

Ивлиева Ирина Валентиновна

канд. филол. наук, профессор

Научно-технологический университет штата Миссури

г. Ролла, Соединенные штаты Америки

DOI 10.21661/r-562772

РОЛЬ МИКРО- И МАКРОСТРУКТУРЫ СЛОВАРЯ ТЕРМИНОВ В РЕШЕНИИ ЗАДАЧ СИНТЕЗА РУССКОГО ГЛАГОЛА

***Аннотация:** цель статьи – представить новые подходы к построению терминологического словаря. Рассматривается идея создания наиболее полного справочника по синтезу русского глагола, в структуру которого интегрирована обучающая база. Через микро- и макроструктуру словаря пользователь получает необходимые инструменты для решения актуальных задач синтеза. Разработанные алгоритмы представляют собой этапы решения поставленной задачи: от микроструктуры словаря к приложениям и блокам глагольных форм в макроструктуре. При этом доказано, что активные обучающие инструменты (формулы синтеза и готовые алгоритмы поиска), встроенные в структуру словаря, помогают решать основные задачи словообразовательного синтеза: создание нового значения от исходного с помощью заданных параметров; определение семантического потенциала глаголов; поиск максимально точного состава существующих в языке и узусе глагольных модификаций. Важным результатом работы является интерактивная связь словаря с Матричной системой форм глаголов звучания (около 12000 единиц), разработанной автором, и базой Национального корпуса русского языка (около 2 млрд слов).*

***Ключевые слова:** словообразовательный синтез русского глагола, глаголы звучания, семантическая модификация, семантический модификатор, словарь терминов, лексикографические параметры, макроструктура словаря, микроструктура словаря, Национальный корпус русского языка, НКРЯ.*

Проектируемый словарь-справочник словообразовательного синтеза является *аспектным* по своему типу (раскрывает различные аспекты изучения модификаций русского глагола и характер отношений между единицами синтеза),

лингвистическим по составу словника, *экспериментальным* по структуре и подаче материала и *обучающим* по назначению. Словарь создается с целью обучения носителей английского языка словообразовательному синтезу русского глагола. Алгоритмы словаря помогают формировать видение процесса синтеза нового значения: определять семантический потенциал того или иного глагола; устанавливать состав глагольных модификаций с помощью базы НКРЯ и *Матричной системы форм глаголов звучания* (разработана автором) [18–20].

Особенности строения терминологического словаря, значимость его в системе обучения русскому языку обсуждались в нашей предыдущей статье [8], где подробно рассматривалась структура словаря и его обучающая направленность; проводилось сравнение с классическими трудами по лексикографии [1] и новаторской работой Т.В. Жеребило, в которой было введено понятие *интерактивного компонента* при составлении словарей [6].

В настоящей статье внимание сосредоточено на новом аспекте – интеграции обучающей базы в состав лексикографического издания – когда в структуру словарной статьи вводится алгоритм исследовательской задачи. Термин *исследовательский* здесь употреблен не случайно: пользователь словаря терминов синтеза (носитель и / или не-носитель языка) выступает одновременно и как получатель информации, и как исследователь. В распоряжение читателя предоставлены механизмы научного анализа и синтеза, например, *компонентный анализ* – при рассмотрении словарных статей и вычленении элементарных смыслов (сем) и лексико-семантических вариантов (ЛСВ), *словообразовательный синтез* – при определении и построении модификаций, *лингвистическое моделирование* – при поиске схем и алгоритмов решения лингвистических задач.

Пользователь словаря при этом получает всю необходимую информацию для той или иной задачи: комплекс глагольных форм и модификаций, систему модификаторов глаголов звучания, схемы синтеза внутри различных лексических групп; свод сведений о лакунах и кодифицированных вариантах, а также другие необходимые сведения, которые отсутствуют в лингвистических словарях.

Структура словаря (его статьи, зоны) составлена так, чтобы читателю как можно меньше пришлось решать рутинных задач по поиску данных. В этом заключается *практическая направленность* и *значимость* работы по составлению словаря.

Научная новизна исследования заключается в создании и вводе в текст лексикографического издания поисковых алгоритмических формул. Для разработки методики составления и использования этих алгоритмов (выстраивания маршрутов задач – поиска путей от микроструктуры словаря к приложениям и блокам глагольных форм в макроструктуре), описанию их релевантных механизмов требуется решить следующие задачи:

1. Определить общий алгоритмический маршрут поиска нового значения по заданным параметрам (классическая задача синтеза);
2. Разработать методику нахождения модификаций, входящих в разные семантические области (разных глаголов);
3. Составить алгоритм поиска максимально точного состава глагольных модификаций (существующих в языке и узусе), используя *Матричную систему форм глаголов звучания* (около 12000 единиц) и *базу Национального корпуса русского языка* (около 2 млрд слов).

Методы исследования данной проблематики включают в себя все типы лингвистических методик: математическая статистика при обработке таблиц; лингвистическое моделирование – при построении схем решения задач и формул синтеза; лексикографические методики – при построении словарной статьи нового формата с включением формул в микроструктуру; структурирование макроструктуры словаря в соответствии с разработанными алгоритмами.

Теоретической базой исследования послужили работы по словообразованию и словообразовательному синтезу И.С. Улуханова [16], И.Г. Милославского [12; 13]; труды по аспектологии Н.С. Авиловой [1], А.В. Бондарко [3], М.Я. Гловинской [5], Ю.С. Маслова [11], М.А. Шелякина [17] и др. Последовательность изложения материала в статье повторяет алгоритм действий пользователя не-носителя языка: 1) знакомство с терминологией синтеза, словообразования,

аспектологии; 2) решение задач по синтезированию нового значения различной сложности; 3) применение на практике алгоритмов поиска глагольных модификаций.

