

УДК 510.62

Н. В. РЯБОВА

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ СУФФИКСАЛЬНОГО
СЛОВООБРАЗОВАНИЯ ОТСУБСТАНТИВНЫХ СУЩЕСТВИТЕЛЬНЫХ
С ОБЩИМ ЗНАЧЕНИЕМ «НОСИТЕЛЬ ПРЕДМЕТНОГО ПРИЗНАКА»**

Опишем один из возможных вариантов построения математической модели внутрикатегориального суффиксального словообразования существительных, имеющих, согласно [1], общее значение «носитель предметного признака». В качестве математического аппарата формального описания используется алгебра

конечных предикатов [2], позволяющая записывать отношения на множестве словообразовательных морфов в виде лингвистических уравнений.

В работе [3] введено понятие текстового отношения L , фрагмента текста Y , переменного вектора смысла фрагмента текста X . Поскольку в данном случае в роли фрагмента текста выступает словообразовательный суффикс (часть слова), будем рассматривать морфологическое отношение $L(X, Y, Z)$, где $X = (x_1, x_2, \dots, x_n)$ и $Z = (z_1, z_2, \dots, z_m)$ — соответственно наборы семантических (смысловых) и грамматических (морфонологических) переменных, связывающих уравнения, которые описывают рассматриваемый фрагмент текста $Y = (y_1, y_2, \dots, y_k)$, с уравнениями, описывающими остальную часть текста («дальний» и «ближний» текст). Таким образом, в роли фрагмента текста выступает словообразовательный суффикс существительного, в роли его смысла — набор грамматических и семантических признаков. С помощью вводимых признаков устанавливается связь между производящей основой и словообразовательным суффиксом, а также между буквами самого суффикса.

В целях конкретизации построения математической модели часть значений грамматических признаков зафиксируем. Это позволит сократить и представить в более компактном виде лингвистические уравнения.

Пусть зафиксированы следующие значения признаков: $\xi_1 = c$ (словообразование), $\xi_2 = c$ (суффиксальный способ словообразования), $\xi_3 = c$ (часть речи производящей основы — существительное), $\xi_4 = c$ (часть речи производного слова — существительное), $\xi_5 = o$ (признак одушевленности со значением «одушевленный»), $\xi_6 = m$ (род производного слова со значением «мужской»). Зафиксировав таким образом признаки, мы тем самым выделили все словообразовательные суффиксы, участвующие в процессах суффиксального внутрикатегориального словообразования одушевленных имен существительных мужского рода.

Введем также переменные семантические признаки и признаки, определяющие влияние производящей основы (грамматические и морфонологические признаки). Необходимо отметить широко распространенную омонимию словообразовательных суффиксов отсубстантивных существительных, наделенных предметным признаком. Как следствие этого явления, имеем большое количество вариантов (морфов) одного суффикса (словообразовательной морфемы), когда выбор конкретного морфа в конечном итоге определяется морфонологическими характеристиками производящей основы. Таковы, например, суффиксы *ан-ин(-ян-ин)* -*чан-ин* (*парижанин, селянин, харьковчанин*), *-чик(-щик)-овщик* (*паркетчик, флейтчик, весовщик*) и т. д.

Перечислим признаки, определяющие влияние производящей основы: z_1 — признак последней буквы основы со значениями всех букв русского алфавита; z_2 — признак предпоследней

буквы основы со значениями всех букв алфавита; z_3 — род производящей основы, со значениями: м — мужской, ж — женский, с — средний $z_3^m \vee z_3^j \vee z_3^c = 1$; z_4 — признак одушевленности производящей основы, со значениями: о — одушевленная, н — неодушевленная $z_4^o \vee z_4^n = 1$; z_5 — признак односложности основы, со значениями: о — односложная, т. е. однослоговая основа, н — нет; z_6 — признак происхождения основы, со значениями: р — русская, и — иностранная основа $z_6^p \vee z_6^i = 1$; z_7 — вид основы, со значениями: т — топонимичная основа, и — имя собственное, о — все остальные виды основ $z_7^t \vee z_7^i \vee z_7^o = 1$; z_8 — признак характеристики основы при опосредованной мотивации, со значениями: г — глагольная характеристика основы, а — адъективная; z_9 — признак твердости (мягкости) последней буквы основы, со значениями: т — твердая, м — мягкая $z_9^t \vee z_9^m = 1$; z_{10} — признак производности основы, со значениями: п — производная основа, н — нет $z_{10}^p \vee z_{10}^n = 1$; z_{11} — признак трансформации основы в процессе словообразования, со значениями: у — основа претерпевает усечение, н — происходит наращение основы, ч — чередование в основе $z_{11}^y \vee z_{11}^n \vee z_{11}^c = 1$.

