

УДК 510.62

М. Ф. БОНДАРЕНКО, канд. техн. наук, *Н. М. БУРКОВА*, *Н. В. ШАРОВА*
НОВА

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ СУФФИКСАЛЬНОГО СЛОВООБРАЗОВАНИЯ ГЛАГОЛОВ

В общей системе словообразования русского языка глагол занимает очень важное место в основном из-за продуктивного в словообразовании глаголов способа префиксации, когда создаются глаголы с иным лексическим значением, чем то, которое было в производящем слове. Однако в словообразовательную систему глаголов входят также и суффиксы, хотя их значения здесь не так богаты и разнообразны, как в именах существительных и прилагательных. Тем не менее, без них описание системы словообразования глаголов было бы неполным, односторонним. Изучение словообразовательной системы дает возможность не только понять законы ее функционирования, установить нормы современного словообразования, но и дать математическую интерпретацию тем или иным процессам в естественном языке. Формализация процессов словообразования является одной из самых актуальных проблем при решении целого ряда практических задач, связанных с переработкой информации на естественном языке с помощью современных ЭВМ.

Под словообразовательной системой понимаются те элементы, из которых состоят слова, т. е. значимые единицы сло-

ва (корни, приставки, суффиксы, окончания), их роль в слово-производстве, сама структура слов, употребляемых в языке (слова производные и непроизводные), а также те законы, по которым одни слова (производные) мотивируются другими (производящими), способы образования слов и те словообразовательные типы (внутри способов образования), которые активно участвуют в образовании новых слов.

В словообразовательной системе глаголов особо следует различать словообразовательные и формообразующие аффиксы. Словообразовательные аффиксы (мы будем рассматривать лишь суффиксы) используются для образования слов с новым лексическим значением, например, *батрак* — *батрачить*, *асфальт* — *асфальтировать* и т. п. Формообразующие суффиксы новых слов не образуют, они не меняют лексического значения слова, а используются для образования форм того же самого слова: *говорить* — *говорящий*, *читать* — *читал*, *читала*, *читали* и т. п. Мы остановимся на рассмотрении словообразовательных суффиксов глаголов. В нашем понимании словообразовательный суффикс — это служебная морфема, находящаяся после корня (или после суффикса производящей основы) и служащая для образования новых слов.

Рассматривая естественный язык как детерминированный, дискретный и конечный объект [1] и выделяя в тексте такой его фрагмент, как суффикс, который часто сам поддается дальнейшему фрагментированию, введем отношение $L(X, Y, Z) = 1$, связывающее между собой мотивирующее слово X , мотивируемое слово Y (продукт деривации) и набор грамматических и семантических признаков Z , позволяющих однозначно из X получить Y . Принцип однозначности состоит в том, что фрагмент текста (суффикс или его фрагмент) однозначно зависит от его смысла (определение смысла дано в [1]). Принцип однозначности можно интерпретировать как требование полноты набора признаков (компонентов смысла), с помощью которого осуществляется выбор фрагмента текста.

Математическое описание суффиксального словообразования глаголов сводится к формированию суффикса $s = s_1 s_2 \dots s_m$ в зависимости от набора грамматических и семантических признаков, где m — число букв в суффиксе.

Все словообразовательные суффиксы глаголов целесообразно представить в виде так называемого нами *регистра суффиксов*.

На основании фактов, изложенных в [2], можно разделить суффиксы глаголов на фрагменты подобно тому, как это делалось для суффиксов имен существительных. Полагаем при этом, что фрагмент суффикса несет в себе ту же семантику, что и суффикс, в который он входит, т. е. самостоятельным значением фрагмент суффикса обладает лишь тогда, когда он совпадает с простым суффиксом.

При решении задачи фрагментации суффиксов глаголов (как видно из регистра суффиксов) выделяются следующие четыре группы фрагментов суффиксов (основной критерий — наличие одинаковой буквы, которую мы назовем фиксированной): 1) фрагменты с фиксированной второй буквой «в»: *ов—ев, ив—ыв—*—в;

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
а	в	и	р	о	в	а	—
и	в	—	р	о	в	а	—
—	—	и	р	о	в	а	—
—	—	—	—	о	в	а	—
—	—	—	—	е	в	а	—
—	—	—	—	и	в	а	—
—	—	—	—	ы	в	а	—
—	—	—	—	—	в	а	—
—	—	—	—	—	—	а	—
—	—	—	—	—	—	я	—
—	—	—	—	—	—	е	—
—	—	—	—	и	—	—	—
—	—	—	н	и	ч	а	—
—	—	а	н	и	—	у	—
—	—	—	н	—	—	у	—

2) фрагменты с фиксированной второй буквой — (пробел): *а—*—я, *р—*—, *у—*; 3) фрагменты с фиксированной второй буквой «н»: *—н, ач*; 4) фрагменты с фиксированной первой буквой «и»: *из, ир, ич, и—*.