Прежде чем перейти к рассмотрению конкретных задач, которые решаются в словаре, важно остановиться на вопросе формирования терминологии синтеза, в частности – глагольных модификаторов. Потребуется объяснение того, как соотносятся термины *синтеза*, *словообразования* и *аспектологии*, почему важно учитывать их в комплексе при составлении словника (части макроструктуры словаря).

Особенностью формирования терминов синтеза, в отличие, например, от терминов *аспектологии* и *словообразования*, является приоритет семантического значения над номинацией – звуковым оформлением терминоединицы. Номинация объектов и новых явлений в синтезе отличается особым маркированием – по звуковой форме термина нельзя сразу определить его семантическое наполнение. Компоненты значений модификаторов и модификаций не имеют особых наименований, отражающих суть их значений, таких, например, как *начинательность*, *однократность*, *спонтанность / неподготовленность действия* и т. д., а снабжены только цифровым и буквенным кодом: Mod1, Mod2,... Mod30, зная который по данным словаря можно определить их семантику и компонентность (например: «начало + один раз»; «начало + один раз + сильно + постепенно» и т. д.) [9, с. 28–31].

В этом главное отличие терминологии синтеза от терминов *аспектологии* (теории вида) и *словообразования*, где каждый термин именован в соответствии с его значением и так или иначе передает заложенный в нем смысл (ср., например, названия способов глагольного действия: *однократность длительного действия*, *однократность действия, не характеризованного по длительности* и т.п. [4], или названия групп глаголов: *глаголы со значением доведения кого-, чего-либо до нежелательного состояния* [16]).

На уровне исследования модификаций их номинативная функция не играет особой роли, более важным, по задаче синтеза, является компонентность модификации (одно- или многокомпонентный состав) и ранжирование внутри самой

структуры модификатора (модификации). Глагольные модификации проиндексированы по релевантным параметрам: а) по сложности компонентов (одно-, двух-, трех-, четырехкомпонентные); б) по конкретным смыслообразующим формантам (*на-*, *по-*, *раз-ся* и т. п.); в) по степени интеграции с корнями (индивидуальные модификации, стандартные); г) по семантическому потенциалу (узкий – образует до 5 модификаций, широкий – до 200 модификаций).

Все модификаторы проиндексированы числовыми обозначениями от 1 до 30 в зависимости от своего ранга в исследуемой группе – способности соединяться с глаголами звуковой семантики. Таким образом, числовой маркер при модификаторе – это не просто порядковый показатель в классификации – он наделяется смыслом – показать семантическую активность модификатора. Ранжирование происходит от меньшего показателя к большему – так, например, модификатор Mod1 занимает первую строчку в рейтинге, так как обладает большей активностью в группе, чем Mod 5 или Mod 30.

Кроме того, в отличие от терминов аспектологии и словообразования, термины синтеза передают заложенные в них идеи через наименования глагольных корней и аффиксов. Так, идея (значение) *однократности продолжительного действия* передается присоединением префикса *про-* в двух вариантах: *про₁-* и *про₂-*, которые могут указывать как на кратковременный характер действия – *про₂-* (вдруг *прокричал*, *пролаял*, *проговорил*), так и напротив, на характер чрезмерно длительного процесса – *про₁-* (*прокричал*, *пролаял*, *проговорил* целый час). Что подтверждается многочисленными примерами из НКРЯ (рис. 1):

ПРО₁:- *Эй, Парашка, Фомич, Кирюшка! - и прочая. Она прокричала почти целый календарь. Пазухин молча стоял перед нею и даже не улыбнулся. Н. В. Кукольник. Сержант Иван Иванович Иванов, или все заодно. "Мы всю ночь проговорили с ним о " Мертвых душах ", - вспоминает он, - и читали их, в который раз - не помню. Игорь Золотусский. «Записки сумасшедшего» [НКРЯ].*

ПРО₂:- *Соседский петух дважды прокричал и смолк; горлица где-то рядом стонала убаюкивающе и сонно: сти да сти; а вот говор людской тревожил непонятностью: кто-то разговаривал совсем рядом. Борис Екимов. Пиночет (1999). И что тогда не проговорили, замяли, заболтали по человеческой тупости, теперь проговорено, и нет уже ни зависти, ни ревности. Н. Б. Черных. Слабые, сильные [Там же].*

Рис. 1

Важно отметить, что в теории синтеза *идея однократности* не привязана к какому-либо одному префиксу или суффиксу, она входит в различные смысловые комбинации, что не позволяет закреплять за ней определенный модификатор и особое название. На Рис. 2 представлен состав модификаторов, куда входит значение однократности – *сема один раз*.

1-компонентный модификатор

один раз = про₂-, с- = Mod3 = *пропеть, спеть, просвистеть*

2-х компонентные модификаторы

один раз + повторно = пере- = Mod 22 = *перепеть*

один раз + начало = за- = Mod2 = *заговорить, запеть, засвистеть*

один раз + некоторое время = по-, про₁- = Mod1 = *поохать, проохать, позвенеть*

один раз + до некот. точки = до- = Mod28 = *донпеть, договорить*

один раз + мгновенно = -ну- / -ану- = Mod4 = *стукнуть, стукануть, крикнуть*

один раз + конец = от- = Mod9 = *отговорить, откричать, отпеть*

один раз + тщательно = вы- = Mod21 = *выпеть, выговорить*

3-х компонентные модификаторы

один раз + сверх предела/нормы + негат. оценка = из-...-ся, об-...-ся = Mod12 = *изреветься, исплакаться, искричаться*

