

Национальный словарный фонд: новые возможности и перспективы

Мария Леонидовна Каленчук, Институт русского языка им. В. В. Виноградова РАН
(Россия, Москва), mkalenchuk@yandex.ru

DOI: 10.31857/S0131611724040037

АННОТАЦИЯ: По поручению Президента РФ начата работа над созданием Национального словарного фонда (НСФ). НСФ — цифровой ресурс, представляющий в электронном виде зафиксированные в словарях данные о функционировании норм русского языка. Необходимость разработки НСФ вызвана потребностью создания единого, полного свода научных знаний, достоверной и объективной информации о нормах русского языка в их актуальном состоянии и исторической динамике в определенную эпоху его развития, а также истории словарного состава современного русского литературного языка со времени его появления до наших дней. В рамках этого проекта будет осуществлено объединение системы словарей русского языка в единую сеть, действующую в режиме непрерывного развития и представляющую интерактивную динамическую модель лексической системы современного русского литературного языка. Предполагается открытый доступ к информации через удобный цифровой инструмент с простой навигацией, НСФ будет оснащен лингвистической разметкой и системой поиска необходимой информации в режиме онлайн. НСФ не просто предлагает пользователю возможность увидеть ту или иную словарную статью из конкретного словаря, но предоставляет экспертно разработанный инструмент извлечения из словарей различных типов языковой информации.

КЛЮЧЕВЫЕ СЛОВА: русский язык, лексикография, словари, нормы языка, электронный ресурс

для цитирования: Каленчук М. Л. Национальный словарный фонд: новые возможности и перспективы // Русская речь. 2024. № 4. С. 34–41. DOI: 10.31857/S0131611724040037.

Issues of Modern Russian Language

National Dictionary Fund: New Opportunities and Prospects

Maria L. Kalenchuk, Vinogradov Russian Language Institute of the Russian Academy of Sciences
(Russia, Moscow), mkalenchuk@yandex.ru

ABSTRACT: On behalf of the President of the Russian Federation, we have started working on the creation of the National Dictionary Fund (NDF). NDF is a digital resource that contains electronically recorded data on the functioning of the norms of the Russian language from various dictionaries. The development of NDF is due to the need to create a single, complete body of scientific knowledge, reliable and objective information about the norms of the Russian language in their current state and historical dynamics in a certain era of its development, as well as the history of the vocabulary of the modern Russian literary language from the time of its appearance to the present day. The system of Russian dictionaries will be integrated into a single network operating in a continuous development mode and representing an interactive dynamic model of the lexical system of the modern Russian language. It is supposed to have open access to information through a convenient digital tool with simple navigation. The NDF will be equipped with linguistic markup and an online search system for necessary information. NDF will offer the user not only the opportunity to see the information from a particular dictionary, but will also provide an expertly developed tool for extracting various types of language information from dictionaries.

KEYWORDS: Russian language, lexicography, dictionaries, language norms, electronic resource

FOR CITATION: Kalenchuk M. L. National Dictionary Fund: New Opportunities and Prospects. Russian Speech = Russkaya Rech'. 2024. No. 4. Pp. 34–41. DOI: 10.31857/S0131611724040037.

В течение долгого времени словари были литературой исключительно для специалистов, и только с середины XX в. ситуация изменилась — словари стали востребованы обществом, людям стало понятно, что словарь не просто книга, в которой можно узнать значение непонятого слова или получить помощь при разгадывании кроссворда, а целая энциклопедия знаний и о нормах языка, и о культуре, и об истории («все-ленная в алфавитном порядке», как говорил А. Франс).

В России, начиная с конца XVIII в., была заложена традиция создания лингвистических словарей как особого жанра научной и научно-популярной литературы. В наше время параллельно с бумажными словарями широко используются и электронные. Компьютерная лингвистика — одно из наиболее динамично развивающихся направлений современной науки о языке; в последние десятилетия было создано много разнообразных электронных ресурсов, посвященных русскому языку. При этом среди таких цифровых ресурсов в основном представлены два типа — корпуса текстов и библиотеки словарей. Они различаются не только источниками предоставляемой информации, но и заложенным в их основу принципом: корпуса — принципиально ненормативны; словари литературного языка фиксируют нормы языка, то есть правила, по которым строится речь образованных людей.

