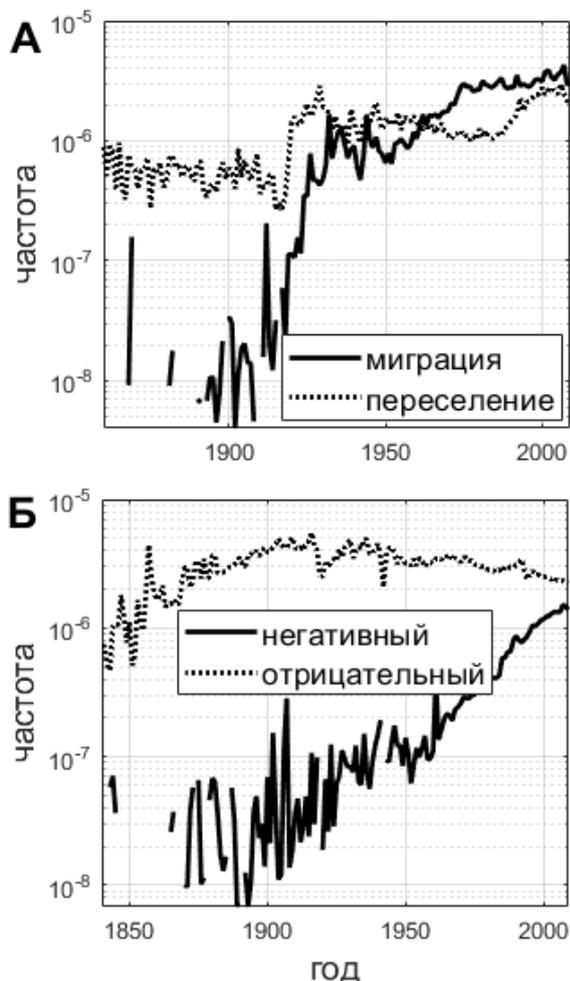


Рис. 3. Корреляция частотности в диахронии слов *миграция* и *переселение* (А); *негативный* и *отрицательный* (Б).

2) особенности ономастики – «русизм + имя собственное», топоним (*переселиться в Турцию, в Петербург*) // «интернационализм + имя нарицательное» (*мигрировать в города, в страны, на юг / север*);

3) синтаксическая сочетаемость, «наречия славянского происхождения + наречия» (совершенно / совсем+наречие – *совсем недавно, близко, рядом; совершенно ясно, справедливо, правильно*), «наречия-интернационализмы + прилагательные» (абсолютно+прилагательное – *абсолютно сухой, чёрный, белый*);

4) аксиология, эспрессивность – славянизмы часто сочетаются с оценочными словами и словами, для которых характерна выразительность, в том числе эмоциональная (*ужасное происшествие*), в то время как заимствования сочетаются со стилистически нейтральной лексикой (*неприятный инцидент*).

Как было сказано выше, самостоятельной темой исследования стало составление хронологии вхождения иноязычных слов в словарный фонд русского языка (например, слово *аргумент* впервые встречается в документе 1771 г.). Зависимость частотности интернационализма в книжной речи от даты его появления в письменных текстах не была обнаружена. Единственное наблюдение связано с тем, что большое количество слов кон. XVIII – нач. XIX вв. возглавляют рейтинг абсолютной частотности словоупотребления. Представляется необходимым применение оригинальных лексико-статистических исследовательских методик, включающих анализ продолжительности функционирования слова в книжной речи и его частотности, а также рассмотрение каждого отдельного слова в историко-этимологическом аспекте.

На основе визуального анализа 125 графиков было определено, что типичное время замещения составляет от 40 до 60 лет. Эти показатели отражают скорость вытеснения заимствованием исконных слов.

Изученные нами явления, происходящие на протяжении десятилетий и отражающиеся, в первую очередь, в изменениях частотности слов, можно назвать микроэволюцией языка. Таким образом, в данной статье была предпринята попытка разработки методологии работы со сверхобъёмными корпусами текстов и построении на их основе объяснительных и предсказательных моделей языковой микроэволюции.

Список источников

1. *Ляшевская О. Н., Шаров С. А.* Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009, 1087 с.
2. *Егорова Т. В.* Словарь иностранных слов современного русского языка. М.: «Аделант», 2014, 800 с.
3. Давайте говорить правильно! Новейшие и наиболее распространённые заимствования в современном русском языке: краткий словарь-справочник. СПб.: Филологический факультет СПбГУ; М.: Издательский центр «Академия», 2004, 217 с.
4. *Lieberman E., Michel J.-B., Jackson J., Tang T., Nowak M.A.* Quantifying the evolutionary dynamics of language // *Nature*. 2007. Vol. 449. Pp. 723–716. URL: <http://www.nature.com/nature/journal/v449/n7163/abs/nature06137.html> (дата обращения: 15.12.2020).
5. OpenCorpora. URL: <http://opencorpora.org> (дата обращения: 25.06.2024).
6. *Дуличенко А. Д.* Русангл, или: Будем ли мы говорить на интеррусском? // *Русский язык конца XX столетия* / Werner Lehfeldt. Munchen: Verl. Otto Sagner, 1994. С. 315–336.

7. Wellman B. Little Boxes, Glocalisation, and Networked Individualism. – Berlin: Springer-Verlag, 2002. Pp. 11–25.

8. Нецименко Г. П. Заимствования как проявление культурно-языковых контактов и их функционирование в языке-реципиенте // Встречи этнических культур в зеркале языка (в сопоставительном лингвокультурном аспекте). Научн. совет по истории мировой культуры. М., 2002. 479 с.