один раз + слабо + дополнительно = при-...-ну-, под- = Mod13 = *притопнуть*

один раз + начало + спонтанно = вз-/вс- = Mod14 = *вскричать, взговорить*

один раз + конец + ниже предела/нормы = недо- = Mod18 = *недонпеть*

один раз + начало+ мгновенно = вз-...-ну- = Mod19 = *взвизгнуть, вскрикнуть*

один раз + некот. вр + слабо = на- = Mod26 = *напеть, наговорить*

один раз + сверх предела/нормы + позит. оценка = про-...-ся, вы-...-ся, на-...-ся = Mod8 = *прокричаться, выплакаться, нареветься*

4-х компонентные модификаторы

один раз + начало + сильно+ постепенно = раз-...-ся = Mod5 = *разахаться, раскричаться*

Рис. 2

Как показано на рис. 2, одно и то же значение передается несколькими сочетаниями сем (элементарных смыслов), что ведет к множеству возможных номинаций. Дополнительно к непростой системе терминов синтеза прибавляются словообразовательная и аспектологическая терминологии, в которых нет единства мнений и очевидных определений, где параллельно существует большое количество дублетов, омонимичных терминов и синонимов. Потому в терминах синтеза наиболее корректным правилом номинации на сегодняшний день остается цифровое маркирование (до тех пор, пока не будут найдены более точные определители). Однако в терминологическом словаре следует учитывать и сопоставлять терминологию разных языковых уровней, демонстрировать таким образом различия и сходства в наименованиях одних и тех же глагольных действий и процессов.

Рассмотрим теперь на практике, в каких терминах аспектологии (способах глагольного действия, или теории вида) представлены перечисленные выше единицы синтеза – *модификации* и их значения.

В описании способов глагольного действия (СГД) релевантными являются следующие характеристики, которые во многом перекликаются с выше-описанными модификаторами: *субъективная оценка действия говорящим* (экспрессивность); *интенсивность действия*; *результативность действия*; *кратность действия* (однократность, многократность) и т. п. В описании глаголов названные характеристики часто объединены в одном названии: *интенсивно-результативные*, *интенсивно-кратные* и т. д.

Однако наши исследования синтезированных значений звуковых глаголов показывают, что далеко не всегда эти характеристики и термины применимы к глаголам различных семантических групп.

Так, например, *интенсивно-результативный* способ глагольного действия, включающий глаголы со значениями исчерпанности и полноты – *выскрести*, *выпарить*, в семантике звуковых глаголов приобретает другие смыслы. В приведенных ниже примерах в глаголе *выпеть* релевантным

является больше смысл «тщательно» (i), а не значение исчерпанности и полноты, хотя оно также присутствует (ii):

(i) «... *бесчувственное, топорное прикосновение в силах разбить, даже без их ведома, лучшую драгоценность сердечную, медвежьей лапой ударить по тончайшим струнам душевным, данным на то, чтобы выпеть небесные звуки...*» Н.В. Гоголь. Выбранные места из переписки с друзьями [НКРЯ].

(ii) «*Только она им не пела своих любимых песен, эти песни она все про себя пела, словно берегла их, чтоб не выпеть, не израсходовать.*» Н. С. Лесков. Житие одной бабы [Там же].

Эта вполне очевидная для носителя языка информация в то же время является далеко не очевидной для изучающего язык, создает непреодолимые трудности в понимании материала. Одни и те же модификаторы, соединяясь с глаголами звучания разных семантических подгрупп, дают разные результаты – от абсолютного соединения, как, например, у модификатора *вы-...ива-* с антропофонами (глаголами, называющими звуки человека) – *выпевать, высвистывать, выкрикивать, выплакивать* – до полной их несовместности, как в зоофонах (звуках, издаваемых животными) – *!вымяукивать, !выкрикивать, !выкаркивать* и т. д. (здесь и далее знаками (!) и (!!)) обозначены некорректные варианты модификаций).

Однако важно отметить и связь между терминами разных языковых уровней. Особенно явно она проявляется при сопоставлении *окказиональных вариантов* глагольных форм в словообразовании и *потенциальных модификаций* в теории синтеза. По сути, и в том, и другом случае – это поиск некодифицированных дериватов, потенциальной части системы языка, которая может реализовываться по мере формирования актуальных запросов. Так же, как мы ищем в синтезе потенциальные лексемы на границе с лакунами, так и в словообразовании описываются те же варианты, только с точки зрения анализа и производности.

На рис. 3 приведен список *окказиональных глаголов* из книги И.С. Улуханова «Глагольное словообразование современного русского языка» [16],

который сравним с формами модификаций, не зафиксированных словарями, но для которых мы собрали доказательные примеры в НКРЯ [9, 18, 19].

ВЫСВИСТЕТЬ - «В это время певчий дрозд, управляющий зарей, вдруг отчетливо **высвистел** членораздельное слово: любии-и, любии-и!» М. Пришвин. Дневники;

ВЫЗВЕНЕТЬ, окказ. (все звуки) - «Р-равня-айсь! Х-хмиррна! - наконец взлетал над сбившимися в строй красноармейцами **вызвенеший** голос помкомзвода Яшкина». В. Астафьев. Прокляты и убиты.

ВЫРЫДАТЬ - «И от вопросов мужа: «Как ты?» - ей хотелось **разрыдаться**, уже не сдерживаясь, от души. **Вырыдать** все. Терпела». Роман Сенчин. Елтышевы.

ВЫРЫДЫВАТЬ, окказ. - «Поэты за чаем **вырыдывали** отрывки из вынашиваемых поэм». А. Г. Малышкин. Люди из захолустья.