Несколько лет назад в соответствии с поручением Президента РФ была начата работа над созданием государственной информационной системы «Национальный словарный фонд» (далее НСФ). НСФ — интернет-ресурс, представляющий в электронном виде закреплённые в словарях разных типов и времени создания данные о функционировании норм русского языка в их актуальном виде и исторической динамике.

Необходимость разработки НСФ вызвана потребностью создания единого, полного свода научных знаний, достоверной и объективной информации о нормах современного русского литературного языка в их актуальном состоянии и исторической динамике в определенную эпоху его развития, а также истории словарного состава современного русского литературного языка со времени его появления до наших дней. В рамках

этого проекта предполагается объединение системы словарей русского языка в единую сеть, действующую в режиме непрерывного развития и представляющую интерактивную динамическую модель лексической системы современного русского литературного языка. Предполагается открытый доступ к информации через удобный цифровой инструмент с простой навигацией, НСФ будет оснащен лингвистической разметкой и системой поиска необходимой информации в режиме онлайн.

Лексикографическая деятельность, ориентированная на печатное издание словарей, не дает возможности представить актуальное состояние языка, неизбежно отстает от реального состояния норм русского языка и культурно-языковых потребностей общества. Кроме того, тексты многих словарей XVIII–XIX вв. труднодоступны для пользователей.

В этой связи ключевую роль в представлении лексического материала, позволяющего фиксировать реальное функционирование языка и отражать происходящие в нем процессы и изменения, играют информационно-коммуникационные технологии.

В сети Интернет имеется достаточное количество сайтов, предоставляющих доступ к текстам различных словарей. Но, во-первых, сами эти тексты в большинстве случаев «грязные», неverifiedированные, полные опечаток и невычитанных после сканирования ошибок. А во-вторых, подобные ресурсы действуют просто как собрание словарей, как своеобразная библиотека, на электронных полках которой стоят лексикографические труды, подобранные в случайном порядке. НСФ же не просто предлагает пользователю возможность увидеть ту или иную словарную статью из конкретного словаря, но предоставляет экспертно разработанный инструмент извлечения из словарей различных типов языковой информации. Подобного ресурса в нашей стране до сих пор не было, хотя аналоги функционируют в некоторых странах, например в Великобритании или Германии.

Мотивация обращения к НСФ может быть весьма различной — от необходимости решить практическую задачу использования языковой единицы в устной или письменной речи до желания расширить свой кругозор, от лингвистической любознательности до потребности поиска материалов в учебных или научных целях.

Состав словарей НСФ

Для включения в НСФ авторитетными экспертами было отобрано 33 словаря русского языка, которые можно разделить на две большие группы: описывающие нормы современного русского литературного языка и представляющие уже ушедшие нормы в их исторической

динамике. Кроме того, используются материалы Словаря русских народных говоров.

При отборе современных словарей русского языка необходимо было подобрать лексикографические источники таким образом, чтобы были представлены разные языковые аспекты функционирования слова. В НСФ входят словари следующих типов: толковые, орфографические, орфоэпические, словари ударений, грамматические, активные, словари синонимов, универсальные, словарь морфем, фразеологические, этимологические. Этим набором вовсе не исчерпывается разнообразие видов словарей русского языка, но выбранные источники позволяют извлечь из них различную информацию, достаточную для описания языковых норм любого типа. Например, в НСФ отсутствует словарь антонимов, но антонимы к конкретным словам приводятся в некоторых толковых словарях, в активном и универсальном словарях русского языка. Информация об устойчивых сочетаниях слов в русском языке не только будет приводиться из фразеологического словаря, но с помощью лингвистической разметки будет отбираться из толковых, орфоэпических и др. словарей.

