

УДК 004.4'412

АВТОМАТИЧЕСКАЯ КОРРЕКЦИЯ ОРФОГРАФИЧЕСКИХ ОШИБОК.

Гниловская Л.П.

Факультет вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова, Москва, Россия

Гниловская Н.Ф.

Крымская Республиканская универсальная научная библиотека им. И. Франко, Симферополь, Украина

Предоставление читателям библиотек возможности использования каталогов и полнотекстовых документов на электронных носителях облегчает поиск необходимой информации, экономит время, позволяет произвести наиболее полный отбор и изучение материалов. При создании баз данных производится ввод текстовой информации. Орфографические ошибки, появляющиеся в таких текстах, не редкость. Современные текстовые редакторы предлагают варианты для их исправления, что требует вмешательства пользователя. Автор статьи предлагает автоматическую коррекцию орфографических ошибок, основанную на морфемном анализе.

Granting to readers of libraries of an opportunity of use of catalogues and text-through documents on electronic carriers facilitates search of the necessary information, saves time, allows to make the fullest selection and studying of materials. At creation of databases input of the text information is made. The spelling errors appearing in such texts, not a rarity. Modern text editors offer variants for their correction that demands intervention of the user. The author of clause offers the automatic correction of spelling errors based on the morphemic analysis.

Проникновение электронной информации во все слои человеческой деятельности заставляет библиотечные учреждения активно осваивать современные информационные технологии, внедрять автоматизированные системы, создавать электронные библиотеки, развивать Интернет-услуги. Цивилизованный библиотечный мир однозначно воспринимает эти изменения как одну из характерных тенденций общества будущего, как систему, облегчающую доступ пользователей к информационному ресурсу.

Нарастание потоков электронной информации во всех сферах человеческой деятельности резко повышает статус библиотеки как информационного и общественного института. Чтобы полно и оперативно удовлетворять информационные запросы современного пользователя, библиотека должна быть в центре внедрения современных технологий и использования электронной информации. Реальностью сегодняшнего дня стали электронные издания, число которых постоянно увеличивается. Библиотеки, не имеющие в фонде тех или иных электронных изданий и предоставляющие к ним доступ через Интернет, уже с полным на это основанием включают их библиографические описания в свои каталоги и предоставляют их пользователям.

Дальнейшая информатизация общества и расширение сферы применения информационных технологий в библиотечной практике в современных условиях порождают новые виды и формы обслуживания пользователей библиотек. Главная особенность — самостоятельность пользователя в работе с компьютерными ресурсами библио-



теки. Это требует обеспечения определенного уровня обслуживания и заставляет библиотекарей более тщательно относится к системе автоматизации, к возможностям информационных технологий, внедряемых в библиотечную практику.

Научная библиотека Московского университета им. М. В. Ломоносова - одна из старейших библиотек России: основана в 1756 году. Она является подразделением МГУ, самоуправляемого государственного вуза Российской федерации. Ежегодно здесь обслуживается более 60 тысяч постоянных читателей, из которых 52 тысячи - сотрудники и учащиеся университета. Количество посещений составляет 1,8 млн. в год. Книговыдача - 4 млн. экземпляров. За год читателям выдается более 80 тыс. справок. К их услугам 16 абонементов и 60 читальных залов на 3300 мест. В штате библиотеки более 700 сотрудников, работающих в 38 отделах: 19 отделах обслуживания, 17 функциональных, 2 научных - редких книг и рукописей и компьютеризации информационнобиблиотечных процессов. Универсальный фонд научной и учебной литературы насчитывает 8,5 млн. единиц хранения, в том числе 2,5 млн. на иностранных языках. В отделе редких книг и рукописей более 200 тыс. единиц хранения: средневековые, греческие, латинские, французские, немецкие, восточные и славянские рукописи, русские рукописи XVIII-XIX вв., архивные материалы XVIII-XIX вв., западноевропейские инкунабулы и палеотипы, славянские и русские старопечатные книги; редкие русские издания XVIII-XX вв. Библиотека обеспечивает литературой работу 21 учебного подразделения МГУ: механико-математический факультет, факультет вычислительной математики и кибернетики, физический факультет, химический факультет, биологический факультет, факультет фундаментальной медицины, факультет почвоведения, геологический факультет, географический факультет, высший колледж наук о материалах, исторический факультет, филологический факультет, факультет иностранных языков, философский факультет, социологический факультет, экономический факультет, юридический факультет, факультет журналистики, факультет психологии, институт стран Азии и Африки, институт государственного управления и социальных исследований. Одним главных направлений работы библиотеки является пополнение ее фондов. Комплектование отечественной литературой ведется на основе государственного бесплатного обязательного экземпляра, который библиотека получает с 1920 года, прямых контактов с издательствами, книготорговыми организациями. Фонд иностранной литературы в значительной степени пополняется за счет изданий, получаемых по международному книгообмену, (библиотека имеет около 1000 партнеров по обмену в 63 странах), благодаря дарам различных организаций и частных лиц, а также в результате валютной подписки на периодические издания.