ВЫСТАНЫВАТЬ, несов. - «Дома, затперев двери и окна, она забралась под суконное одеяло и всю ночь бессонно продрожала, сунув голову под подушку и лишь изредка в отчаянии **выстанывая**: «Господи Боже мой!..» Ю. Буйда. У кошки девять смертей.

ВЫХРИПНУТЬ - «- Меженин, Меженин... - **выхрипнул** Никитин и перевел дыхание: ему не хватило воздуха». Юрий Бондарев. Берег.

ВЫШЕПТАТЬ - «Губы **вышептали** имя супруги, протянутая рука застыла, пробуя что-то невидимое на ощупь». Анатолий Азольский. Монахи.

Рис. 3

Модификации и дериваты типа *выхлюпываться* мы не найдем ни в словарях, ни в НКРЯ, но некорректным будет делать вывод, что их не существует. Таким образом, и в словообразовательном, и в модификационном аспекте отсутствие *формы / модификации / деривата* в словаре или в НКРЯ – не показатель отсутствия слова в системе. Этот вывод следует из примеров, приведенных выше, и именно он побуждает к поиску и пополнению лексических баз новыми элементами. В поиске редких значений, не зафиксированных словарями, видим одну из основных задач синтеза. И ее решение также заложено в проектируемый словарь.

Рассмотрим теперь, как строится *алгоритм*, или *алгоритмический маршрут* поиска значения по заданным параметрам.

В нашей статье «Принципы лексикографического описания терминологии словообразовательного синтеза» [8] было описано принципиальное строение разрабатываемого терминологического словаря синтеза и дан ответ на вопрос, при какой макроструктуре наиболее ясно проявляются связи между терминами разных уровней (корневого, словообразовательного, веб-извлечения), отсылки к

таблицам, Грамматическому словарю А.А. Зализняка [7], Национальному корпусу русского языка [14], приложениям и т. д.

В продолжение этой темы рассмотрим, как в данную структуру вписываются различные алгоритмы решения задач синтеза. С помощью структурных элементов словаря (таблицы модификаторов и модификаций, модификационные матрицы, словники исходных глаголов и т. п.) покажем маршруты поиска решений и вовлечение пользователя в процесс указанной маршрутизации.

Поиск нового значения, или классическая задача синтеза, сформулированная еще в 1980 году И.Г. Милославским, заключается в том, чтобы «выяснить, как от слова, лежащего в вершине гнезда, можно образовать новое слово, если дано семантическое различие между исходным словом и искомым, а также с помощью каких формальных средств этот процесс можно осуществить» [13]. Актуальность задачи, поставленной более 40 лет назад, не уменьшается и сегодня, о чем говорит недавнее издание автора, дополненное новыми аргументами в слабой предсказуемости результатов синтеза: «...значения выражаются исключительно словообразовательными средствами, нерегулярность равно как и исключительно широкая омонимичность которых делают почти невозможным определение правил сочетаемости этих средств с модифицируемой основой» [12].

Принимая во внимание проблематичность самой идеи синтеза слова, ее недостаточную разработанность и неоднозначность, легко предположить, что исследования в синтезе трудноосуществимы без специальных знаний и навыков. Они требуют серьезной лингвистической компетенции, которой, разумеется, не обладают не-носители языка – целевая аудитория словаря. Потому разработка алгоритмов с готовыми поисковыми решениями может стать единственным инструментом, который поможет пользователю продвинуться в понимании данной проблемы.

Однако словарь предоставляет не только готовые решения (формулы, алгоритмы, таблицы, ссылки и т. д.), но и более абстрактные схемы, заполняя которые, можно получить искомый результат. Абстрактные схемы расположены по всему корпусу словаря и, чтобы воспользоваться ими, необходимо пройти по

заданному поисковому маршруту (по схемам маршрутизации, заложенным в словарных статьях и задачах).

Маршрутизация – это, говоря терминами сетевого распределения – процесс принятия решений более высокого уровня, который направляет пользователя от источника материала к месту назначения через промежуточные блоки информации. В нашем случае маршрутизация так же предполагает передачу пользователю данных – тех, которые необходимы на определенном этапе поиска решения.

С помощью маршрутизации в словаре решаются все перечисленные задачи синтеза, корпус которых задан в начальном разделе – *Задачи синтеза*. Рассмотрим одну из них – поиск *синтезированного значения от исходного глагола по заданному параметру*.

Формулировка задачи предполагает использование элементарных математических символов и интуитивно понятных формул. Пример формулировки задачи и алгоритма ее решения представлен на Рис. 4.

Задача 1. Поиск модификаций со значением «начало 1».
Алгоритм решения: $N_{1,2,3,..} + за- ... -ива- = Mod_{1,2,3,..}$

Рис. 4

Условие задачи записывается в виде обобщенной формулы поиска модификации и алгоритма ее решения. Формула, в свою очередь, предполагает подстановку любого глагола звучания. В условии задан *модификатор* с искомым значением и прописана его индексация – *начало 1*, которая соответствует заданному форманту *за-...-ива-*. Цифровая индексация при модификаторе означает, что значение «начало» представлено в нескольких вариантах (модификаторах).

По аналогии с элементарным (однозначным) модификатором *один раз*, который входит в состав 16 более сложных сочетаний, модификатор «начало» промаркирован от 1 до 6 в соответствии с количеством вхождений, показанных на рис. 5.

- Начало 1* (за-ива-)
Начало 2 + однократно (за-)
Начало 3 + спонтанно (вз-...-ива-)
Начало 4 + однократно + спонтанно (вз- / вс-)
Начало 5 + однократно + мгновенно (вз-...-ну-)
Начало 6 + однократно + сильно + постепенно (раз-...-ся)

Рис. 5

В каждом из данных сочетаний значения *начало 1*, *начало 2*, *начало 3*, *начало 4*, *начало 5*, *начало 6* дают различные результаты, так как осложнены дополнительными смыслами. Покажем это на примере глаголов *кричать* и *щелкать*.

