

УДК 519.7



МОДЕЛЬ УСТНОЙ РЕЧИ

Н. Е. Русакова¹¹ХНУРЭ, г. Харьков, Украина

Статья посвящена разработке методов формального описания структуры естественного языка с помощью алгебры предикатов и алгебры предикатных операций, а также аппарата логических сетей; рассматривается модель устного языка на примере звукового варианта склонения полных непряжательных имен прилагательных. Актуальность этой темы определяется перспективностью применения полученных методов для разработки систем общения с компьютером на естественном языке.

АЛГЕБРА ПРЕДИКАТОВ, ЛОГИЧЕСКАЯ СЕТЬ, ЕСТЕСТВЕННЫЙ ЯЗЫК, ПАРАДИГМАТИЧЕСКАЯ ТАБЛИЦА

Введение

В настоящее время все более совершенствуется компьютерная обработка естественного языка. Лингвистические труды содержат неформализованные описания морфологических явлений, которые непросто перевести на математический язык. Эти описания адресованы для понимания человеку. Математические же описания морфологии языка адресованы вычислительной машине, они должны быть понятны ей. В качестве объектов математического описания будем рассматривать тексты, то есть представленные в виде строчек знаковые речевые сообщения, которыми люди обмениваются в процессе интеллектуальной деятельности. Важно то, что далеко не любая строчка, составленная из знаков, может быть признана текстом [1].

Целью работы является развитие формальных средств описания естественного языка — разработка методов алгебрологического аппарата логических сетей, а также их применение для формализации морфологических структур на примере построения логической сети для звукового варианта склонения полных непряжательных имен прилагательных.

1. Звуковой вариант склонения полных непряжательных имен прилагательных

В статье [2] для примера рассматривается, формально описывается на языке алгебры предикатов и доводится до воспроизведения на логической сети механизм склонения полных непряжательных имен прилагательных. Здесь предлагается письменный вариант, который более прост по сравнению со звуковым. Это обусловлено тем, что звуковая речь более древняя, чем письменная и вполне может обходиться без письменной, в то время как письменная без звуковой обойтись не может. Звуковой вариант предельно регулярен, правила его достаточно просты, этот механизм языка уникален, благодаря своей простоте и привычности для любого человека. Исследование письменного варианта склонения имён прилагательных невозможно без привлечения факторов звуковой речи (ударность-безударность, твердость-мягкость). Исследование же звукового варианта возможно без опоры на элементы письменного языка. Предшественники до

сих пор поступали наоборот, рассматривали письменный вариант, а звуковым вообще не интересовались. Грамматика русского языка как академическая, так и школьная, строятся по следующему принципу: вначале немного говорится о звуковом варианте речи, а затем всё описание механизма языка ведется для варианта письменной речи.

Для рассмотрения выбираем полные непряжательные имена прилагательные. Достаточно основательная проработка этого раздела производится в языкознании. Притяжательные прилагательные сложнее, поскольку имеют беглость, например *лисий-лисья*, и несколько типов склонения. В полных непряжательных прилагательных всего лишь один тип склонения, имеющий семь четких вариантов склонения. Склонение полных непряжательных имен прилагательных образуют в грамматике компактный раздел морфологии языка [3]. Склонение происходит с помощью парадигматической таблицы, которая для звукового варианта представлена табл. 1.

Таблица 1

Падеж, x_3	Число, x_2				Множественное
	Единственное				
	Род, x_1				
	мужской	женский	средний		
именительный	<i>слабый</i> _	<i>слабайя</i>	<i>слабойе</i>	<i>слабые</i>	
родительный	<i>слабово</i>	<i>слабой</i> _	<i>слабово</i>	<i>слабых</i> _	
дательный	слабому	слабой_	слабому	<i>слабым</i>	
винительный	<i>слабый</i> _	<i>слабую</i>	<i>слабойе</i>	<i>слабые</i>	
	<i>слабово</i>			<i>слабых</i> _	
творительный	<i>слабым</i> _	<i>слабой</i> _	<i>слабым</i> _	<i>слабыми</i>	
		<i>слабую</i>			
предложный	<i>слабом</i>	<i>слабой</i>	<i>слабом</i>	<i>слабых</i>	

Вводим необходимые для формального описания предметные переменные: x_4 — признак одушевленности формы слова со значениями O — одушевленный, H — неодушевленный; x_1 — род формы слова со значениями M — мужской, J — женский, C — средний; x_2 — число формы слова со значениями E — единственное, M — множественное; x_3 — падеж формы слова со значениями I — именительный, P — родительный, D — дательный, V — винительный, T — творительный, $Π$ — предложный.

