

УДК 510.62

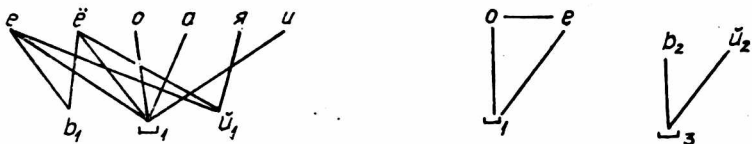
А. Ф. ОСЫКА, канд. техн. наук, В. А. ЧИКИНА,
А. И. ЧУГУН, канд. техн. наук

О МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ ЯВЛЕНИЙ ЧЕРЕДОВАНИЯ

При математическом моделировании русского языка возникает необходимость в рассмотрении такого рода явлений, как чередования, так как значительная часть словоформ им подвержена. Чередования могут наблюдаться при формообразовании (конец — кон₁ц[а, резать — режу), а также при словообразовании (соб₁рать — собер[у, тускл[ый — тускнеть, кни-з[а—книжиц[а). Для различных частей речи при словоизменении рассматриваемые явления имеют свои особенности: у существительных они проявляются при склонении, у глаголов — при спряжении, у прилагательных — при переходе от полных форм к кратким.

В данной работе рассмотрим один из возможных подходов к математическому моделированию явлений чередования в именах существительных русского языка при словоизменении.

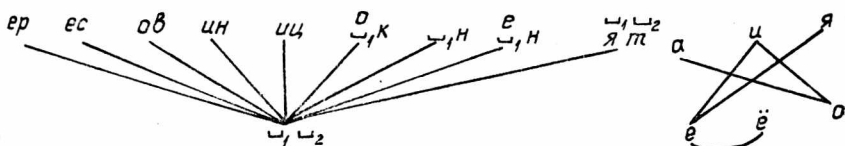
Под различными явлениями чередования будем понимать следующие лингвистические явления.



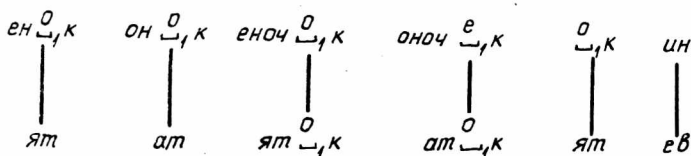
1. Беглость гласной или чередование гласной с ₁, обозначающим «нуль» гласной (слиток — слит₁к[а).

2. Беглое смягчение основы (словоформа без окончания). Рассматриваются случаи беглости в конце основы *ь*, *й* (огонь — огнЕя, музей — музеЕя).

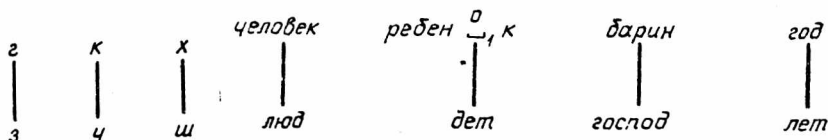
3. Беглость суффикса. В академической грамматике [1] рассматривается как наращение и усечение основы (гражданин — гражданинЕе, знамЕя — знамён[а]).



4. Чередование гласных (зар[я — зор[и]).



5. Чередование суффиксов (хозяин — хозяев[а]).



6. Чередование согласных (друг — друг[я]).

7. Супплетивизм (человек — люд[и]).

В работе рассматривается письменная форма русского языка. Вследствие этого такие понятия как чередование, альтернативный ряд [1] будем применять для букв. Буквы (гласные, согласные, ь, й) или последовательности букв, принимающие участие в чередовании, составляют альтернативный ряд, а члены альтернативного ряда — альтернанты.

Математическим аппаратом для построения модели является алгебра конечных предикатов [2, 3]. Под математической моделью будем понимать определенным образом построенное морфологическое отношение $L(X, Y)$, выражающее зависимость фрагмента текста Y от его смысла X . Под фрагментом текста Y будем понимать альтернант $Y = (y_1, y_2, \dots, y_n)$, где y_1, y_2, \dots, y_n — соответственно 1-я, 2-я, ..., n -я буквы альтернанта, нумерация слева направо.