Введем также систему переменных семантических признаков таким образом, чтобы каждому набору значений признаков можно было поставить в соответствие один словообразовательный суффикс в его конкретной семантической роли. Перечислим эти признаки: x_1 — признак классификации представителей живого мира, со значениями: л — лица, ж — другие представители живого мира (животные, птицы, рыбы и др.) $x_1^l \vee x_1^j = 1$; x_2 — тип характеристики лиц, со значениями: л — личностные характеристики, п — профессионально-функциональные характеристики $x_2^l \vee x_2^p = 1$; x_3 — признак классификации личностных характеристик, со значениями: т — топонимические характеристики (раса, национальность, место жительства или нахождения), в — характерные особенности внешнего облика $x_3^t \vee x_3^v = 1$; x_4 — признак классификации профессионально-функциональных характеристик, со значениями: о — характеристика по объекту деятельности лица (*сапожник, паркетчик, флейтист*), м — характеристика по месту деятельности лица (*тюремщик, дворник, типографщик*), с — характеристика по сфере деятельности лица (*композитор, навигатор, революционер*), к — характеристика по характерному свойству, качеству (*сплетник, завистник, причудник*), д — характеристика лица по выполняемому действию (*помощник, завистник, взяточник*), $x_4^o \vee x_4^m \vee x_4^c \vee x_4^k \vee x_4^d = 1$; x_5 — признак характеристики объекта деятельности лица, со значениями: д — лицо, действующее с помощью предмета, названного в основе (*лебедчик, бетонщик, гитарист*), п — лицо, производящее то, что названо в основе, или лицо, конечным продуктом деятельности которого является предмет,

названный в основе (*виноградарь, бочар, сапожник*) $x_5^n \vee x_5^n = 1$; x_6 — признак стилевой характеристики производного слова, со значениями: к — книжный, с — специальный, р — разговорный, о — областной, у — устаревший, п — просторечье.

Следует отметить, что при разработке системы семантических признаков использовались данные традиционной грамматики [1], а также новейшие разработки в области прикладной лингвистики [4, 5]. Кроме того, мы руководствовались следующими принципами: семантические признаки вводились как классы эквивалентности семантических оттенков. При объединении выявленных оттенков в признаки использовался метод логического противопоставления, т. е. в один признак могут компоноваться только те оттенки, которые не встречаются одновременно ни в одной текстовой ситуации. При выявлении оттенков использовалась конструкция, называемая нами «*пробный текст*», в которой устанавливается однозначная зависимость между производящей основой X , словообразовательным суффиксом Y и производным словом XU .

Приведем примеры выявления оттенков с помощью пробных текстов (ПТ). Оттенок «характеристика по месту деятельности лица» выявляется с помощью ПТ вида « XU обычно находится при X » (*штабник, дворник, тюремщик*). Оттенок «характеристика лица по его свойствам, качеству» выявляется с помощью ПТ « XU обладает характерным качеством X ». Оттенок «характеристика лица, действующего с помощью предмета», выявляется в ПТ « XU действует с помощью X » (*завистник, причудник; лебедчик, паркетчик, гитарист*).

Задача состоит в математическом описании отношения $L(X, Y, Z)$. В силу принципа однозначности [3] это отношение задает функцию, которую возможно описать в явном виде: $y_i^j = F_{\sigma}(x_1, x_2, x_3, x_4, x_5, x_6, z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9, z_{10}, z_{11})$, где $\sigma \in \{a, e, и, o, я, в, к, н, р, с, т, ц, ч, ш, ь\}$ y_i — буквы суффикса. Математическое описание модели суффиксального словообразования сводится к формированию суффикса $Y = y_1, y_2, y_3, y_4, y_5, y_6$ (максимальное количество букв суффикса в данном случае равно 6) в зависимости от набора грамматических и семантических признаков.

Все рассматриваемые словообразовательные суффиксы целесообразно представить в виде конструкции, называемой нами *лингвистическим регистром суффиксов* (таблица). Суффиксы представляются в регистре в сегментированной форме. Для одного сегмента выделяется три позиции s_1, s_2, s_3 . Структура сегмента (гласная), (согласная), (пробел или «ь») обусловлена тем, что подавляющее большинство суффиксов имеет такую структуру или легко членятся на сегменты такого типа.