В ячейках таблицы указаны формы слова в зависимости от их рода x_1 , числа x_2 и падежа x_3 . Таблица составлена для конкретного слова *слабый*.

Для полных имен прилагательных имеются типы склонений [3]. Ниже приведены таблицы окончаний (табл. 2-7) для всех типов склонения с указанием их номеров, которые сопровождаются примерами слов, склоняемых по данному типу.

В звуковом варианте получилось меньше типов склонения – не 7, а 6. Исчез 4-й тип склонения. Если ударная основа оканчивается на звуки *ж*, *ш*, то слово склоняется по 3-му типу. Если основа оканчивается на *ч*, *щ*, то слово склоняется по 5-му типу. В парадигматических таблицах определяем для каждого склонения переменные: y_1 – последний звук основы слова, y_2 – признак ударности-безударности, y_3 – мягкости-твердости, u – первичный класс последнего звука основы слова.

Таблица 2

1 – <i>слабый</i>			
<i>ый_</i>	<i>айя</i>	<i>ойе</i>	<i>ыйе</i>
<i>ово</i>	<i>ой_</i>	<i>ово</i>	<i>ых_</i>
<i>ому</i>	<i>ой_</i>	<i>ому</i>	<i>ыми</i>
<i>ый_</i>	<i>уйю</i>	<i>ойе</i>	<i>ыйе</i>
<i>ово</i>			<i>ых_</i>
<i>ым_</i>	<i>ой_</i>	<i>ым_</i>	<i>ыми</i>
	<i>ойю</i>		
<i>ом_</i>	<i>ой_</i>	<i>ом_</i>	<i>ых_</i>
$y_1 \in \{\text{б, в, д, з, л, м, н, т, н, р, с}\};$ $y_2 = \text{у}; y_3 = \text{т}; u = \text{б}$			

Таблица 3

2 – <i>чужой</i>			
<i>ой_</i>	<i>айя</i>	<i>ойе</i>	<i>ыйе</i>
<i>ово</i>	<i>ой_</i>	<i>ово</i>	<i>ых_</i>
<i>ому</i>	<i>ой_</i>	<i>ому</i>	<i>ыми</i>
<i>ой_</i>	<i>уйю</i>	<i>ойе</i>	<i>ыйе</i>
<i>ово</i>			<i>ых_</i>
<i>ым_</i>	<i>ой_</i>	<i>ым_</i>	<i>ыми</i>
	<i>ойю</i>		
<i>ом_</i>	<i>ой_</i>	<i>ом_</i>	<i>ых_</i>
$y_1 \in \{\text{б, в, д, з, м, н, т, н, р, с, ж, ш}\};$ $y_2 = \text{б}; y_3 = \text{т}; u = \text{ж}$			

Таблица 4

3 – <i>горячий</i>			
<i>ий_</i>	<i>айя</i>	<i>ейе</i>	<i>ийе</i>
<i>ево</i>	<i>ей_</i>	<i>ево</i>	<i>их_</i>
<i>ему</i>	<i>ей_</i>	<i>ему</i>	<i>ими</i>
<i>ий_</i>	<i>юйю</i>	<i>ейе</i>	<i>ыйе</i>
<i>ево</i>			<i>их_</i>
<i>им_</i>	<i>ей_</i>	<i>им_</i>	<i>ими</i>
	<i>ейю_</i>		
<i>ем</i>	<i>ей_</i>	<i>ем_</i>	<i>их_</i>
$y_1 \in \{\text{н, р, с, ч, щ, _}\}; y_2 = \text{у}; y_3 = \text{м}; u = \text{ч}$			

Рассмотрим формирование первичных, вторичных и третичных классов последнего звука основы слова.