В 1987 году началась компьютеризация библиотеки МГУ. Был организован отдел компьютеризации информационно-библиотечных процессов, создано современное программное обеспечение, позволяющее комплексно решить вопрос перевода библиотечного дела на компьютерную технологию. В эксплуатации находится система "Биб-



лиотека 4.0", получившая широкое распространение в библиотеках России. Данная система использует формат US MARC библиотеки Конгресса США. Она может применяться как на отдельном персональном компьютере, так и в локальной сети, позволяет компьютеризировать все библиотечные процессы, начиная с комплектования литературы и заканчивая информационным обслуживанием читателей.

С 1990 года все новые поступления книг заносятся в электронный каталог. В настоящее время объем электронного каталога составляет около 280 тысяч названий. Кроме основного каталога книг ведутся отдельные каталоги по различным областям знаний на основе росписи журнальных статей. Имеется полнотекстовая база данных учебников. В помещениях научной библиотеки МГУ читатели имеют доступ к полным текстам статей журналов Научной Электронной Библиотеки **elibrary.ru**. Предоставление читателям возможности пользоваться электронными каталогами, различными по тематике базами данных и полными текстами учебников и отраслевых журналов значительно облегчает тематический поиск необходимой информации, экономит время, позволяет произвести наиболее полный отбор и изучение документов.

При создании различных по назначению баз данных производится ввод текстовой информации, осуществляемый двумя способами – набором вручную или сканированием. Как правило, при этом возникают одни и те же проблемы. Если набирать текст вручную, то возможны опечатки и появляется необходимость вычитывать весь текст. При этом непроизводительно тратится уйма времени. Другой способ – сканирование. Время, которое уходит на ввод в компьютер информации, значительно сокращается. Однако и в этом случае возможны орфографические ошибки.

Современные текстовые редакторы (например, WinWord) при проверке текстов не корректируют ошибки, а предлагают варианты для их исправления. Это требует вмешательства пользователя, что не всегда удобно. Автоматическая коррекция орфографических ошибок, основанная на морфемном анализе, может быть более эффективным средством минимизации опечаток и их исправлений при создании текстовых файлов. Данная тема и стала целью исследования.

Морфемой называется минимальная значимая часть слова. Морфемный анализ — это операция, цель которой состоит в определении инвентаря морфем того или иного языка. Для достижения этой цели необходимо решить задачи, состоящие в выявлении и систематизации морфем в языке, в установлении законов из функционирования и соединения. Известно, что в русском языке около 5 000 морфем позволяют построить примерно 52 000 слов, то есть, разлагая слово на морфемы, можно достичь более чем десятикратного сокращения словаря.

Слова русского языка с точки зрения морфологической структуры делятся на слова, имеющие формы словоизменения и не имеющие словоизменения.

Слова первой группы распадаются на две части: основу и окончание, или флексию, слова второй группы представляют собой чистую основу.



Основа – это часть слова, которая выражает его лексическое значение. Основа выделяется путём вычета окончания.

Окончание, или флексия, — это изменяемая часть слова, которая указывает на отношение данного слова к другим, то есть является средством выражения синтаксических свойств слова в предложении. Окончание может быть нулевым.

Основа слова распадается на отдельные значимые части: приставка, корень, суффикс.

Корень слова – общая часть всех родственных слов.

Слова, происходящие от одного корня, составляют словообразовательное гнездо. Корень может присоединять к себе различные аффиксы.