Под смыслом текста X будем понимать совокупность признаков, соответствующие наборы значений которых характеризуют условия существования альтернанта.

В именах существительных чередования представлены в финалях — конечных элементах основы, в частном случае они могут быть морфемами. В качестве финалей выступают как

суффиксы (или их части) (рисунок — рисунок₁к[а], воронёнок — воронён₁к[а]), части корня (лёд — льд[а]), так и отдельные буквы (музей — музе₁я, око — оч₁и).

Все рассмотренные альтернативные ряды имен существительных выделены в регистр из 8-ми разрядов:

согл.₄ глас.₃ согл.₃ глас.₂ согл.₂ \bar{y}_2 глас.₁ согл.₁ \bar{y}_1 .

Такой регистр выбран исходя из удобства математического описания, хотя приведенной последовательности букв нет у одного альтернанта. Самый длинный альтернант представлен явлением супплетивизма, например, человек — люд[и]. Альтернативный ряд в регистре размещен следующим образом:

$$y_9^4 y_8^e y_7^d y_6^0 y_5^p y_4^- y_3^c y_2^k y_1^- - y_9^- y_8^- y_7^- y_6^- y_5^d y_4^- y_3^0 y_2^d y_1^-.$$

Пустые места в соответствующих разрядах регистра заполняются пробелами.

Для современного языка наличие чередования в большей степени обуславливается буквенным окружением альтернанта (в частном случае соседними морфемами). Это влияние описывается вектором смысла $X = (x_1, x_2, \dots, x_m)$, где x_1, x_2, \dots, x_m — компоненты вектора смысла, а именно: x_1 — падеж со значениями: и — именительный, р — родительный, д — дательный, в — винительный, т — творительный, п — предложный,

$x_1^i \vee x_1^p \vee x_1^d \vee x_1^v \vee x_1^t \vee x_1^p = 1$; x_2 — число со значениями:

e — единственное, m — множественное, $x_2^e \vee x_2^m = 1$; x_3 — одушевленность со значениями: 0 — одушевленный, н — неодушевленный, $x_3^0 \vee x_3^n = 1$; x_4 — ударность альтернанта со значениями: y — ударен, б — безударен, $x_4^y \vee x_4^b = 1$; x_5 — мягкость последней буквы усеченной основы (основа без финали) со значениями: м — мягкая, т — твердая, $x_5^m \vee x_5^t = 1$; x_6 — вариант чередования со значениями: 0 — основной, н — неосновной, $x_6^0 \vee x_6^n = 1$. К x_6^n относятся словоформы, имеющие нестандартную чередующуюся гласную (ва₁й[я — ваий, кочан₁коча₁н[а, семен[а — семян, заяц — зайц[а]). Все остальные существительные имеют $x_6 = d$.

x_7 — тип склонения альтернантов, $x_7^1 \vee x_7^2 \vee \dots \vee x_7^{42} = 1$

Этот признак осуществляет классификацию чередований внутри конкретного явления. Выделяются отдельным типом склонения и словоформы, не имеющие чередований.

Компоненты вектора смысла X можно разбить на два подмножества: $x_1 - x_3$ учитывают влияние дальнего текста, а $x_4 - x_7$ — влияние ближнего текста, т. е. буквенного состава словоформы. Опосредованное влияние дальнего текста через окончание не учитывается.

Возможные соотношения альтернантов показаны на схемах 1—7. В схемах учтены и изолированные альтернационные ряды.

Индексы при a, b, γ указывают на различное местонахождение их в основе. Как видно из приведенных схем, альтернационные ряды состоят из двух альтернантов (за некоторым исключением: сапожок — сапож \perp к[а — сапожек). Случаи, когда в основе встречаются более чем два альтернанта, не редки, например, мышонок — мышон \perp к[а — мышат[а. Но в связи с тем, что мы расклассифицировали рассматриваемые явления, нам удалось избежать такого рода сложностей при построении модели.