Введем следующие обозначения: s_1 — первая буква фрагмента суффикса со значениями *о, е, и, ы, —, а, я, у*; s_2 — вторая буква фрагмента со значениями *в, —, н, р, з, ч*.

Рассмотрим последовательно каждую из четырех выделенных групп фрагментов. В каждой из них наблюдаются чередова-

ния гласных букв при фиксированной согласной или пробел (—), или чередования согласных букв (в последней группе) при фиксированной гласной. На выбор буквы фрагмента влияют грамматические признаки основы Z , семантические признаки основы W и класс чередований μ , который выделяется при выборе буквы фрагмента, т. е. $S = f(Z, W, \mu)$.

Класс чередования μ зависит от грамматических признаков слова Z и семантических признаков W : $\mu = \varphi(Z, W)$. Влияние слова Z на выбор буквы фрагмента суффикса может быть охарактеризовано следующим набором признаков: $Z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7)$, z_1 — часть речи мотивирующего слова со значениями *с* — имя существительное; *п* — имя прилагательное; *г* — глагол; *м* — местоимение; *ч* — числительное; *з* — звукоподражательные междометия. z_2 — признак последней буквы основы со значениями *а, б, б', в, в', г, г', д, д', е, ж, з, з', и, к, л, л', м, м', н, н', о, п, п', р, р', с, с', т, т', у, ф, х, ц, ч, ч, ы, ы, э, ю, я* (буква со штрихом обозначает мягкую согласную); z_3 — признак происхождения основы со значениями *р* — русская основа, *и* — иностранная основа; z_4 — признак наличия или отсутствия следующего фрагмента суффикса со значениями *д* — есть следующий фрагмент, *н* — нет следующего фрагмента; z_5 — признак наличия или отсутствия префикса со значениями *п* — есть префикс, *б* — нет префикса; z_6 — признак ударности фрагмента со значениями *у* — ударный, *б* — безударный; z_7 — признак изменчивости состояния со значениями *д* — состояние изменяется, *н* — состояние не изменяется.

Семантические признаки W описываются следующим набором: $W = (w_1, w_2, w_3)$ где w_1 — совершать действие, имеющее отношение к тому, что названо мотивирующим существительным: a — действовать с помощью предмета (*плужить*), b — наделять предметом (*асфальтировать*), v — приводить в состояние (*восторгать*), z — совершать действие, свойственное лицу или животному (*бродяжничать*), d — делать кем-нибудь (*пролетаризировать*), e — совершать действие, свойственное для данного места (*иллюзовать*), $ж$ — помещать в данное место (*госпитализировать*), $з$ — проводить время (*зимовать*), $и$ — проявлять признак лица или предмета (*сиротеть*); w_2 — совершать действие, имеющее отношение к тому, что названо мотивирующим прилагательным: a — приобретать признак (*белеть*), b — наделять признаком (*активизировать*), v — проявлять признак (*лютовать*), w_3 — совершать действие, имеющее отношение к тому, что названо мотивирующим глаголом: a — однократно совершать действие (*ключнуть*), b — интенсивно совершать действие (*долбануть*).

Рассмотрим первый класс фрагментов суффиксов с фиксированной буквой «в». При выборе гласной s_1 выделяются следующие классы чередований: $\mu_1 = s_1^0 V s_1^e$, $\mu_2 = s_1^v V s_1^h V s_1^r$. Для данного класса фрагментов $\mu_1 V \mu_2 = 1$. Записываем правила проявления того или иного значения внутри каждого класса чередования.

Для $\mu_1 = s_1^0 V s_1^e$. Буква «о» появляется после всех согласных, кроме *ч, ц, ш, щ, ж*. Например, *плутовать, иллюзовать*. Уравнение алгебры конечных предикатов будет иметь следующий вид: $s_1^0 = P_1 V P_3 V z_2^0$, $P_1 = z_2^0 V z_2^v V z_2^h V z_2^r V z_2^v V z_2^h V z_2^p V z_2^c V z_2^r$, $P_3 = = z_2^r V z_2^k V z_2^x$. Буква «е» появляется после букв *а, е, ж, о, у, ч, ш, щ, я, о', в', д', з', л', м', н', п', р', т'*. Например, *циклевать, горевать*.

