

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ АКЦЕНТНЫХ ХАРАКТЕРИСТИК РУССКОГО ЯЗЫКА

Перспективным направлением совершенствования техники синтеза речи, а также методов фонетического анализа в системах автоматического понимания речевых сообщений является широкое использование лингвистической информации, относящейся к закономерностям фонетической организации высказывания, характерным для данного языка. Применение сведений о языке системы в процессе распознавания речи дает возможность скорректировать ошибки и устранить неопределенности, неизбежно возникающие при автоматическом распознавании входного акустического сигнала. В процессе формирования речевого сообщения фонемы получают конкретные акустические характеристики, соответствующие синтаксической и семантической структуре предложения, местоположению ударения, границам слов и фонемному окружению. В связи с этим необходимо полное, однозначное и эксплицитное описание характеристик звуков речи, правил произношения, характерных для данного языка, закономерностей следования фонем, мелодических и акцентных характеристик, которое предназначено для кибернетических машин и сделает доступным для них свойственные человеку операции по обработке речевых сообщений.

В настоящей работе исследуются акцентологические нормы в области образования форм слов русского языка и строится их действующая математическая модель с применением средств теории интеллекта [1, 2].

Можно предположить, что связь между значением условного ударения X , акцентной характеристикой слова Y и набором грамматических признаков Z может быть описана в виде тернарного отношения $F(X, Y, Z)$. Акцентуационное отношение, как и всякое другое, целиком и полностью определяется своей характеристической функцией $F(X, Y, Z)$, которая задается следующим образом:

$$F(X, Y, Z) = \begin{cases} 1, & \text{если } F(X, Y, Z) \text{ выполняется,} \\ 0, & \text{в противном случае,} \end{cases} \quad (1)$$

т. е. если грамматические признаки Z_1, Z_2, \dots, Z_n , акцентная характеристика Y и значение условного ударения X согласуются в соответствии с нормами языка, то предикат $F(X, Y, Z)$ принимает значение 1, в противном случае — 0. Таким образом, задача заключается в разработке способа задания и определения вида функции $F(X, Y, Z)$ для склоняемых и спрягаемых слов современного русского литературного языка.

При образовании форм одного и того же слова ударение может перемещаться с одного слога на другой, с одной морфемы на другую, например: *де́ло, дела́, ве́рчу, вертя́т*. Эта особенность называется подвижностью словесного ударения. Подвижное ударение служит дополнительным средством, которое сопровождает основное средство образования грамматических форм — аффиксацию. При построении главной акцентуационной характеристики парадигм вводится особый акцентуационный признак отдельной словоформы — условное ударение. Для его обозначения введем переменную x . В отличие от ударения в обычном смысле слова условное ударение не приписывается какой-либо определенной букве или определенному слогу словоформы; оно считается принадлежащим морфологическому компоненту (основе или флексии в целом) и поэтому переменная x может принимать два значения: n — условное ударение на основе (неконечное) и k — условное ударение на флексии (конечное). Формы слов с разным морфемным составом объединяются по признаку наличия в них неконечного или конечного ударения. Для изменяемых слов основной единицей классификации по ударению является акцентный тип. Акцентный тип — это общая схема размещения ударения в словоформах определенного множества слов, относящихся к данной части речи. Распределение русских имен (в особенности существительных) по акцентным типам подчиняется определенным закономерностям, однако в некоторых случаях обнаружить какую-либо закономерность не удастся и приходится прибегать к составлению списков.