Первичные классы:

$$\begin{aligned}
 y_1^{\text{б}} \vee y_1^{\text{в}} \vee y_1^{\text{д}} \vee y_1^{\text{з}} \vee y_1^{\text{л}} \vee y_1^{\text{м}} \vee y_1^{\text{н}} \vee y_1^{\text{т}} &= u^{\text{б}}; \\
 y_1^{\text{ж}} \vee y_1^{\text{ш}} \vee y_1^{\text{ч}} &= u^{\text{ж}}; \\
 y_1^{\text{жс}} \vee y_1^{\text{шс}} &= u^{\text{жс}}; \\
 y_1^{\text{ш}} &= u^{\text{ш}}; \\
 y_1^{\text{ч}} \vee y_1^{\text{щ}} \vee y_1^{\text{ц}} &= u^{\text{ч}}; \\
 y_1^{\text{ц}} \vee y_1^{\text{к}} \vee y_1^{\text{х}} &= u^{\text{ц}}.
 \end{aligned}$$

Вторичные классы:

$$\begin{aligned}
 u^{\text{жс}} &= v^{\text{жс}}; \\
 u^{\text{жс}} \vee u^{\text{ш}} &= v^{\text{ш}}; \\
 u^{\text{б}} \vee u^{\text{ж}} &= v^{\text{б}}; \\
 u^{\text{ж}} \vee u^{\text{ч}} &= v^{\text{ч}}; \\
 u^{\text{ц}} &= v^{\text{ц}}.
 \end{aligned}$$

Таблица 5

4 – <i>строгий</i>			
<i>ий_</i>	<i>айя</i>	<i>ойе</i>	<i>ийе</i>
<i>ево</i>	<i>ой_</i>	<i>ово</i>	<i>их_</i>
<i>ому</i>	<i>ой_</i>	<i>ому</i>	<i>ими</i>
<i>ий_</i>	<i>уйю</i>	<i>ойе</i>	<i>ийе</i>
<i>его</i>			<i>их_</i>
<i>им_</i>	<i>ой_</i>	<i>им_</i>	<i>ими</i>
	<i>ойю</i>		
<i>ом_</i>	<i>ой_</i>	<i>ом_</i>	<i>их_</i>
$y_1 \in \{\text{з, к, х}\}; y_2 = \text{у}; y_3 = \text{м}; u = \text{з}$			

Таблица 6

5 – <i>тугой</i>			
<i>ой_</i>	<i>айя</i>	<i>ойе</i>	<i>ийе</i>
<i>ово</i>	<i>ой_</i>	<i>ово</i>	<i>их_</i>
<i>ому</i>	<i>ой_</i>	<i>ому</i>	<i>ими</i>
<i>ой_</i>	<i>уйю</i>	<i>ойе</i>	<i>ийе</i>
<i>ово</i>			<i>их_</i>
<i>им_</i>	<i>ой_</i>	<i>им_</i>	<i>ими</i>
	<i>ойю</i>		
<i>ом_</i>	<i>ой_</i>	<i>ом_</i>	<i>их_</i>
$y_1 \in \{\text{з, к, х}\}; y_2 = \text{б}; y_3 = \text{т}; u = \text{з}$			

Таблица 7

7 – <i>куцый</i>			
<i>ый_</i>	<i>айя</i>	<i>эйе</i>	<i>ыйе</i>
<i>эво</i>	<i>эй_</i>	<i>эво</i>	<i>ых_</i>
<i>эму</i>	<i>эй_</i>	<i>эму</i>	<i>ыми</i>
<i>ый_</i>	<i>уйю</i>	<i>эйе</i>	<i>ыйе</i>
<i>эво</i>			<i>ых_</i>
<i>ым_</i>	<i>эй_</i>	<i>ым_</i>	<i>ыми</i>
	<i>эйю</i>		
<i>эм_</i>	<i>эй_</i>	<i>эм_</i>	<i>ых_</i>
$y_1 \in \{\text{ж, щ, ц}\}; y_2 = \text{у}; y_3 = \text{т}; u = \text{ц}$			

Третичные классы:

$$\begin{aligned}v^{\bar{o}} &= w^{\bar{o}}; \\v^{\bar{o}} \vee v^{\bar{ж}} &= w^{\bar{ж}}; \\v^u &= w^u; \\v^e &= w^e; \\v^y &= w^y.\end{aligned}$$

Производим нумерацию ячеек парадигматической таблицы. Полученные номера ячеек в дальнейшем будут использованы для обозначения влияний контекста. Некоторые из влияний контекста получат более одного имени. Однако от этого недостатка мы впоследствии постараемся избавиться. Для каждой ячейки парадигматической таблицы вводим свой номер r (табл.8).