Кроме перечисленных выше признаков введем также переменную t — номер сегмента в регистре, со значениями 1, 2, 3: $t^1 \vee$

t^1			t^2			t^3		
s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3
[[[[н	[и	к	[
[[[[ч	[и	к	[
[[[[щ	[и	к	[
о	в	[[щ	[и	к	[
[[[и	с	[[т	[
[[[[[[а	ч	[
[[[[[[и	ч	[
о	в	[[[[и	ч	[
[[[а	н	[и	ч	[
[[[я	н	[и	н	[
[[[я	н	[и	н	[
н	ч	[я	н	[и	н	[
[т	[[[[и	н	[
[[[е	н	[е	ц	[
о	в	[а	н	[е	ц	[
[[[[[[е	ц	[
[[[[[[а	н	[
[[[[[[а	р	[
[[[[[[а	р	ь
[[[а	т	[я	р	[
[[[и	т	[о	р	[
[[[[т	[о	р	[

$\vee t^2 \vee t^3 = 1$. Для удобства описания введем также именные предикаты: ГЛ (z), принимающий значение 1 только тогда, когда z — гласная буква, и равный нулю в остальных случаях; СГ (z), принимающий значение 1 для всех согласных букв и равный нулю в остальных случаях.

Рассмотрим первую s_1 , вторую s_2 и третью s_3 буквы каждого сегмента. Области значений этих переменных задаются уравнениями

$$(s_1^o \vee s_1^n \vee s_1^-) t^1 \vee (s_1^a \vee s_1^r \vee s_1^h \vee s_1^c \vee s_1^-) t_2 \vee (s_1^a \vee s_1^r \vee s_1^c \vee s_1^h \vee s_1^o \vee s_2^-) t^3 = 1, \quad (1)$$

$$(s_2^- \vee s_2^b \vee s_2^q \vee s_2^r) t^1 \vee (s_2^h \vee s_2^q \vee s_2^m \vee s_2^c \vee s_2^r \vee s_2^-) t^2 \vee (s_2^k \vee s_2^q \vee s_2^h \vee s_2^p \vee s_2^h \vee s_2^r) t^3 = 1, \quad (2)$$

$$(s_3^- t^1 \vee s_3^- t^2 \vee (s_3^- \vee s_3^b)) t^3 = 1. \quad (3)$$

Запишем в явном виде уравнения для букв каждого сегмента, а затем, с учетом декомпозиции модели, легко можно записать общие уравнения для всего регистра. Рассмотрим первый сегмент. Первая буква s_1 принимает значения {о, и, —}. s_1 принимает значение о при образовании наименований лиц с типом

характеристики «личностный» (ЛХ) от именных основ (*Иванович*), а также при образовании наименований лиц с типом характеристики «профессионально-функциональный» (ПФХ) от односложных основ, или русских основ, для которых свойственна интерфиксация или интерференция в процессе словообразования (*весовщик, кайлощик*):

$$s_1^0 = x_1^l (x_2^l x_3^l z_7^0 \vee x_2^p (z_5^0 \vee z_{11}^n)) = p_1. \quad (4)$$

Если образуется наименование лица с ЛХ от односложной основы-топонима или с ПФХ и характеристикой по месту нахождения лица, то первая буква первого сегмента принимает значение и (*островитянин, псковитянин*):

$$s_1^n = x_1^l (x_2^l x_3^t z_5^0 z_7^t \vee x_2^p x_4^m) = p_2. \quad (5)$$

В остальных случаях первая буква сегмента отсутствует, принимает значение —:

$$\bar{s}_1 = x_1^l (x_2^t (z_7^t (z_1^c z_2^c \vee x_6^b) \vee z_7^0 x_3^b) = p_3. \quad (6)$$

Для второй буквы первого сегмента область определения {в, ч, т, —}. Условия появления на втором месте буквы «в» те же, что и для s_1^0 . Поэтому уравнение будет иметь вид $s_2^v = p_1$ (7).