Введем переменную y , представляющую акцентный тип слова. Областью определения переменной будем считать множество чисел $\{1, 2, \dots, t\}$, где t — общее число схем ударения. Имеет место соотношение $y^1 \vee y^2 \vee \dots \vee y^{18} = 1$ (2). Для математического описания правил, приписывающих каждой первичной словоформе одно из двух значений признака условного ударения, необходимо ввести следующие переменные: z_1 — часть речи со значениями s — существительное, g — глагол, p — прилагательное, ch — числительное, n — неизменяемая часть речи; z_2 — тип склонения существительных со значениями a — адъективное, c — субстантивное; z_3 — признак одушевленности существительных со значениями o — одушевл., n — неодушевл.; z_4 — форма прилагательных со значениями p — полная, k — краткая. Через $z_i (i=5, 10)$ обозначим переменные, характеризующие словоизменительные категории, где z_5 — число со значениями e — ед., m — мн.; z_6 — падеж со значениями i — им., p — родит., d — дат., v — вин., t — творит., p — предл.; z_7 — наклонение глаголов со значениями i — изъявит., p — повелит., c — сослагат.; z_8 — время со значениями n — наст., p — прош., b — буд.; z_9 — род со значениями m — муж., zh — жен., c — ср.; z_{10} — лицо со значениями 1 — первое, 2 — второе, 3 — третье. Значения переменных

введены в соответствии с грамматикой русского языка [3]. Для рассматриваемых переменных справедливы уравнения

$$z_1^c \vee z_1^r \vee z_1^n \vee z_1^t \vee z_1^- \quad (3); \quad z_2^a \vee z_2^c \quad (4); \quad z_3^o \vee z_3^h \quad (5); \quad z_4^n \vee z_4^k \quad (6); \\ z_5^e \vee z_5^m \quad (7); \quad z_6^h \vee z_6^p \vee z_6^a \vee z_6^t \vee z_6^n \quad (8); \quad z_7^h \vee z_7^n \vee z_7^c \quad (9); \quad z_8^h \vee z_8^n \vee z_8^o \quad (10); \quad z_9^m \vee z_9^k \vee z_9^c \quad (11); \quad z_{10}^1 \vee z_{10}^2 \vee z_{10}^3 \quad (12).$$

Сочетаемость значений введенных переменных для существительных опишем уравнением

$$z_1^c \supset (z_2^a \vee z_2^c)(z_3^o \vee z_3^h)(z_5^e \vee z_5^m)(z_6^h \vee z_6^p \vee z_6^a \vee z_6^t \vee z_6^n). \quad (13)$$

Для глагольных форм расширенной парадигмы имеет место соотношение

$$z_1^2 \supset z_7^h (z_8^h (z_5^e \vee z_5^m)(z_{10}^1 \vee z_{10}^2 \vee z_{10}^3) \vee z_8^n (z_5^m \vee z_5^e (z_9^m \vee z_9^c \vee z_9^k)) \vee z_8^o) \vee z_7^c (z_5^m \vee z_5^e (z_9^m \vee z_9^k \vee z_9^c)) \vee z_7^n (z_5^e \vee z_5^m). \quad (14)$$

Предикат вида

$$z_1^n \supset z_4^n (z_5^e (z_9^m \vee z_9^k \vee z_9^c) \vee z_5^m)(z_6^h \vee z_6^p \vee z_6^a \vee z_6^b \vee z_6^t \vee z_6^n) \vee z_4^k \times \\ \times (z_5^m \vee z_5^e) (z_5^m \vee z_9^k \vee z_9^c) \quad (15)$$

описывает сочетаемость переменных, характеризующих словоизменительные категории прилагательных.

Перейдем к формальному описанию нормативных акцентных характеристик склоняемых и спрягаемых слов современного русского литературного языка. Математическую модель акцентных характеристик целесообразно представить в виде системы уравнений F_1, F_2, \dots, F_k , выражающих грамматические правила на языке алгебры конечных предикатов.

Для существительных зависимость значений признака «условное ударение» от акцентной характеристики слова и значений грамматических признаков выражается следующими правилами.

Акцентный тип (1), охватывающий весь основной массив существительных, характеризуется неподвижным ударением на одном и том же слоге основы во всех формах: $y^1 \supset x^n$ (16). Акцентный тип (2) характеризуется неподвижным ударением на флексии во всех словоформах: $y^2 \supset x^k$ (17).