Таблица 8

$x_2 \backslash x_1$		ME	$ЖЕ$	CE	$-M$	
x_3	I	1	8	15	21	
	P	2	9	16	22	
	$Д$	3	10	17	23	
	B	4	11	18	24	x_4
		5			25	
	T	6	12	19	26	x_5
			13			
	Π	7	14	20	27	

Выразим номера r ячеек парадигматической таблицы через признаки $x_1 - x_5$.

Преобразование блока контекста слова имеет следующий вид:

$$\begin{aligned}x_1^M x_2^E x_3^I &= r^1; x_1^M x_2^E x_3^P = r^2; x_1^M x_2^E x_3^Д = r^3; \\x_1^M x_2^E x_3^B &= r^4; x_1^M x_2^E x_3^O = r^5; x_1^M x_2^E x_3^T = r^6; \\x_1^M x_2^E x_3^П &= r^7; x_1^Ж x_2^E x_3^I = r^8; x_1^Ж x_2^E x_3^P = r^9; \\x_1^Ж x_2^E x_3^Д &= r^{10}; x_1^Ж x_2^E x_3^B = r^{11}; x_1^Ж x_2^E x_3^T = r^{12}; \\x_1^Ж x_2^E x_3^A &= r^{13}; x_1^Ж x_2^E x_3^П = r^{14}; x_1^C x_2^E x_3^I = r^{15}; \\x_1^C x_2^E x_3^P &= r^{16}; x_1^C x_2^E x_3^Д = r^{17}; x_1^C x_2^E x_3^B = r^{18}; \\x_1^C x_2^E x_3^T &= r^{19}; x_1^C x_2^E x_3^П = r^{20}; x_2^M x_3^I = r^{21}; \\x_2^M x_3^P &= r^{22}; x_2^M x_3^Д = r^{23}; x_2^M x_3^B x_4^H = r^{24}; \\x_2^M x_3^B x_4^O &= r^{25}; x_2^M x_3^T = r^{26}; x_2^M x_3^П = r^{27}.\end{aligned}$$

Составляем таблицу типов склонений в зависимости от признаков y_2, y_3 , где s — тип склонения.

Таблица 9

w	$УТ$	$УМ$	$БТ$	$БМ$
$\bar{b}$	1	—	—	—
$\bar{ж}$	—	—	2	—
$\bar{ч}$	—	3	—	—
$\bar{ц}$	6	—	—	—
$\bar{з}$	—	4	5	—

В итоге получаем блок слова, имеющий вид:

$$\begin{aligned}w^{\bar{o}} y^Y y^T &= s^1; w^{\bar{ж}} y^B y^T = s^2; w^u y^Y y^M = s^3; \\w^e y^Y y^M &= s^4; w^e y^B y^T = s^5; w^y y^Y y^T = s^6,\end{aligned}$$

который зависит от последнего звука основы слова и признаков ударности-безударности и твердости-мягкости.

Связь внешних переменных представлена в виде сети (рис. 1).

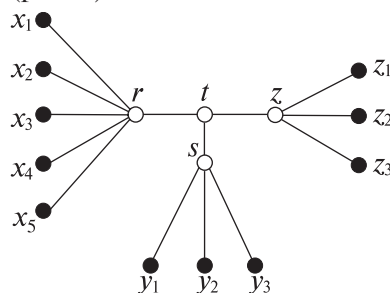


Рис. 1. Логическая сеть

2. Классификация первых звуков окончания

Механизм языка иерархичен, следовательно, исследование языка необходимо вести в определенном порядке [4]. Механизм устной речи, появляющийся при склонении имён прилагательных, можно описать, не привлекая письменную речь, в то время как для изучения письменной речи нужны признаки смягчения и ударности звука. Если изучать сначала письменный вариант склонения имён прилагательных, то создаётся впечатление, что буквы e и o противопоставлены друг другу: *ом-ем, ого-его* и т.д. Но если рассмотреть сначала устный вариант, то становится ясным, что звук e противопоставлен звуку $\bar{э}$, а звук o — звуку $\bar{ё}$. Звук $\bar{ё}$ из письменного варианта склонения прилагательных вовсе исключен. Но необходимость в звуке $\bar{ё}$ возникает при рассмотрении потенциального варианта письменной речи для склонения прилагательных. После изучения устного варианта речи легче понять её письменный вариант и глубже его изучить. Гласные звуки явно соединяются в пары родственных звуков, которые противопоставлены друг другу по признаку y_3 (твёрдый-мягкий). Появилась необходимость выделить пять звуков. Обозначим их переменной L со значениями $a, o, y, ы, \bar{э}$. Классификация этих звуков представлена в табл. 10.