Далее полагаем, что мы умеем решать задачу выделения альтернанта из словоформы и помещать его в регистр, а также обратную ей задачу из регистра помещать в словоформу.

Посмотрим на явления чередования с формальной точки зрения. Такой подход приводит к тому, что мы получаем 9 предикатов, описывающих появление альтернанта в зависимости от тех компонентов вектора смысла X , которые характеризуют влияние дальнего текста. Далее, идя по пути упрощения, заметим, что при разложении предикатов по числу (переменная x_2) мы получим предикаты более компактного вида, а именно:

$$\begin{aligned}
 P_1(x_1, x_2, x_3, \beta, \gamma) &= x_2^e P_{11} \vee x_2^m P_{14}; P_2(x_1, x_2, x_3, \beta, \gamma) = \\
 &= x_2^e P_{12} \vee x_2^m P_{14}; P_3(x_1, x_2, x_3, \beta, \gamma) = x_2^e P_{12} \vee x_2^m P_{21}; P_4(x_1, x_2, x_3, \\
 &\beta, \gamma) = x_2^e P_{13} \vee x_2^m P_{14}; P_5(x_1, x_2, x_3, \beta, \gamma) = x_2^e P_{11} \vee x_2^m P_{21}; \\
 P_6(x_1, x_2, x_3, \beta, \gamma) &= x_2^e P_{13} \vee x_2^m P_{22}; P_7(x_1, x_2, x_3, \beta, \gamma) = \\
 &= x_2^e P_{11} \vee x_2^m P_{22}; P_8(x_1, x_2, x_3, \beta, \gamma) = x_2^e P_{14} \vee x_2^m P_{21}; P_9(x_1, x_2, \\
 &x_3, \beta, \gamma, \xi) = x_2^e P_{11} \vee x_2^m P_{23}; P_{10}(x_1, x_2, x_3, \alpha) = P_{15}. \quad (1)
 \end{aligned}$$

Поясним, что представляют собой предикаты, стоящие в правых частях (1):

$$\begin{aligned}
 P_{11}(x_1, x_3, \beta, \gamma) &= t^1 \beta \vee (t^2 \vee t^3 \vee t^4) \gamma; P_{12}(\beta) = \beta; \\
 P_{13}(x_1, x_3, \beta, \gamma) &= (t^1 \vee t^4) \beta \vee (t^2 \vee t^3) \gamma; \\
 P_{14}(\gamma) &= \gamma; P_{15}(\alpha) = \alpha; P_{21}(x_1, x_3, \beta, \gamma) = t^2 \gamma \vee (t^1 \vee t^3 \vee t^4) \beta; \\
 P_{22}(x_1, x_3, \beta, \gamma) &= t^4 \beta \vee (t^1 \vee t^2 \vee t^3) \gamma; P_{23}(x_1, x_3, \gamma, \xi) = t^2 \xi \vee \\
 &\vee (t^1 \vee t^3 \vee t^4) \gamma; P_{24}(\gamma) = \gamma; P_{25}(\alpha) = \alpha, \text{ заметим, что } P_{14} = P_{24}, \\
 P_{15} &= P_{25}. \text{ Таким образом нет необходимости с введением пре-} \\
 &\text{дикатов } P_{14}, P_{15} \text{ вводить предикаты } P_{24}, P_{25}.
 \end{aligned}$$

Константы t^1, t^2, t^3, t^4 используются для сокращения математической записи: $t^1 = x_1^a \vee x_1^b x_3^b$; $t^2 = x_1^p \vee x_1^b x_3^0$; $t^3 = x_1^a \vee x_1^b$; $t^4 = x_1^r$.