Аффикс – общее название всех значимых частей слова, за исключением корня. Аффиксы подразделяются на приставки или префиксы – части слова, стоящие перед корнем, суффиксы – части слова, стоящие между корнем и окончанием, и окончания.

Например, слово **поведение** имеет корень – вед – (ос*вед*омляться, с*вед*едение, введение и так далее), приставку по–, суффикс –ениј–, и окончание –е. По–, –ениј– и –е являются аффиксами.

Каждая значимая часть слова – приставка, корень, суффикс или окончание – называются морфемой. Слово **поведение** состоит из четырех морфем: по–вед–ениј–е.

Аффикс –ся, –сь имеет особое значение – частица, потому что часто помещается после других частей слова, в том числе и окончания.

В большинстве сложных слов выделяется ещё одна часть – словосоединительная гласная (интерфике): пар–о–воз, сам–о–лёт, труб–о–провод, нитк–о–вдеватель.

Выделение морфем из состава слова и определение их значения производится на основе его сопоставления с другими словами и с другими формами данного слова.

Учет морфемной структуры слов позволяет компактно представлять совокупность словоформ, группируя их в словообразовательные гнезда. Такое представление реализуется в виде словаря морфем, содержащего три части – корневую, префиксальную и суффиксальную.

В корневой части морфемного словаря корни расположены в лексикографическом порядке. В словарной статье под каждым заглавным корнем приводятся однокоренные слова, расчлененные на морфемы.

Аффиксальная часть словаря состоит из префиксальной и суффиксальной частей.

В префиксальной части словаря под каждым заглавным префиксом дается в алфавитном порядке перечень всех аффиксальных окружений корня, в которых встречается данный префикс; рядом перечисляются все корни, употребляющиеся в соответствующем окружении. Слова, начинающиеся прямо с корня, приводятся вначале.

Префиксы даются в словаре по следующей схеме:



- сначала идут модели, начинающиеся на заглавный префикс, который является единственным в слове;
- потом приводятся модели, начинающиеся с заглавного префикса, но сопровождаемые одним или двумя другими префиксами, стоящими по алфавиту;
- далее идут модели, в которых заглавный префикс расположен не на первом месте в словаре; такие модели располагаются строго по алфавиту префиксов.

В суффиксальной части словаря под заглавным суффиксом помещены модели, в которых данный суффикс встречается сначала в беспрефиксальных моделях на 1-ом месте, затем идут все суффиксы на 2-ом, 3-ем и так далее, вплоть до 6-ого места в словаре (по отношению к корню).

Выделяют четыре типа случайных ошибок:

- 1. пропуск символа (асемблер);
- 2. замена одного символа любым другим (конпьютер);
- 3. удвоение символа (клаввиатура);
- 4. перестановка двух соседних символов (аглоритм).

Алгоритмы исправления этих ошибок, основанные на хранении в памяти компьютера всех возможных словоформ, имеют существенные недостатки.

Первый недостаток состоит в том, что каждая словоформа из списка должна сопоставляться со словами словаря. Если слово состоит из L букв, то пропущенная буква могла стоять на одном из L+1 мест. Поскольку неизвестно, какая буква была пропущена, то на каждом месте может стоять почти любая буква алфавита. То есть ошибка первого типа порождает (28+32L) словоформ. Ошибка второго типа порождает (27+31(L-1)) словоформ. Для установления ошибки третьего типа нужно поочередно вычеркивать каждую из L букв, этот процесс добавит ещё L словоформ. Если возникает ошибка четвёртого типа, то при перестановке всех пар соседних букв появляется L-1 словоформ. Всего список кандидатов содержит (65L + 33) словоформ, для каждой из которых нужно осуществлять сопоставление со словарём. Знание длины слов позволяет не делать сравнение со всеми словами словаря, а выбирать из словаря только слова той же длины. Кроме того, лексикографическое упорядочение слов словаря позволяет использовать индекс для поиска всех словоформ, кроме тех, которые призваны исправить ошибку в первой букве. Учитывая, что средняя длина слова в русском языке составляет 6 букв, можно ожидать, что этот недостаток не будет приводить к тупиковым ситуациям.

Второй недостаток таких алгоритмов в том, что список словоформ-кандидатов будет содержать более одного слова из словаря. В этом случае ошибка должна исправляться оператором.