Таблица 10

y_3	a	o	y	$ы$	$\bar{э}$	L
T	a	o	y	$ы$	$\bar{э}$	z_1
M	$\bar{я}$	$\bar{ё}$	$\bar{ы}$	$\bar{и}$	$\bar{е}$	

В соответствии с классификацией первого звука окончания можно произвести классификацию окончания в целом (табл. 11).

Таблица 11

T	$\bar{а}йя$	$уйю$	$ыми$	$\bar{э}йе$	$ойе$	$ыйе$	$\bar{ый}_-$	$ой_-$	$\bar{эй}_-$
M	$\bar{я}йя$	$\bar{ю}йю$	$\bar{ими}$	$\bar{е}йе$	$\bar{ё}йе$	$\bar{ийе}$	$\bar{ий}_-$	$\bar{ёй}_-$	$\bar{ей}_-$
T	$\bar{э}во$	$\bar{ово}$	$\bar{э}му$	$\bar{ому}$	$\bar{ым}_-$	$\bar{эм}_-$	$\bar{ом}_-$	$\bar{ых}_-$	
M	$\bar{е}во$	$\bar{ё}во$	$\bar{ему}$	$\bar{ё}му$	$\bar{им}_-$	$\bar{ем}_-$	$\bar{ём}_-$	$\bar{их}_-$	
T	$\bar{эйю}$	$\bar{ойю}$							
M	$\bar{ейю}$	$\bar{ёйю}$							

Окончание z расчленяется естественным образом на левую (z_l) и правую (z_r) части. Основанием для разделения окончания на левую и правую части является то, что правая часть окончания определяется исключительно влиянием контекста (признаки $x_1 - x_5$), тогда как левая часть окончания определяется, кроме того, еще и типом склонения слова s . Левая часть окончания состоит из первой буквы окончания. Правая же часть окончания состоит из двух компонентов — второй и третьей букв окончания.

Теперь рассмотрим зависимость правой части окончания от первого звука окончания (табл. 12), а также расчленением правую часть окончания на звуки (табл. 13).

Таблица 12

z_1	йа	йю	ми	йе	й_	во	му	м_	х_
а	айя	—	—	—	—	—	—	—	—
у	—	уйю	—	—	—	—	—	—	—
ы	—	—	ыми	ыйе	ый_	—	—	ым_	ых_
э	—	эйю	—	эйе	эй_	эво	эму	эм_	—
о	—	ойю	—	ойе	ой_	ово	ому	ом_	—

Таблица 13

$z_2 z_3$	я	ю	и	е	—	о	у
й	йя	йю	—	йе	й_	—	—
м	—	—	ми	—	м_	—	му
в	—	—	—	—	—	во	—
х	—	—	—	—	х_	—	—

Проведя классификацию звуков окончания и рассмотрев связь звуков между собой, можно записать в формульном виде эту связь.