9. Щитова О. Г. Функционально-стилевая миграция заимствований как один из критериев их ассимиляции в языке-реципиенте // Вестник Томского государственного ун-та. 2007. № 294. С. 102–108.

10. Бизюков Н. В. Манипулятивный потенциал коннотации интернационализмов (на материале публицистического и рекламного дискурса) // Вестник КГПУ им. В. П. Астафьева. № 2 (32). 2015. С. 203–206.

11. Пфандль Х. О силе и бессилии пуризма. Англицизмы и интернационализмы и их возможные альтернативы (на материале русского, словенского и хорватского языков) // Вопросы языкознания. 2003. № 6. С. 108–122.

12. Kia L.S., Dan Su'Ad A. A study of education-related chinese words used in Malaysia-based computer corpus. 2019, № 37 (1). Pp. 83–107.

References

1. Lyashevskaya, O. N., Sharov, S. A. (2009). *Chastotnyi slovar' sovremennogo russkogo yazyka (na materialakh Natsional'nogo korpusa russkogo yazyka)* [Frequency Dictionary of the Modern Russian Language (based on materials from the National Corpus of the Russian Language)]. 1087 p. Moscow, Azbukovnik. (In Russian)

2. Egorova, T. V. (2014). *Slovar' inostrannykh slov sovremennogo russkogo yazyka* [Dictionary of Foreign Words in the Modern Russian Language]. 800 p. Moscow, "Adelant". (In Russian)

3. *Davaite govorit' pravil'no! Noveishie i naibolee rasprostranennye zaimstvovaniya v sovremenном russkom yazyke: kratkii slovar'-spravochnik* (2002) [Let's Speak Correctly! The Newest and Most Common Borrowings in Modern Russian: A Short Dictionary-Reference Book]. 217 p. St. Petersburg, Filologicheskii fakul'tet SPBGU; Moscow, Izdatel'skii tsentr "Akademiya". (In Russian)

4. Lieberman, E., Michel, J.–B., Jackson, J., Tang, T., Nowak, M. A. (2007). *Quantifying the evolutionary*

dynamics of language. Nature. Vol. 449, pp. 723–716. URL: <http://www.nature.com/nature/journal/v449/n7163/abs/nature06137.html> (accessed: 15.12.2020). (In English)

5. *OpenCorpora*. URL: <http://opencorpora.org> (accessed: 25.06.2024). (In English)

6. Dulichenko, A. D. (1994). *Rusangl, ili: Budem li my govorit' na interrusskom?* [Rusangl, or: Will We Speak InterRussian?]. *Russkii yazyk kontsa XX stoletiya*. Werner Lehfeldt. Pp. 315–336. Munchen, Verl. Otto Sagner. (In Russian)

7. Wellman, B. (2002). *Little Boxes, Glocalisation, and Networked Individualism*. Pp. 11–25. Berlin, Springer-Verlag. (In English)

8. Neshchimenko, G. P. (2002). *Zaimstvovaniya kak proyavlenie kul'turno-yazykovykh kontaktov i ikh funktsionirovanie v yazyke-retsipiente* [Borrowings as a Manifestation of Cultural and Linguistic Contacts and Their Functioning in the Recipient Language]. *Vstrechi etnicheskikh kul'tur v zerkale yazyka (v sopostavitel'nom lingvokul'turnom aspekte)*. Nauchn. sovet po istorii mirovoi kul'tury. 479 p. Moscow. (In Russian)

9. Shchitova, O. G. (2007). *Funktsional'no-stilevaya migratsiya zaimstvovaniya kak odin iz kriteriev ikh assilyatsii v yazyke-retsipiente* [Functional and Stylistic Migration of Borrowings as One of the Criteria for Their Assimilation in the Recipient Language]. *Vestnik Tomskogo gosudarstvennogo un-ta*. No. 294, pp. 102–108. (In Russian)

10. Bizyukov, N. V. (2015). *Manipulyativnyi potentsial konnotatsii internatsionalizmov (na materiale publitsisticheskogo i reklamnogo diskursa)* [The Manipulative Potential of the Connotation of Internationalisms (Based on Journalistic and Advertising Discourse)]. *Vestnik KGPU im. V. P. Astaf'eva*. No. 2 (32), pp. 203–206. (In Russian)

11. Pfandl', Kh. (2003). *O sile i bessilii purizma. Anglitsizmy i internatsionalizmy i ikh vozmozhnye al'ternativy (na materiale russkogo, slovenskogo i khorvatskogo yazykov)* [On the Power and Impotence of Purism. Anglicisms and Internationalisms and Their Possible Alternatives (Based on the Russian, Slovenian and Croatian Languages)]. *Voprosy yazykoznaneya*. No. 6, pp. 108–122. (In Russian)

12. Kia, L. S., Dan Su'Ad, A. (2019). *A Study of Education-Related Chinese Words Used in Malaysia-Based Computer Corpus*. No. 37 (1), pp. 83–107. (In English)

The article was submitted on 09.08.2023

Поступила в редакцию 09.08.2023

Галеев Тимур Ильдарович,
кандидат филологических наук,
доцент,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
tigaleev@kpfu.ru

Galeev Timur Ildarovich,
Ph.D. in Philology,
Associate Professor,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
tigaleev@kpfu.ru

Бочкарев Владимир Владимирович,
научный сотрудник,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
vbochkarev@mail.ru

Ван Мэняо,
магистр,
Казанский федеральный университет,
420008, Россия, Казань,
Кремлевская, 18.
mengyao@bk.ru

Bochkarev Vladimir Vladimirovich,
researcher,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
vbochkarev@mail.ru

Wang Mengyao,
Master's student,
Kazan Federal University,
18 Kremlyovskaya Str.,
Kazan, 420008, Russian Federation.
mengyao@bk.ru