Уравнение имеет следующий вид: $s_1^e = P_2 V P_4 V P_5 V P_6 V P_7$, $P_2 = z_2^a V z_2^o V z_2^y$, $P_4 = z_2^x V z_2^v V z_2^w V z_2^z$, $P_5 = z_1^i V z_2^r V z_2^h V z_2^p$, $P_6 = z_2^m V z_2^p$, $P_7 = z_2^o V z_2^b V z_2^c V z_2^r$.

Для $\mu_2 = s_1^v V s_1^h V s_1^r$. Если основа оканчивается на *г, ж, к, о, х, ч, ш, щ, з, л', н', с', р'*, то первая буква фрагмента суффикса «и», например: *обменивать, проталкивать*; если же основа оканчивается на *б, в, д, з, л, м, н, п, р, с, т*, то первая буква «ы», например: *обрабатывать, завербовать*. Буква «_» появляется в том случае, если последняя буква основы *а, е, и, у, ы*, например: *утаивать, девать*.

Уравнение алгебры конечных предикатов будет иметь следующий вид: $s_1^r = P_3 V P_5 V P_4 V z_2^o V z_2^c$, $s_1^h = P_1$; $s_1^r = P_2 V z_2^h V z_2^z$.

В случае присоединения фрагмента суффикса «ив» в последней букве основы происходит чередование $t \rightarrow ч$ (*засекретить — засекречивать, чертить — ычерчивать*). Если обозначить последнюю букву основы через t_1 , то можно записать, что $t_1^r V t_1^c = \tau_1$, где τ — класс чередования в последней букве основы.

Выбор класса чередования μ в первой букве фрагмента зависит от части речи мотивирующего слова. Если глагол образован от имени существительного, тогда выбирается $\mu_1 = z_1^c Vz_1^n$. Если глагол образован от глагола, то $\mu_2 = z_1^f$.

Рассмотрим второй класс фрагментов суффиксов с фиксированной буквой « $_$ ». Это фрагменты: $a _ - я _, e _, y _$.

При присоединении фрагмента « $a _$ » происходит чередование в последней букве основы $г \rightarrow ж$ (*долг — должать*), $к \rightarrow ч$ (*венок — венчать*), $х \rightarrow ш, щ$ (*страх — стращать, ветхий — ветшать*). Выделяем следующие классы чередований в последней букве основы: $t_1^f V t_1^k = \tau_1$, $t_1^k V t_1^n = \tau_2$, $t_1^x V (t_1^w V t_1^w) = \tau_3$, $\tau_1 V \tau_2 V \tau_3 = 1$.

При выборе первой буквы s_1 фрагмента выделяются следующие классы чередований: $\mu_1 = s_1^a Vs_1^a$, $\mu_2 = s_1^e$, $\mu_3 = s_1^y$. Для данного класса фрагментов можно записать $\mu_1 V \mu_2 V \mu_3 = 1$.

Записываем правила появления того или иного значения внутри каждого класса чередований первой буквы фрагмента.

Для $\mu_1 = s_1^a Vs_1^a$. Буква «а» появляется после букв *б, в, г, д, ж, з, к, л, м, н, п, р, с, т, х, ц, ч, ш, щ*, т. е. после всех согласных, кроме *ф*, например: *пеленать, седлать*.

Уравнение имеет следующий вид: $s_1^a = P_1 V P_3 V P_4 V z_2^n$. Если последняя буква основы *а, е, и, у, ъ, ь, з', л', м', н', п', р', с'*, то выбирается буква «я», например, *роять, нырять*; $s_1^a = P_2 V P_5 V P_6 V P_8 V z_2^c Z_2^n$, $P_8 = z_2^b V z_2^b$.

На выбор класса чередования влияет часть речи, от которой образован глагол, наличие или отсутствие префикса и семантика.