Разложение словоформы на морфемы, помимо того, что существенно сокращает объем словаря, позволяет сокращать объем комбинаторного перебора при генерации кандидатов.



Для алгоритма коррекции ошибок, опирающегося на результаты морфемного анализа, достаточно взять только префиксальную часть словаря, разбив её на три уровня:

- 1. словарь первого уровня содержит наиболее часто используемые слова языка;
- 2. словарь второго уровня является подсловарем словаря третьего уровня и создается в процессе обработки конкретного документа либо пакета документов из одной и той же предметной области;
- 3. словарь третьего уровня должен содержать все слова языка, за исключением тех, что вошли в словарь первого уровня.

Этот словарь необходимо дополнить следующим образом: каждой основе сопоставить группу окончаний, совместимых с данной основой. Известно, что имена существительные в русском языке имеют три типа склонения. В пределах одного типа каждый падеж чаще всего имеет одно окончание, общее для всех слов, входящих в этот тип. Однако не редки и колебания в использовании определенных падежных окончаний. Например, в первом и втором склонениях различается склонение слов на твердую и мягкую основы. Кроме того, в первом склонении различается склонение имен существительных мужского рода и имен существительных среднего рода. Десять имен существительных на —мя, такие, как знамя, имя, семя и другие, находятся вообще вне этих трёх типов склонений так же, как и слово путь.

Аналогичная ситуация наблюдается в других изменяемых частях речи. Более точную информацию о формах изменяемых частей речи можно получить, разбив их на флективные классы слова, имеющие одинаковые окончания во всех формах изменения, объединяются в один класс. Каждому классу сопоставляется номер и список окончаний, добавляемых к основе при склонении, спряжении, изменении рода и числа.

Пример. Модель на $-\sqrt{-a-n-b-H-biŭ}^1$, $-s^2$; $-\kappa o s^2 - 54$, $-v^1 - 66$ представляет два слова, одно из них – существительное **наковальня** – получается, если вместо символа радикала подставить корень $-\kappa o s - u$, поскольку корень ков снабжен верхним индексом 2, то берем окончание с верхним индексом 2, то есть -s. Цифра 54 после – $\kappa o s - u$ означает, что получающееся после подстановки этого корня и окончания слово относится к пятьдесят четвертому флективному классу. Подставляя вместо символа радикала корень -u - u взяв в качестве окончания -u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный – u - u взяв в качестве окончания u - u получим начальный u - u взяв в качестве u - u взяв в качестве u - u получим начальный u - u взяв в качестве u - u

Другое изменение, вносимое в префиксальную часть словаря морфем, состоит в том, что несколько последовательно стоящих префиксов рассматривается как один составной префикс (этот термин – составной префикс вводится для удобства алгоритмизации и изложения, но ничего общего с лингвистикой не имеет). В связи с этим заглавный префикс всегда стоит на первом месте. Таким образом, представляемый в компью-



тере префиксальный словарь морфем имеет порядок словарных статей, определяемый лексикографическим упорядочением составных префиксов, начинающимся с нулевого префикса.

Помимо упрощения алгоритма поиска слова в словаре это приводит к сокращению словаря: ведь в словаре морфем словарная статья встречается столько раз, из скольких префиксов состоит составной префикс. Например, слово **перевооружение** встречается в префиксальной части словаря морфем 3 раза: сначала в форме пере–ВО– о–√–ени –е, –руж–, дальше – в форме ПЕРЕ–во–о–√–ени–е, –руж– и ещё в форме пере–во–О–√–ени–е, –руж–.

Определение. Будем называть отображением ошибки категории err_i (err_i = 1, 2, 3, 4) данной словоформы множество словоформ, порождаемых этой словоформой в результате всех возможных ошибок категории err_i.

Определение. Будем называть полным отображением одиночной ошибки данной словоформы множество словоформ, порождаемых этой словоформой в результате всех возможных ошибок четырех категорий.

Таким образом, если через SL_1 (словоформа) обозначить множество словоформ, порождаемых данной словоформой в результате ошибки категории 1 (отображение ошибки категории 1), через SL_2 — отображение ошибки категории 2, через SL_3 — отображение ошибки категории 3 и через SL_4 — отображение ошибки категории 4, то полное отображение одиночной ошибки SL (словоформа) есть объединение всех четырёх множеств:

SL (словоформа) = SL_1 (словоформа) \cup SL_2 (словоформа) \cup $\cup SL_3$ (словоформа) \cup SL_4 (словоформа).