Кричать, щелкать + Начало 1 = *!закрикивать, !защелкивать*

Кричать, щелкать + Начало 2 = *закричать, защелкать*

Кричать, щелкать + Начало 3 = *вскрикивать, !вщелкивать*

Кричать, щелкать + Начало 4 = *вскричать, !вщелкать*

Кричать, щелкать + Начало 5 = *вскрикнуть, !вщелкнуть*

Кричать, щелкать + Начало 6 = *раскричаться, расщелкаться*

Из приведенных выше сочетаний следует сделать вывод, что *начало 1* в изолированном виде не присоединяется к глаголам *кричать* и *щелкать*, а звуковые глаголы *!закрикивать* и *!защелкивать* не могут служить показателем положительного синтеза. Здесь попутно отметим, что смысл *начало 1* в неосложненном чистом виде крайне редко присоединяется к глаголам звучания, словарями зафиксированы единичные случаи: *запевать, заговаривать, завывать, засвистывать*. Значения *!вщелкивать* и *!вщелкнуть* также нельзя признать положительным результатом синтеза с образованием модификации. Здесь следует привести важное наблюдение из практики преподавания: если носитель языка без лингвистической подготовки без труда сделает вывод о неправильности каких-то полученных модификаций, то не-носитель языка на данном этапе исследования не способен сделать корректное заключение о положительном или отрицательном результате синтеза. Не обладая достаточной языковой компетенцией, он не сможет дифференцировать полученные значения по признаку *правильно /*

неправильно и глаголы *!защелкивать* и *!закрикивать* для него такие же «правильные», как *раскричатся* и *расщелкаться*.

Проблема дифференциации полученных значений по признаку *правильности / неправильности* решается в словаре следующим образом. После того как определены исходные элементы задачи, в данном случае, это *глагол* и *модификатор*, вводится зона *Маршрут*, где подробно обозначены поэтапные действия по поиску необходимой информации, которая позволит пользователю делать корректные выводы.

Маршрут:

Исходный глагол → *Словник глаголов звучания;*

Модификатор Начало → *Индекс модификаторов;*

Формула синтеза → *Словарь терминов, Толковый словарь модификаций*

Модификация Mod_{1, 2, 3...} → *Матричная таблица*

Запись маршрута означает следующие этапы действий:

1) выбрать требуемый глагол из *Словника* (глаголы звучания в исходной форме: *ахать, бахать, барабанить...* и т. п.);

2) выбрать индекс требуемого модификатора в разделе *Индекс модификаторов* (*начало 1, начало 2, начало 3...* и т. п.);

3) в *Словаре терминов* и в *Толковом словаре модификаций* выбрать требуемую формулу синтеза: *исходный глагол + формант модификатора;*

4) ввести выбранные глаголы и форманты в формулу и зафиксировать предварительные модификации: *ахать, бахать, барабанить... + за-...-ива- = заахивать, забахивать, забарабанивать...* и т. п.;

5) проверить полученные модификации *Mod1* (*заахивать*), *Mod2* (*забахивать*), *Mod3* (*забарабанивать*) и т. п. на наличие их в системе языка (кодифицированность в словарях) и в узусе (распространение в речи) по *Матричной таблице* [18; 19] и базе НКРЯ [14].

Далее на основе полученных данных следует сделать вывод о *наличии / отсутствии* искомой модификации. Если полученные модификации, как показано на рис. 6, дают отрицательный результат в НКРЯ, но находятся в

Матричной таблице, следует сделать вывод, что данная модификация потенциально существует в языке и является «правильной».

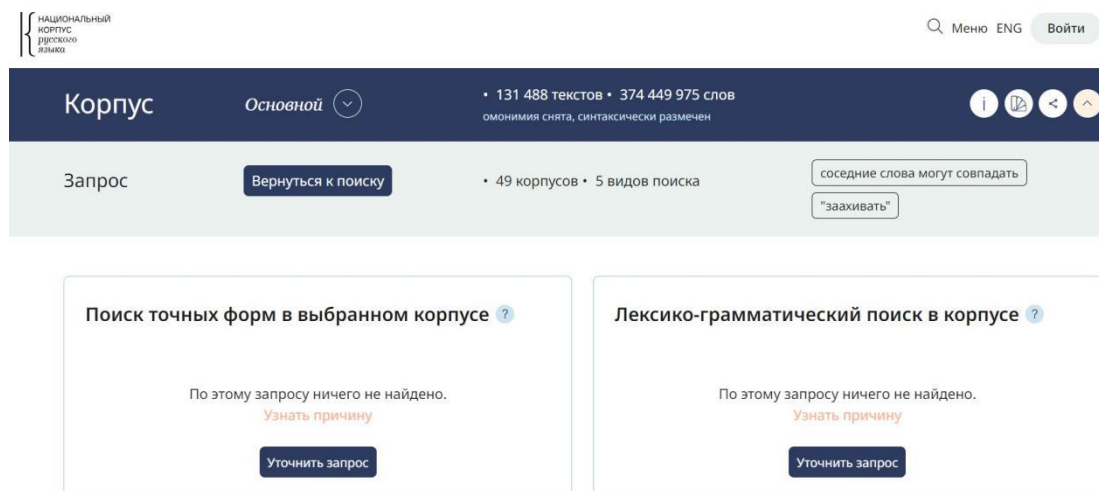


Рис. 6

Заахивать: по запросу ничего не найдено

Забахивать: по запросу ничего не найдено

Забарабанивать: по запросу ничего не найдено

Защелкивать: *наручники, замки* и т. п. механизмы; по запросу о звуковых глаголах ничего не найдено.