$$\begin{aligned}
 z^1 &= z_1^a z_2^y z_3^y; \quad z^2 = z_1^a z_2^y z_3^y; \\
 z^3 &= z_1^y z_2^y z_3^y; \quad z^4 = z_1^y z_2^y z_3^y; \\
 z^5 &= z_1^b z_2^m z_3^u; \quad z^6 = z_1^b z_2^m z_3^u; \\
 z^7 &= z_1^z z_2^y z_3^e; \quad z^8 = z_1^z z_2^y z_3^e; \\
 z^9 &= z_1^o z_2^y z_3^e; \quad z^{10} = z_1^b z_2^y z_3^e; \\
 z^{11} &= z_1^u z_2^y z_3^e; \quad z^{12} = z_1^b z_2^y z_3^e; \\
 z^{13} &= z_1^u z_2^y z_3^e; \quad z^{14} = z_1^z z_2^y z_3^e; \\
 z^{15} &= z_1^e z_2^y z_3^e; \quad z^{16} = z_1^o z_2^y z_3^e; \\
 z^{17} &= z_1^z z_2^y z_3^e; \quad z^{18} = z_1^e z_2^y z_3^e; \\
 z^{19} &= z_1^o z_2^y z_3^e; \quad z^{20} = z_1^z z_2^y z_3^e; \\
 z^{21} &= z_1^e z_2^y z_3^e; \quad z^{22} = z_1^o z_2^y z_3^e; \\
 z^{23} &= z_1^b z_2^y z_3^e; \quad z^{24} = z_1^u z_2^y z_3^e; \\
 z^{25} &= z_1^z z_2^y z_3^e; \quad z^{26} = z_1^e z_2^y z_3^e; \\
 z^{27} &= z_1^o z_2^y z_3^e; \quad z^{28} = z_1^b z_2^y z_3^e; \\
 z^{29} &= z_1^u z_2^y z_3^e; \quad z^{30} = z_1^z z_2^y z_3^e; \\
 z^{31} &= z_1^e z_2^y z_3^e; \quad z^{32} = z_1^o z_2^y z_3^e.
 \end{aligned}$$

Выводы

Для построения моделей механизмов естественного языка нужны более мощные алгебрологический аппарат и методы формального описания, чем те, которые разработаны к настоящему времени. Поэтому в статье уделяется внимание описанию применяемых в этой модели существующих алгебрологических средств и разработке новых. В работе алгебрологические методы не только применяются на практике, но и развиваются далее до того уровня, при котором моделирование механизмов языка становится возможным. Главными ограничивающими факторами при моделировании языка являются недостаточное развитие алгебрологических средств и недостаточное развитие методов физического (объективного) исследования механизмов естественного языка.

В процессе рассмотрения методов можно делать вывод, что мы разрабатываем не актуальный, а потенциальный вариант устного склонения прилагательных. Потенциальный вариант разделяется на ближний (псевдослова), более далекий (введение звука ϕ в конце слова) и еще более далекий (введение ϵ в качестве первого звука окончания).

В дальнейшем предполагается исследование совершенной дизъюнктивной нормальной формы предиката модели звукового варианта склонения полных непряжательных имен прилагательных.

Список литературы: 1. Широков, В. А. Феноменология лексикографических систем [Текст] / В. А. Широков. — К.: Наукова думка, 2004. — 327 с. 2. Бондаренко, М. Ф. Модели языка [Текст] / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнарченко // Бионика интеллекта: научн.-техн. журнал. — 2004. — № 1. — С. 27-37. 3. Бархударов, С. Г. Учебник русского языка [Текст] / С. Г. Бархударов, С. Е. Крючков. — М.: Просвещение, 1970. — 296 с. 4. Зализняк, А. А. Грамматический словарь русского языка [Текст] / А. А. Зализняк. — М.: Русский язык, 1977. — 879 с. 5. Бондаренко, М. Ф. Обионике интеллекта [Текст] / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнарченко // Бионика интеллекта научн.-техн. журнал. — 2006 — № 1. — С. 3-5.

Поступила в редколлегию 29.03.2010 г.

УДК 519.7

Модель усной мови / Н. Є. Русакова // Біоніка інтелекту: наук.-техн. журнал. — 2010. — № 1 (72). — С. 94–97.

В статті розглядається проблема формалізації поняття тексту та його математичного описання, тобто представлення тексту у вигляді рядків знакових мовних повідомлень, якими люди обмінюються в процесі інтелектуальної діяльності. Розглядається модель усної мови на прикладі звукового варіанта відмінювання прикметників.

Табл. 13. Лл. 1. Бібліогр.: 5 найм.

UDC 519.7

Model of verbal speech / N. E. Rusakova // Bionics of Intelligence: Sci. Mag. — 2010. — № 1 (72). — P. 94–97.

In the article the problem of formalization of concept of text is examined and him mathematical description, that presentation of text is as lines of sign linguistic reports, which people change in the press of intellectual activity. The model of verbal language is examined on the example of voice variant of conjugation of adjectives.

Tab. 13. Fig. 1. Ref.: 5 items.