Для существительных акцентного типа (3) характерно ударение на основе в форме им. пад. мн. числа (у неодушевл. сущ. — также вин. пад. мн. числа) и на флексии во всех остальных формах:

$$y^3 \supset (x^n \sim z_5^m (z_6^a \vee z_6^b z_3^h))(x^n \vee x^k). \quad (18)$$

Существительные акцентного типа (4) имеют ударение на флексии во всех формах ед. и мн. числа, за исключением форм

вин. пад. ед. числа и им. и вин. пад. мн. числа — с ударением на основе:

$$y^4 \supset (x^h \sim z_5^h (z_6^h \vee z_6^b) \vee z_5^e z_5^b) (x^h \vee x^k). \quad (19)$$

Акцентный тип (5) характеризуется ударением на основе во всех формах ед. числа и на флексии — во всех формах мн. числа:

$$y^5 \supset z_5^e x^h \vee z_5^m x^k. \quad (20)$$

Акцентный тип (6) характеризуется ударением на основе во всех формах ед. числа и в форме им. — вин. пад. мн. числа, ударением на флексии — в формах всех остальных падежей мн. числа:

$$y^6 \supset (z_5^e \vee z_5^m (z_6^h \vee z_6^b) \sim x^h) (x^h \vee x^k). \quad (21)$$

Для акцентного типа (7) характерно ударение на флексии во всех формах ед. числа и ударение на основе — во всех формах мн. числа:

$$y^7 \supset x^k z_5^e \vee x^h z_5^m. \quad (22)$$

Существительные акцентного типа (8) имеют ударение на флексии во всех формах ед. числа, за исключением формы вин. падежа, и ударение на основе — в форме вин. пад. ед. числа и во всех формах мн. числа:

$$y^8 \supset (x^h \sim z_5^e z_6^b \vee z_5^m) (x^h \vee x^k). \quad (23)$$

Перейдем к математическому описанию правил для предикативных глагольных форм. Для глаголов акцентного типа (9) характерно неподвижное ударение на одном и том же слоге основы во всех спрягаемых формах: $y^9 \supset x^h$ (24).

Глаголы, принадлежащие акцентному типу (10), характеризуются ударением на флексии во всех формах наст. — буд. времени и в повелит. наклонении; в прош. времени ударение на последнем слоге основы

$$y^{10} \supset (z_7^h (z_8^h \vee z_8^b) \vee z_7^h \sim x^h) (x^h \vee x^k). \quad (25)$$

Акцентный тип (11) характеризуется в прош. времени ударением на последнем слоге основы; в наст. — буд. времени — на флексии в 1 л. ед. числа и на основе — в ост. формах; в повелит. наклонении ударение, как и в 1 л. ед. числа, — на флексии:

$$y^{11} \supset (z_7^h (z_8^h z_5^e z_{10}^1 \vee z_8^b) \vee z_7^h \sim x^k) (x^k \vee x^h). \quad (26)$$

Для глаголов акцентного типа (12) характерно ударение на флексии во всех спрягаемых формах: $y^{12} \supset x^k$ (27).

Построим математическую модель размещения условного ударения в полных и кратких именах прилагательных. Известно, что акцентуационные характеристики полных и кратких прилагательных во многом не совпадают [3]. Ударение в полных

формах отличается большей устойчивостью. Все полные прилагательные по характеру ударения делятся на два акцентных типа: один из них характеризуется неподвижным ударением на основе во всех падежных формах ед. и мн. числа, для другого характерно неподвижное ударение на флексии во всех словоформах. Указанные схемы размещения ударения в полных формах прилагательных описываются уравнениями (24) и (27) соответственно. Среди прилагательных, имеющих полные и краткие формы, выделяются следующие акцентные типы по соотношению неконечного и конечного ударения в полных и кратких формах.

Акцентный тип (13) характеризуется неподвижным ударением на основе во всех полных и кратких формах: $y^{13} \supset x^n$ (28).