Рассмотрим условия появления s_2 . Это становится возможным, когда от основ-топонимов образуются названия жителей, при условии, что основа неодносложная, заканчивающаяся на -ск-, -св-, -ав- (*минчанин, харьковчанин, полтавчанин*). Запишем уравнение:

$$s_2^a = x_1^l x_2^l x_3^t z_5^h (z_2^k z_1^c \vee z_2^0 z_1^b \vee z_2^a z_1^b) = p_4. \quad (8)$$

Условия появления на втором месте т такие же, как и для s_1^n . Поэтому уравнение запишется так: $s_2^t = p_2$ (9). На втором месте появляется — при тех же условиях, что и на первом месте первого сегмента: $s_2 = p_3$ (10). Третья буква первого сегмента всегда принимает значение —: $s_3 = 1$.

Рассмотрим второй сегмент. Для первой буквы область определения $s_1 \in \{и, а, я, е, —\}$. На первом месте появляется е при образовании лиц с ПФХ, мотивированных отглагольными существительными на -ение (*сопротивление, управление*):

$$s_1^e = x_1^l x_2^p z_8^t z_{11}^y = p_5. \quad (11)$$

s_1 принимает значение и, если образуются названия лиц с ПФХ и характеризующиеся по объекту деятельности лица (*пианист, артиллерист*), а также при образовании от иностранных основ на -ация/-иция (оканчивающихся на мягкую согласную) наименований лиц, характеризующихся по сфере деятельности (*экспедитор*). Уравнение будет иметь вид

$$s_1^и = x_1^l x_2^p (x_4^0 x_5^h \vee x_4^c) \vee x_1^l x_2^p x_4^c z_6^h z_{11}^y z_9^m. \quad (12)$$

Обозначим

$$p_6 = x_1^{\text{л}} x_2^{\text{п}} (x_4^0 x_5^{\text{д}} \vee x_4^{\text{с}}), \quad p_7 = x_1^{\text{л}} x_2^{\text{п}} x_4^{\text{с}} z_6^{\text{н}} z_1^{\text{у}} (z_9^{\text{м}} \vee z_9^{\text{т}}).$$

s_1 принимает значение а, если образуются названия жителей от основ-топонимов (выполняется условие p_4), или названия лиц с ПФХ, от иностранных основ (выполняется условие p_7), оканчивающихся на твердую согласную (*полтавчанин, комбинатор*): $s_1^{\text{а}} = p_4 \vee p_7 z_9^{\text{т}}$ (13). s_1 принимает значение я, если выполняется условие p_2 (*островитянин*): $s_1^{\text{я}} = p_2$ (14). На первом месте второго сегмента будет $_$, если образуется наименование лица с ПФХ, характеризующееся по объекту его деятельности или месту, или характерному свойству, качеству, действию (*сапожник, паркетчик, путник, фокусник*); если образуется название лица с личностной характеристикой и стилевым оттенком «просторечье»:

$$s_1^{_} = x_1^{\text{л}} x_2^{\text{п}} (x_4^0 \vee x_4^{\text{м}} \vee x_4^{\text{д}} \vee x_4^{\text{к}}) z_5^{\text{н}} z_6^{\text{п}} \vee x_1^{\text{л}} x_2^{\text{п}} (x_6^{\text{п}} \vee x_6^{\text{н}}). \quad (15)$$

Рассмотрим вторую букву второго сегмента, область ее определения $s_2 \in \{н, ч, т, с, щ, _ \}$. s_2 принимает значение н, если производное слово имеет семантику «лицо с ПФХ, характеризующееся по объекту его деятельности, месту либо характерному свойству или действию» (*помощник, путник, завистник*). Запишем уравнение: $x_1^{\text{л}} x_2^{\text{п}} x_4^0 = p_8 = s_2^{\text{н}}$ (16). Если s_2 принимает значение щ, то образуется название лица с ПФХ, характеризующееся по объекту деятельности либо месту и имеющее разговорную стилевую характеристику: $x_1^{\text{л}} x_2^{\text{п}} (x_4^0 \vee x_4^{\text{м}}) x_6^{\text{в}} = s_2^{\text{щ}}$ (17). s_2 принимает значение с при тех же условиях, что и $s_1^{_}$ в этом же сегменте. Поэтому уравнение имеет вид $s_2^{\text{с}} = p_6, s_2^{\text{т}} = p_7$ (18). В остальных случаях s_2 принимает значение $_$:

$$s_2^{_} = x_1^{\text{л}} x_2^{\text{п}} (x_6^{\text{п}} \vee x_6^{\text{н}}) x_3^{\text{н}} \vee p_8; \quad p_8 = x_3^{\text{т}} z_7^{\text{л}} (z_1^{\text{с}} z_2^{\text{с}} \vee z_1^0 z_2^{\text{а}}). \quad (19)$$

Третья буква второго сегмента всегда $_1 s_3 _1$.