Если же, как показано выше, ни одна модификация не подтвердилась в НКРЯ и поиск в *Матричной таблице* дал отрицательные результаты, следует сделать вывод о вероятном отсутствии данной модификации в языке.

Это было решение задачи *соединения (синтеза) многих глаголов с одним элементарным модификатором* (смыслом) или группой модификаторов с тем же значением в составе (многокомпонентными модификаторами).

В исследовании представлено около 150 таких глаголов, которые отобраны из Словаря Ожегова [15]. Если же брать выборку слов из больших толковых словарей (17-томного БАС, например), то глаголов, объединенных идеей звучания, будет намного больше, что говорит о незакрытых границах изучаемой группы.

Задачей более высокого уровня сложности является определение семантического потенциала того или иного глагола звучания, то есть его активного (кодифицированного) корпуса, когда требуется соединить исходный глагол со

всеми возможными модификаторами и синтезировать комплекс модификационных значений.

Решение данной задачи также начинается с алгоритма, где требуется обращение к словарю терминов, а именно – к тем единицам, которые уже заявлены в условии и выделены шрифтом – *семантический потенциал, кодификация, кодификационный корпус, исходный глагол, модификатор, модификация, синтезированное значение*.

Микроструктура словаря, то есть строение словарной статьи, позволяет получить всю лингвистическую информацию, требуемую в решении задачи. На Рис. 7 представлена статья одной из терминов – *синтезированное значение*.

СИНТЕЗИРОВАННОЕ ЗНАЧЕНИЕ, Σ . То же, что *модификационное значение* - производное значение глагола в его отдельном лексико-семантическом варианте (см. ЛСВ), полученное по формуле (схеме) синтеза: **исходный глагол** N_{1-142} + **модификатор** $Mod_{1-32} = \Sigma$, где N_{1-142} – любой глагол звучания в исходной форме, выбранный из списка *Словник глаголов звучания*), Mod_{1-32} – любой модификатор, выбранный из *Индекса модификаторов*).

Пример: синтезированное значение (модификация Mod) исходного глагола *лгать* и модификатора *про-* выглядит так: Σ *лгать* + *про-* = *пролгать*.

Рис. 7

Далее предлагается несколько вариантов решения. Самый краткий и простой маршрут – это путь к модификационной таблице (см. Табл. 1), где в алфавитном порядке собраны все возможные модификаторы, актуальные для глаголов звучания. Кроме того, в каждой клетке приведены примеры ее заполнения, где некорректные варианты выделены знаками малой вероятности (!) или невозможности (!!)

Таблица 1

Пример заполнения модификационной таблицы

| <i>Плакать</i> | | | | |
|--|-------------------------------------|-------------------------------------|---|---|
| <i>-ва-/ -ива-/ -ыва- !плакивать</i> | <i>вы-...-ива- !выплакивать</i> | <i>за-...-ива- !заплакивать</i> | <i>недо-...-ива !!недоплакивать</i> | <i>по-, про- поплакать, проплакатьI</i> |

| | | | | |
|---|--|---|---|--|
| <i>вз-/вс- !всплакать</i> | <i>вы-...-ся, вы- плакаться про-...-ся, про- плакаться на-...-ся наплакаться</i> | <i>из-...-ся !исплакаться об-...-ся !!обплакаться</i> | <i>-ну- !!плакнуть</i> | <i>по-...-ива- поплакивать</i> |
| <i>вз-/вс-...-ива- всплакивать</i> | <i>до- !доплакать</i> | <i>на-!наплакать</i> | <i>от- отплакать</i> | <i>под- !!подплакать при-...-ну- !!приплакнуть</i> |
| <i>вз-...-ну- всплакнуть</i> | <i>до-...-ива- !!доплакивать</i> | <i>на-...-ива- наплакивать</i> | <i>от-...-ива- отплакивать</i> | <i>под-...-ива- !подплакивать</i> |
| <i>вз-/вс-...-ива- !всплакивать</i> | <i>до-...-ся доплакаться</i> | <i>на-...-ся наплакаться вы-...-ся выплакаться</i> | <i>пере- !переплакать</i> | <i>при-...-ива- !приплакивать</i> |
| <i>вы- выплакать</i> | <i>за- заплакать</i> | <i>недо- !недоплакать</i> | <i>пере-...-ива- !переплакивать</i> | <i>про₂- проплакать с- !сплакать</i> |
| <i>про-...-ива- проплакивать</i> | <i>про-...-ся см. вы-ся</i> | <i>раз-...-ся расплакаться</i> | <i>с- см. про₂-</i> | |

Второй вариант решения, более сложный, предполагает лингвистическую компетенцию и навыки в решении задач первого типа – это последовательный синтез 30-и возможных модификаций, принцип поиска которых был рассмотрен выше. Здесь остановимся подробнее на определении ЛСВ глаголов, к которым отсылает алгоритм решения задачи.

В маршруте *Исходный глагол → Словник глаголов звучания → Толковый словарь модификаций* заложено логичное обращение к толкованию глагола, а именно – к его лексико-семантическим вариантам, так как не каждое значение глагола содержит сему звука, а, по условию задачи, следует синтезировать только глаголы звучания.