Для прилагательных акцентного типа (14) характерно ударение на основе во всех полных формах и на флексии — во всех кратких формах:

$$y^{14} \supset z_4^n x^n \vee z_4^k x^k. \quad (29)$$

Прилагательные акцентного типа (15) имеют следующие характеристики: ударение на основе во всех полных и кратких формах, кроме краткой формы женского рода, где ударение на флексии:

$$y^{15} \supset (z_4^k z_5^e z_9^j \sim x^k)(x^n \vee x^k). \quad (30)$$

Акцентный тип прилагательных (16) характеризуется неподвижным ударением на флексии во всех полных и кратких формах: $y^{16} \supset x^k$ (31).

Прилагательные акцентного типа (17) имеют ударение на флексии во всех полных формах и в краткой форме жен. рода, на основе — в остальных кратких формах:

$$y^{17} \supset (z_4^l z_4^k z_5^e z_9^j \sim x^k)(x^n \vee x^k). \quad (32)$$

Все существительные, изменяющиеся по адъективному склонению, распределяются по двум акцентным типам полных форм прилагательных.

Ударение числительных представлено тремя акцентными типами. Два из них характеризуются неподвижным ударением на основе во всех формах и неподвижным ударением на флексии и могут быть описаны уравнениями (24) и (27) соответственно. Для числительных, принадлежащих акцентному типу (18), характерно ударение на флексии во всех косвенных падежах и на основе — в форме им. пад.:

$$y^{18} \supset z_6^n x^n \vee (z_6^p \vee z_6^a \vee z_6^b \vee z_6^t \vee z_6^n) x^k. \quad (33)$$

Последним звеном нашей модели является описание предикатов, устанавливающих зависимости между правильными словоформами изменяемых частей речи и соответствующими им

акцентными типами. Для существительных субстантивного и адъективного склонения справедливы соотношения

$$z_1^c z_2^c \supset y^1 \vee y^2 \vee y^3 \vee y^4 \vee y^5 \vee y^6 \vee y^7 \vee y^8 \quad (34); \quad z_1^c z_2^a \supset y^9 \vee y^{12}. \quad (35)$$

Для глаголов имеет место

$$z_1^r \supset y^9 \vee y^{10} \vee y^{11} \vee y^{12}. \quad (36)$$

Полные и краткие формы прилагательных представлены следующими акцентными типами:

$$z_1^n \supset y^9 \vee y^{12} \quad (37); \quad z_1^n \vee z_1^k \supset y^{13} \vee y^{14} \vee y^{15} \vee y^{16} \vee y^{17}. \quad (38)$$

Зависимость между словоформами числительных и соответствующими им акцентными характеристиками выражается предикатом

$$z_1^4 \supset y^9 \vee y^{12} \vee y^{18}. \quad (39)$$

Таким образом, полученная система уравнений представляет собой математическое описание акцентных характеристик русского языка. Разработанная математическая модель должна естественным образом состыковываться с моделью морфологического уровня обработки словоформ для получения информации о грамматических признаках. Результаты экспериментальных исследований, проведенных с применением действующей программы решения уравнений алгебры конечных предикатов на ЭВМ [4], свидетельствуют о безошибочном действии модели.

Список литературы: 1. Шабанов-Кушнаренок Ю. П. О теории интеллекта. — Проблемы бионики, 1979, вып. 22, с. 3—10. 2. Шабанов-Кушнаренок Ю. П. Об уравнениях теории интеллекта. — АСУ и приборы автоматики, 1979, вып. 51, с. 10—15. 3. Грамматика современного русского литературного языка. — М.: Наука, 1970. — 200 с. 4. Бондаренко М. Ф., Бондарев В. М. Программа решения систем уравнений алгебры конечных предикатов. — Фонд алгоритмов и программ, изд. АН СССР, 1980, № 5645, с. 17—23.

Поступила в редколлегию 24.02.83