Во многих случаях в НСФ включается по несколько словарей одного и того же типа, например девять толковых или три орфоэпических. Это делается для того, чтобы можно было продемонстрировать и разные подходы авторов при кодификации норм, и изменение норм во времени. Например, в [Аванесов (ред.) 1983] кодифицируются как равноправные два варианта ударения — *бижутэ́рия* и *бижуте́рия*, *ло́сось* и *лосо́сь*, а в [Касаткин (ред.) 2012], созданном на несколько десятилетий позже, литературным признается только произношение *бижутэ́рия*, *лосо́сь*, а *бижуте́рия* и *ло́сось* маркируются пометой *неправильное*.

Словари, представляющие становление норм в истории русского языка, делятся на несколько групп:

- исторические по времени создания, XVIII — первая половина XX в. (например, «Словарь Академии Российской, по азбучному порядку расположенный» (1806–1822 гг.));
- созданные в наше время, середина XX — начало XXI в. (например, «Словарь языка А. С. Пушкина» (1956–1961));
- этимологические.

Некоторые принципы и правила, положенные в основу НСФ

Пользование НСФ будет свободным и бесплатным. Регистрация и авторизация — по желанию, например, при необходимости создать личный кабинет пользователя и экспортировать в него извлеченные словарные материалы.

Словарная информация о нормах русского языка выдается пользователю по запросу, оформленному в графическом виде и представляющему собой отдельную лексическую единицу — слово или устойчивое сочетание слов.

Все материалы включенных в НСФ словарей предъявляются пользователям в аутентичном виде, в авторских версиях. Создатели ресурса ни при каких обстоятельствах не должны содержательно редактировать лексикографические источники.

Модули НСФ

Языковая информация на платформе НСФ структурируется в различные модули. Пользователь выбирает тот модуль, информация в котором позволяет решить поставленные задачи, и каждый может «нырнуть» на нужную ему «глубину» языковой информации.

Модуль 1. Основной. Как показывает практика, чаще всего человек хочет получить информацию о написании, произношении (включая место ударения), толковании значения и грамматических формах заданного в поиске слова. Это значит, что при обращении к этому модулю пользователю после введения запроса будут предъявлены словарные статьи из орфографического, орфоэпического, толкового и грамматического словарей. Обращаясь к этому модулю, пользователь в большинстве случаев хочет решить конкретную задачу. Обычно это происходит тогда, когда человек ощущает определенное сомнение, чувствует неуверенность в своих знаниях или находится в положении выбора. Как правильно написать — *риелтор* или *риэлтор*, *вебсайт* или *веб-сайт*? Как верно произнести — *брю[н'э]т* или *брю[нэ]т*, *уценённый* или *уцёненный*? Нет *носок* или *носков*, *поехай* или *поезжай* — как образовать грамматическую форму? Проверить себя, правильно ли вы понимаете значение слова или выражения — что такое *каршеринг* или *коворкинг*, что означает выражение *гамбургский счет*? Этот модуль должен предлагать базовую информацию о слове в компактном и доступном виде. Для каждого типа информации, предъявляемой в этом модуле, выбран один — самый авторитетный, актуальный и полный словарь. Но при желании пользователь может нажать кнопку «Посмотреть в других словарях» и расширить список словарей, содержащих нужную информацию.

Модуль 2. Расширенный. Обратившись к этому модулю, пользователь сможет получить более системные словарные данные о запрошенном слове, в которых суммируются базовая информация из предыдущего модуля и сведения о синонимах, антонимах к этому слову, его происхождении, фразеологизмах, в которые он входит, специфике его поведения

в разговорной речи, морфемной структуре и др. Информация в этом модуле позволяет получить целостный «портрет» слова в его актуальном для современных норм виде.

Модуль 3. Историко-этимологический. Предоставляя доступ к словарным статьям исторических словарей разных типов и времени создания, НСФ дает возможность посмотреть информацию о слове во всех исторических словарях с древнерусского периода до наших дней, или в определенном историческом словаре, или в словарях выделенной подгруппы (хронологическая — от дописменного периода до современного или по хронологии создания словарей). Кроме того, будет заложена возможность поиска информации по определенному временному отрезку.