Если глаголы со значениями «наделять признаком», «проявлять признак» образованы от прилагательных, а также если глагол образован от глагола с префиксом и от существительных, имеет значение «действовать с помощью предмета» (*костылять*), «покрывать предметом» (*венчать*), «приводить в состояние» (*восторгать*), «совершать действие, свойственное лицу или животному» (*бродяжничать*), «проявлять свойство» (*малодушничать*), «делать кем-нибудь» (*пролетаризировать*), «совершать действия, свойственные для данного места» (*квартировать*), «помещать в данное место» (*складировать*), «проводить время» (*зимовать*), тогда выбирается $\mu_1 = s_1^a Vs_1^a$: $\mu_1 = z_1^f z_2^n V z_1^n (\omega_2^b V \omega_2^b) V z_1^c (\omega_1^a V \omega_1^b V \omega_1^c V \omega_1^d V \omega_1^e V \omega_1^f V \omega_1^g)$.

$\mu_2 = s_1^e$ выбирается в том случае, если глагол образован от прилагательных со значением «приобретать признак безотносительно к изменчивости состояния» и со значением «приобретать признак лица или животного», образован от существительного $\mu_2 = z_1^p \omega_2^a z_2^n V z_1^c \omega_1^n$.

$\mu_3 = s_1^y$, если глагол со значением «приобретать признак относительно к изменчивости состояния» образован от прилагатель-

ного и если глагол образован от глагола без префикса: $\mu_3 = z_1^1 \omega_2^a z_7^d V z_1^1 z_6^b$.

Рассмотрим третий класс фрагментов с фиксированной буквой «и». В этот класс входят фрагменты *ич*, *ир*, *из*, *и* ____.

В случае присоединения фрагмента суффикса «и ____» происходят следующие чередования в последней букве основы: $г \rightarrow ж$ (*утюг* — *утюжить*), $к \rightarrow ч$ (*чудак* — *чудачить*), $ц \rightarrow ч$ (*кузнец* — *кузнечить*), $х \rightarrow ш$ (*грех* — *грешить*), $ск \rightarrow щ$ (*воск* — *вощить*). Выделяются следующие классы чередований: $t_1^1 V t_1^{ж} = \tau_1$, $(t_1^1 V t_1^{ж}) V t_1^1 = \tau_2$, $t_1^1 V t_1^1 = \tau_3$, $t_2^1 t_1^1 V t_1^1 = \tau_4$; t_2 — предпоследняя буква основы со значениями «с» и «с̄», $\tau_1 V \tau_2 V \tau_3 V \tau_4 = 1$.

Вторая буква фрагмента выбирается в зависимости от того, от какой основы (русской или иностранной) образован глагол. Если основа русская, то «ч» или «____», если основа иностранная, то «р» или «з». На выбор буквы «ч» или «____» влияет наличие или отсутствие следующего фрагмента суффикса «а ____». Если есть следующий фрагмент, то выбирается «ч», если его нет, то «____».

На выбор букв «р» или «з» оказывает влияние ударность фиксированной буквы. Если фиксированная буква под ударением — «р» (*асфальтировать*), без ударения — «з» (*кристаллизировать*): $s_2^r = z_3^p z_4^d$, $s_2^z = z_3^p z_4^h$, $s_2^p = z_3^h z_6^o$, $s_2^z = z_3^h z_6^y$.

Рассмотрим четвертый класс фрагментов с фиксированной буквой «н»: ____ *н*, *ан*. При выборе первой буквы фрагмента выделяются следующие классы чередований: $\mu_1 = s_1^1 V s_1^a$, $\mu^2 = s_1^1$, $\mu_1 V \mu_2 = 1$.

Записываем правила появления того или иного значения внутри каждого класса.

Для $\mu_1 = s_1^1 s_1^a$. Буква «а» выбирается тогда, когда глагол имеет значение «интенсивно однократно совершать действие» (*стегануть*). Буква «____» — если глагол имеет значение «однократно совершать действие» (*клонуть*): $s_1^1 = \omega_3^b$, $s_1^a = \omega_3^a$.

Выбор класса чередования μ зависит от части речи мотивирующего слова. Если глагол образован от глагола, то выбирается μ_1 , если от прилагательного или существительного — μ_2 : $\mu_1 = z_1^r$, $\mu_2 = z_1^c V z_1^п$.

Список литературы: 1. Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П. О математическом описании естественного языка. — Проблемы бионики, 1981, вып. 27, с. 9—13. 2. Бондаренко М. Ф., Шаронова Н. В. Задача фрагментации суффиксов имен существительных. — Проблемы бионики, 1981, вып. 27, с. 79—83. 3. Шабанов-Кушнарченко Ю. П. Об алгебре конечных предикатов. — АСУ и приборы автоматики, 1979, вып. 50, с. 14—20

Поступила в редколлегию 09.10.80.