Таким образом в маршруте заложена следующая последовательность действий: клик на глаголе, например, *ахать*, в *Словнике* или *Матричной таблице* (Рис.8) перемещает в словарную статью (Рис.9), где и уточняется значение

глагола и количество лексико-семантических вариантов (ЛСВ), если глагол многозначный.

| | B | C | D | E | F | G | H |
|----|-------|-------|---------|---------|-------|--------|--------|
| 1 | Root | Word | Pronoun | PRESENT | PAST | FUTURE | IMPER |
| 2 | ахать | ахать | -- | -- | -- | -- | -- |
| 3 | ахать | ахать | -- | -- | -- | -- | -- |
| 4 | ахать | ахать | -- | -- | -- | -- | -- |
| 5 | ахать | ахать | Я | ахаю | ахал | -- | -- |
| 6 | ахать | ахать | Я | ахаю | ахала | -- | -- |
| 7 | ахать | ахать | Я | ахаю | ахало | -- | -- |
| 8 | ахать | ахать | Ты | ахаешь | ахал | -- | ахай |
| 9 | ахать | ахать | Ты | ахаешь | ахала | -- | -- |
| 10 | ахать | ахать | Ты | ахаешь | ахало | -- | -- |
| 11 | ахать | ахать | Он | ахаёт | ахал | -- | -- |
| 12 | ахать | ахать | Она | ахаёт | ахала | -- | -- |
| 13 | ахать | ахать | Оно | ахаёт | ахало | -- | -- |
| 14 | ахать | ахать | Мы | ахаём | ахали | -- | -- |
| 15 | ахать | ахать | Вы | ахаёте | ахали | -- | ахайте |

Рис. 8

- АХАТЬ**, -а ю, -а е ш ь. 1. Произносить, вскрикивать "ах", выражая какое-либо чувство (удивление, восторг, печаль, сожаление) <...>. 2. Издавать громкий, сильный, отрывистый звук при ударе, разрыве, выстреле (обычно о пушках, минометах и т.п.) <...>

Рис. 9

Для чего еще, помимо определения количества ЛСВ, следует уточнять значение глагола? Чтобы понимать, *что именно* (какой из предметов, машин, механизмов) или *кто* (человек, животное) является источником звука, так как данный факт в первую очередь влияет на модификационный потенциал, который требуется определить в задаче.

Так, в случае с глаголом *ахать* имеем два разных источника звука (человек и механизм) и два значения ЛСВ, которые считаем разными словами (*ахать1* «произносить звук» и *ахать2* «издавать звук») с разной сочетаемостью, разной деривационной способностью и, соответственно, на выходе получим разные слова [9].

Когда определены все возможные модификации отдельно взятого глагола звучания по заданным 30-и параметрам, следует обратиться, как и в случае решения первой задачи, к *Матричной таблице*, которая всем своим составом (12000 единиц) привязана к базе НКРЯ (2 млрд слов).

Остановимся подробнее на описании данной интеграции. В макроструктуре словаря алфавитное расположение материала уступает место более

приоритетному в данном случае – блоковому – структура строится только на тех статьях, главах, разделах и т. п., которые являются звеньями решения задач. В этом заключается новизна и особенность проектируемого словаря.

Матричная таблица является одним из блоков словаря, без которого невозможно проверить результаты синтеза и собрать доказательную базу достоверного существования модификаций, без которой, в свою очередь, процесс синтеза нового значения теряет научный смысл и переходит в разряд гипотетических теорий.

Под задачи синтеза разрабатываются оригинальные системы отсылок и дополнения к словарю – приложения, таблицы, схемы, комментарии – которые являются полноценными частями алгоритмов и без которых невозможно решение тех или иных задач [8]. Все разделы словаря, так или иначе, привязаны к *Матричной таблице*:

Словник глаголов звучания → Матричная таблица

Значение глагола → ЛСВ → Матричная таблица

Источник звука → Значение глагола → Матричная таблица

Список модификаторов → Словник глаголов звучания → Матричная таблица

Модификаторы, релевантные для источника → Матричная таблица

Формулы синтеза → Модификации → Матричная таблица

Матричная таблица → Национальный корпус русского языка

При создании обучающего словаря ставится задача – поиск наиболее адекватных (не усложненных и не упрощенных) способов представления материала. Такая подача информации наиболее доступна пользователям разных уровней подготовки. Однако при использовании как Матричной таблицы, так и НКРЯ существуют проблемы, которые пока не решены. Речь идет о таком системном явлении как омонимия, которая присутствует во всех лексических группах, не исключая глаголы звучания.

В словаре терминов на данном этапе проблема дублетов и «двойных значений» решается вводом их обязательного индексирования. Индексы вводятся при

модификациях незвукового содержания при омонимии, переносных значениях и идиоматических сочетаниях: *в ушах!звенит; и звезда с звездю!говорит;!отговорила роща золотая; сердце!застучало* и т. п.

И если в тексте словаря такая индексация введена, то в *Матричной таблице* она находится в стадии разработки. То есть слова омонимы и незвуковые ЛСВ рассматриваются как разные слова с идентичным звуковым оформлением. Дается также информация о том, что не следует путать омонимию и незвуковые модификации с многозначностью звуковых глаголов.

Например, указано, что омонимы и дублиеты рассматриваются как два разных слова. Ср., например, примеры на Рис 10: *чокать1* (говорить на диалекте), *чокать2* (производить выстрел), *чокать3* (издавать звук, о птицах).

(i) Конечно, были в речи старших моих родственников и некоторые различия (наш большой «клан» объединял выходцев из разных районов Прикамья): кто-то цокал (цисто вместо чисто), кто-то **чокал** (курича вместо курица). В. З. Санников. Записки простодушного [НКРЯ].

(ii) Гумана здесь не было, а пули в стволы **чокали** - только щепка летела. Борис Васильев. А зори здесь тихие [Там же].

(iii) Мы весело топтали кусты. Испуганный дрозд **чокал** свое: чок-чок-чок... Владимир Брагин. В стране дремучих трав [Там же].