Пример. Найдем отображение ошибки категории 4 (перестановка двух соседних символов) словоформы село:

 SL_1 (село) = есло, слео, сеол.

Определение. Назовём обратным полным отображением одиночной ошибки данной словоформы множество словоформ, порождающих данную словоформу в результате одиночной ошибки любой из четырех категорий.

Так как в дальнейшем будут рассматриваться только одиночные ошибки, то для краткости будем говорить «обратное полное отображение», опуская «одиночной ошибки».

В основе алгоритма обнаружения и коррекции ошибок, опирающегося на префиксальную часть словаря морфем, лежат два положения.

Во-первых, поскольку в слове по условию возможна только одна ошибка, то она может находиться или в префиксе, или в корне, или в суффиксе, или в окончании. Это сокращает объем необходимых комбинаторных вычислений при поиске правильного слова.



Во-вторых, знания о сочетаемости каждого префикса с определенными корнями слов и основы – с определенными окончаниями довольно часто позволяют делать заключения о правильности того или иного варианта коррекции ошибки.

Алгоритм обнаружения и коррекции одиночных ошибок включает этапы:

Этап 1. На первом этапе предполагается, что в рассматриваемом слове нет префикса, и поэтому в нем пытаемся выделить одну из основ с нулевым префиксом. Если это не удается, то либо в слове есть ошибка, либо в слове есть префикс (этап 2).

Сначала проверяется гипотеза об ошибке в основе. Для этого:

- А) производятся все возможные разбиения анализируемого слова на основу и окончание;
- Б) для каждой из полученных основ находится обратное полное преобразование, то есть генерируются все основы, могущие породить данную в результате одиночной ошибки.

Если сгенерированная основа есть в словаре, то она проверяется на совместимость (принадлежность к флективному классу) с окончанием анализируемого слова. Если сгенерированная основа несовместима с окончанием анализируемой основы, то она из рассмотрения исключается и генерируется следующая основа; если же сгенерированная основа совместима с окончанием, то есть она находится в одном флективном классе с анализируемой основой, то она заносится в список кандидатов и совершается переход к генерации следующей основы. По окончании процесса генерации проверяется число основ в списке кандидатов. Если в этом списке окажется больше одной основы, то после присоединения окончаний (совпадающих с окончанием анализируемого слова), все они выдаются на экран вместе с анализируемым словом и сообщением «Варианты исправления ошибки». Если в списке кандидатов – только одна основа, то после присоединения к ней окончания замещаем этой словоформой анализируемую. Ошибка исправлена. Если же в списке кандидатов не оказалось ни одной основы, то это значит, что окончание было выбрано неверно, нужно испытать другое окончание. Если в результате перебора всех допустимых пар основа-окончание не удалось найти ни одного кандидата, то это значит, что в слове есть префикс, переходим к этапу 2.

Вернёмся теперь к началу этапа и рассмотрим случай, когда в слове удалось выделить основу с нулевым префиксом. В этом случае нужно проверить соответствие окончания анализируемого слова флективному классу основы. Проверка на совместимость основы с окончанием может дать следующие результаты: либо подтвердится, что основа определена правильно, либо обнаружится ошибка в окончании и, возможно, будет исправлена. Сложность исправления ошибки в окончании состоит в том, что в результате ошибки в окончании возможна замена одного окончания на другое из той же самой группы окончаний, определяющие флективный класс.

Этап 2. На втором этапе проверяется наличие в словоформе префикса. Сначала предполагается, что в префиксе нет ошибки. Начиная с префикса наибольшей длины с



той же начальной буквой, что и слово, пытаемся выделить префикс в словоформе. Если это удалось, то среди основ, совместимых с данным префиксом, пытаемся найти ту, которая входит в анализируемую словоформу. Если это удается, то осуществляется проверка окончания на принадлежность к группе окончаний, определяющих флективный класс найденной основы. Положительный результат проверки свидетельствует о том, что в слове нет ошибки, поэтому осуществляется переход к анализу следующего слова; отрицательный результат проверки свидетельствует об ошибке в окончании. Если перебор основ, совместимых с данным префиксом не дал результата, то это значит, что либо в основе есть ошибка, либо префикс выбран ошибочно. Сначала обращаемся к модулю проверки гипотезы об ошибке в основе. Если не удалось найти ошибку в основе, то возвращаемся из модуля для выбора следующего префикса. Если перебор всех префиксов с той же начальной буквой, что и слово, не дает результата, то осуществляется переход к этапу проверки гипотезы об ошибке в префиксе.