Рассмотрим третий сегмент. Первая буква s_1 имеет область определения {и, а, я, е, о, $_$ }. Уравнения запишем с учетом полученных условий появления определенных букв регистра. На первом месте будет и, если выполняется одно из условий p_2, p_4, p_8 , если от именных основ образуются названия лиц с ЛХ (*Иванович*) либо при образовании наименований лиц с ПФХ, характеризующихся по объекту деятельности (*сапожник*), характерному действию или свойству:

$$s_3^{\text{и}} = p_2 \vee p_4 \vee p_8 \vee x_1^{\text{л}} x_2^{\text{п}} z_7^{\text{н}} \vee x_1^{\text{л}} x_2^{\text{п}} (x_4^0 \vee x_4^{\text{д}} \vee x_4^{\text{к}}). \quad (20)$$

s_1 принимает значение а, если производное слово имеет семантику «лицо с ЛХ, характеризующей особенности внешнего облика и несущее просторечную стилевую характеристику (*брюхач*); либо лицо с ПФХ, характеризующееся по объекту деятельности, от основ, оканчивающихся на к (*скрипка — скрипач*)

и подвергающихся усечению; либо, при значении признака x_4^o , лицо, производящее то, что названо в основе (*виноградарь*), либо лицо по месту деятельности (*свинарь*):

$$s_1^a = p_9 \vee p_{10} \vee x_1^l x_2^l x_4^o z_1^k z_{11}^y, \quad (21)$$

где

$$p_9 = x_1^l x_2^l x_3^b x_6^b, \quad p_{10} = x_1^l x_2^l (x_4^o x_5^l \vee x_4^m).$$

s_1^a имеет место, если образуется наименование лица с ПФХ, характеризующее лицо по объекту деятельности (*гусляр, столяр*), от основ, оканчивающихся на мягкую согласную:

$$s_1^a = x_1^l x_2^l x_4^o z_9^m. \quad (22)$$

На первом месте будет е, если производное слово имеет семантику «лицо с ЛХ, образованное от основ-топонимов и обозначающее жителя того места, которое названо в основе (*ленинградец*)», а также если выполняется условие p_5 :

$$s_1^e = x_1^l x_2^l x_3^t z_7^t \vee p_5. \quad (23)$$

s_1^o имеет место при выполнении условия p_7 (*композитор*):

$$s_1^o = p_7 \quad (24), \quad s_1^- = p_6. \quad (25)$$

Рассмотрим вторую букву третьего сегмента $s_2 \in \{к, т, ч, н, ц, р\}$. s_2^k имеем, когда образуется название лица с ПФХ, характеризующееся по объекту деятельности, либо месту, либо характерному действию или свойству:

$$s_2^k = x_1^l x_2^l (x_4^o \vee x_4^m \vee x_4^l) z_6^b. \quad (26)$$

s_2^t имеет место при выполнении условия p_6 : $s_2^t = p_6$ (27). Если выполняется условие p_8 или p_9 , то на втором месте будет ч: $s_2^c = p_8 \vee p_9$ (28). s_2^h имеет место при выполнении одного из условий p_2 или p_4 : $s_2^h = p_2 \vee p_4$ (29). На втором месте третьего сегмента будет ц, если выполняется условие p_4 или p_{11} : $s_2^u = p_4 \vee p_{11}$ (30). s_2 принимает значение р, когда выполняется уравнение $s_2^o = p_7 \vee p_{10}$ (31).

Для третьей буквы третьего сегмента область определения $s_3 \in \{_, б\}$. Уравнения запишутся так:

$$s_3^b = p_{11}, \quad s_3^- \vee s_3^b = 1. \quad (32)$$

Список литературы: 1. *Русская грамматика*. — М.: Наука, 1980. — 783 с. 2. *Шабанов-Кушнаренко Ю. П.* Об алгебре конечных предикатов. — АСУ и приборы автоматики, 1979, вып. 50, с. 14—20. 3. *Бондаренко М. Ф., Шабанов-Кушнаренко Ю. П.* О математическом описании естественного языка. — Пробл. библики, 1981, вып. 27, с. 9—13. 4. *Кубрякова Е. С.* Типы языковых значений. — М.: Наука, 1981. — 200 с. 5. *Милославский И. Г.* Вопросы словообразовательного синтеза. — М.: Изд-во Моск. ун-та. — 246 с.

Поступила в редколлегию 02.03.84.