Рис. 10

Главное отличие рассматриваемых выше явлений – в словарной интерпретации: зона омонимии в словаре всегда представлена как зона отрицательной информации.

Вторая трудность представления информации в Матричной таблице касается такого распространенного явления в русском языке, как употребление глаголов будущего времени в роли прошедшего, что порождает множество дублетов, неразличимых по форме и трудных для понимания и идентификации: *а он возьми да и скажи, возьми да и спроси, а они возьми да и смолкни, Рихтер возьми да и сыграй на бис, возьми да и разревись, а он возьми да и расскажи, она ему возьми да и брякни* [16] и т. д.

Выводы

1. Исследование показало, что термины трех языковых уровней словообразовательного синтеза, аспектологии и теории словообразования взаимосвязаны и должны быть отражены в терминологическом словаре глагольного синтеза. Терминологию разных уровней объединяет поиск некодифицированных дериватов, потенциальной части системы языка, которая может реализовываться по мере формирования актуальных запросов.

2. Основные задачи синтеза – поиск модификационного значения от исходного глагола по заданному параметру могут решаться с помощью алгоритмов, вписанных в словарную статью в виде готовых поисковых решений – маршрутов. Маршрутизация – это инструмент, которым может воспользоваться не-носитель языка.

3. Матричная таблица является одним из блоков словаря, без которого невозможно проверить результаты синтеза и собрать доказательную базу достоверного существования модификаций.

Список литературы

1. Авилова Н.С. Вид глагола и семантика глагольного слова: дис. ... д-ра филол. наук / Н.С. Авилова. – М., 1976. – 408 с.
2. Ахманова О.С. Словарь лингвистических терминов / О.С. Ахманова. – М.: Советская энциклопедия, 1966. – 608 с.
3. Бондарко А.В. Вид и время русского глагола (значение и употребление) / А.В. Бондарко. – М.: Просвещение, 1971. – 239 с. EDN WENUJL
4. Гайдуков В.Г. Значение однократности действия и средства его выражения в современном русском языке: автореф. дис. канд. филол. наук / В.Г. Гайдуков. – М., 1989. – 20 с.
5. Гловинская М.Я. Семантические типы видовых противопоставлений русского глагола / М.Я. Гловинская. – М.: Наука, 1982. – 155 с. EDN PXNNHN
6. Жеребило Т.В. Словарь лингвистических терминов / Т.В. Жеребило. – 5-е изд., испр. и доп. – Назрань: Пилигрим, 2010. – 486 с. – EDN RUSOUC

7. Зализняк А.А. Грамматический словарь русского языка. Словоизменение / А.А. Зализняк. – М.: Русский язык, 1980. – 881 с. [Электронный ресурс]. – Режим доступа: <https://www.morfol.ru/словоформа/ахать> (дата обращения: 08.07.2024). EDN UFUIPL
8. Ивлиева И.В. Принципы лексикографического описания терминологии словообразовательного синтеза (на материале глаголов звучания русского языка) / И.В. Ивлиева // Интерактивная наука. – 2023. – №6 (82).
9. Ивлиева И.В. Экспериментальный модификационный словарь русского языка (на материале глаголов звучания) / И.В. Ивлиева. – М.: Азбуковник, 2013. – 467 с.
10. Ивлиева И.В. Русские глагольные модификации (опыт составления словаря) / И.В. Ивлиева. – М.: Элпис, 2008. – 275 с.
11. Маслов Ю.С. Избранные труды. Аспектология. Общее языкознание / сост. и ред. А.В. Бондарко, Т.А. Майсак, В.А. Плунгян. – М.: Языки славянской культуры, 2004. – 840 с.
12. Милославский И.Г. Современный русский язык. Культура речи и грамматика: учебное пособие для вузов / И.Г. Милославский. – М.: Юрайт, 2024. – 163 с.
13. Милославский И.Г. Вопросы словообразовательного синтеза / И.Г. Милославский. – М.: Наука, 1980. – 295 с.
14. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа: <http://www.ruscorpora.ru/> (дата обращения: 08.07.2024).
15. Ожегов С.И. Словарь русского языка / С.И. Ожегов; под ред. Н.Ю. Шведовой. – М.: Русский язык, 1988. – 750 с.
16. Улуханов И.С. Глагольное словообразование современного русского языка / И.С. Улуханов. – М.: Азбуковник, 2017. – 516 с.
17. Шелякин М.А. Категория вида и способы действия русского глагола: теоретические основы / М.А. Шелякин. – Таллин: Валгус, 1983. – 216 с.
18. Koob P., Ivliyeva I. An Assessment of the Propagation of Contemporary Russian Mechanical Phonation Verbs to Online Dictionaries. Web scraping project by

Missouri S&T, IT and ALP departments [Electronic resource]. – Access mode: https://scholarsmine.mst.edu/artlan_phil_facwork/157/

19. Koob P., Ivliyeva I. Russian National Corpus Web Scraping Project 2019–2020 [Electronic resource]. – Access mode: https://scholarsmine.mst.edu/artlan_phil_facwork/158/ (дата обращения: 08.07.2024).

20. Ivliyeva Irina V., Koob Perry. Multi-dimensional scaling of web-scraping results from the A.A Zalizniak Grammatical Dictionary and the Russian National Corpus. Creating a corpus fragment of all possible word-forms of modified Russian sound verbs using web-scraping methodology. Compilation of a summary table for the present tense, past tense, future tense, imperative, imperfective, and perfective gerund forms [Electronic resource]. – Access mode: https://scholarsmine.mst.edu/research_data/12 (дата обращения: 08.07.2024).