Этап 3. Проверку гипотезы об ошибке в префиксе начинаем с того, что пытаемся отделить префикс от основы. Для этого последовательно убираем начальные буквы слова, и для каждого начального отрезка из 1, 2, ... к букв находим обратное полное отображение (ОПО). Каждый элемент этого отображения рассматривается как возможный правильный префикс. Эта последовательность символов сравнивается со всеми префиксами языка. В случае совпадения начинается перебор основ, совместимых с этим префиксом, и осуществляется проверка вхождения основ с этим префиксом, и осуществляется проверка вхождения основ в анализируемое слово. Если найдется основа, входящая в анализируемое слово, то проверяется принадлежность окончания флективному классу этой основы. При положительном результате проверки найденный префикс присоединяется вместо удалённых букв, анализируемое слово заменяется полученным, осуществляется переход к анализу следующего слова. Если окончание не принадлежит флективному классу, то считается, что в окончании ошибка, а это значит, что в слове более одной ошибки.

Вернемся к случаю, когда элемент ОПО не совпал ни с одним из префиксов языка. В этом случае осуществляется переход к генерации следующего элемента ОПО. Если генерация всех элементов ОПО для начального отрезка словоформы длиной s не дала элемента, совпадающего с одним из префиксов, то длина отрезка, рассматриваемого как возможный префикс (но с ошибкой), увеличивается на единицу и процесс повторяется.

Эвристически предполагается, что в любом слове, помимо префикса, есть ещё не менее трех букв. Поэтому, если в остатке словоформы после удаления s букв окажется менее трех букв, то либо слова нет в словаре, либо в нем более одной ошибки.



Вывод: алгоритм коррекции орфографических ошибок, основанный на морфемном анализе, намного лучше справляется с данной задачей, нежели алгоритмы, основанные на хранении в памяти компьютера всех возможных словоформ. Его с успехом можно использовать при создании электронных каталогов и полнотекстовых баз данных, в подготовке словарей, рубрикаторов, справочников.

Литература и источники:

- 1. Маршак Б. И. Современные проблемы разработки и внедрения автоматизированных библиотечно-информационных систем: системный подход и оценка программного окружения. // Научные и технические библиотеки. 2003. №2. С. 5-11.
- 2. Шварце Х., Хольцгрефе Г. Использование компьютеров в регулировании и управлении. М., 1990.
- 3. Шрайберг Я. Л. Библиотеки, компьютерные технологии и информационное общество: год прошедший и год грядущий. // Научные и технические библиотеки. 2003. №1. С. 28-58.
- 4. Файн В. С., Рубанов Л. И. Машинное понимание текстов с ошибками. М., 1991
- 5. www.lib.msu.su

УДК 811.161.1 + 81'373.7 + 81'374

ХУДОЖЕСТВЕННОЕ ЗНАЧЕНИЕ В СИСТЕМЕ ИНДИВИДУЛЬНО-АВТОРСКОГО СЛОВАРЯ

Петрова Л.А.

Институт русского языка им. А.С. Пушкина, Москва, Россия

В статье рассматриваются особенности формирования художественного значения слова, которое является одной из единиц описания индивидуально-авторского словаря. Выявляются семантические процессы, соотносящие художественное значение с лексическим значением. Обосновывается возможность создания индивидуально-авторского словаря в электронной версии.

Ключевые слова: лексическое значение, художественное значение, семантико-стилистическая система писателя, электронные словари

The article deals with the features of formation of art meaning of a word. This art meaning is one of units of the description of the individual - author's dictionary. The semantic processes, correlating art meaning with lexical meaning, come to light.

Keywords: lexical meaning, art meaning, semantical-stylistic system of the writer, the electronic dictionaries