Модуль 4. Посмотреть в выбранном словаре. При запросе слова показываются только те словари, в которых это слово имеется, а далее есть выбор: посмотреть в конкретном словаре; посмотреть во всех словарях; посмотреть в словарях определенного типа (толковые, орфографические, орфоэпические, этимологические и др.) — при выборе типа разворачиваются все словари этой группы.

Модуль 5. Сопоставительный. Это самый сложный для исполнителей проекта модуль. Он позволяет сравнить словарные данные по выбранным аспектам, например: место ударения; написание; структура значения; сочетаемость; фразеология и пр. в разных словарях. Например, пользователя интересуют устойчивые сочетания слов с опорным словом *ухо*. Выбрав сопоставительный модуль и указав аспект сравнения — фразеология, он может получить информацию не только о списках подобных фразеологизмов из словарей разного типа и времени создания (*тугой на ухо; говорить, шептать* и др. *на ухо; слон (медведь) на ухо наступил; за уши не оттащишь; быть по уши в чем-л.; влюбиться по уши; вешать лапшу на уши; поставить всех на уши* и мн. др.) и узнать их значение, но и увидеть динамику их изменений в разные языковые эпохи.

Обращение к НСФ позволит пользователям решать задачи совершенно разного уровня. Среди них будут чисто практические: например, какова норма правописания — *апелляция* или *апелляция*, *Интернет* или *интернет*; как правильно произносить *логин* или *логín*, *вёган* или *вегán*. А преподаватель, готовясь к занятиям, сможет проверить, как делятся на морфемы, скажем, слова *земляника* или *последователь*, а также подобрать, например, все фразеологизмы, со словом *рука* и т. д. Но новый цифровой ресурс даст возможность решать и гораздо более сложные и часто многоуровневые задачи, демонстрирующие устройство и функционирование языковой системы. Как увлекательно, сравнивая материалы разных словарей, узнать, что слово *негодяй* раньше означало *человек, не годный к воинской службе*; что слово *артиллерия* в XVIII в.

писалось семью разными способами, и проследить этапы становления современной нормы его написания; увидеть, как в разные эпохи конкурировали различные значения слов *базар* и *рынок* и мн. др. Каждый пользователь Национального словарного фонда сможет найти в новом электронном ресурсе и полезное, и интересное.

Источники

Словарь Академии Российской, по азбучному порядку расположенный. СПб., 1806–1822.

Литература

Аванесов Р. И. (ред.). Орфоэпический словарь русского языка: произношение, ударение, грамматические формы / С. Н. Борунова, В. Л. Воронцова, Н. А. Еськова. М.: Русский язык, 1983. 703 с.

Виноградов В. В. (отв. ред.). Словарь языка Пушкина. Т. I–IV. М.: Наука, 1956–1961.

Касаткин Л. Л. (ред.). Большой орфоэпический словарь русского языка. Литературное произношение и ударение начала XXI века: норма и ее варианты / М. Л. Каленчук, Л. Л. Касаткин, Р. Ф. Касаткина. М.: АСТ-ПРЕСС, 2012. 1001 с.

References

Avanesov R. I. (ed.) *Orfoepicheskii slovar' russkogo yazyka: proiznoshenie, udarenie, grammaticheskie formy* [Orthoepical dictionary of the Russian language: pronunciation, stress, grammatical forms]. Moscow, Russkii Yazyk Publ., 1983. 703 p.

Vinogradov V. V. (resp. ed.) *Slovar' yazyka Pushkina* [A. S. Pushkin Language Dictionary]. Vol. I–IV. Moscow, Nauka Publ., 1956–1961.

Kasatkin L. L. (ed.) *Bol'shoi orfoepicheskii slovar' russkogo yazyka. Literaturnoe proiznoshenie i udarenie nachala XXI veka: norma i ee varianty* [A large orthoepical dictionary of the Russian language. Literary pronunciation and stress of the beginning of the XXI century: the norm and its variants]. Moscow, AST-PRESS Publ., 2012. 1001 p.