Сформулируем на формальном уровне правила для введенных предикатов $P_1 \div P_{10}$: предикат P_1 описывает альтернативный ряд из двух альтернантов, один из которых (β) присутствует в основах в единственном числе именительном падеже и винительном неодушевленном, а другой (γ) — во всех оставшихся падежах единственного и множественного числа (орёл, орл[а, орл[у, орл[а/(орёл), орл[ом, орл[орл[ы, орл[ов, орл[ам, орл[ов/(орл[ы), орл[ам, орл[ах]); в предикате альтернанты (β, γ) противопоставлены по числу (судн[о, судн[а, судн[у, (судн[а)/судн[о, судн[ом, судн[суд[а, суд[ов, суд[ам, (суд[ов)/суд[а, суд[ами, суд[ах]); предикат P_3 описывает появление одного из альтернантов γ в основах во множественном числе родительном падеже и винительном одушевленном, другого β — в оставшихся падежах множественного числа и в единственном числе (масл[о, масл[масл[у, (масл[а/(масл[о, масл[ом, масл[е; масл[масел, масл[ам, (масел)/масл[а, масл[ами, масл[ах); предикат P_4 описывает альтернативный ряд из двух альтернантов, один из которых β присутствует в основах в единственном числе в именительном падеже, винительном неодушевленном и творительном, другой γ — в оставшихся падежах единственного числа и во множественном (дверь, двер[и, двер[и, (двер[и]дверь, дверь[ю; двер[и; двер[и, двер[ей, двер[ям, (двер[ей]двер[и, двер[ями, двер[ях); в предикате P_5 альтернанты противопоставлены следующим образом: именительный падеж и винительный неодушевленный в единственном и множественном числе, а также во множественном числе дательный, творительный, предложный падежи β и оставшиеся падежи единственного и множественного числа γ (князь, княз[я, княз[ю, княз[князь), княз[ем, княз[е; князь[я, княз[ей, князь[ям, княз[ей(князь[я), князь[ями, князь[ях); в предикате P_6 один из альтернантов β появляется в основах в единственном числе именительном падеже и винительном неодушевленном, а также в творительном падеже единственного и множественного числа, другой γ — во всех оставшихся падежах единственного и множественного числа (лошадь, лошаd[и, лошаd[и, лошаd[и/(лошадь), лошаd[ю, лошаd[и; лошаd[и, лошаd[ей, лошаd[ям, лошаd[ей/(лошад[и), лошаd[ми, лошаd[ях); в предикате P_7 альтернанты противопоставляются следующим образом: единственное число именительный падеж и винительный неодушевленный и множественное число творительный падеж β и все оставшиеся падежи единственного и множественного числа γ (зверь, звер[я, звер[ю, звер[я/(зверь), звер[ем, звер[е; звер[и, звер[звер[ям, звер[ей/(звер[и), звер[ми, звер[ях); в предикате P_8 один из альтернантов β появляется в единственном числе именительном падеже и винительном одушевленном, другой (β) — во всех оставшихся падежах мно

друг[ом, друг[е; друзь[я, друз[ей, друзь[ям, друз[ей]/(друзь[я), друзь[ями, друзь[ях]; предикат P_9 описывает, в отличие от рассмотренных предикатов $P_1 \div P_8$ альтернативный ряд из трех альтернатив, один из которых (β) появляется в единственном числе именительном падеже и винительном неодушевленном, другой ξ — во множественном числе родительном падеже и винительном одушевленном, а третий γ — во всех оставшихся падежах единственного и множественного числа (сапожок, сапож[к[а, сапож[к[у, (сапож[к[а)/сапожок, сапож[к[ом, сапож[к[е; сапож[к[и, сапожек, сапож[к[ам, (сапожек)/сапож[к[и, сапож[к[ами, сапож[к[ах); предикат P_{10} описывает все основы, в которых не происходит чередований при словоизменении (альтернативный ряд представлен одним альтернативом α).

Далее, если в каждом из предикатов $P_1 \text{—} P_{10}$ определить альтернативы α , β , γ , ξ на неформальном уровне, то это и будет собственно математическая модель явлений чередования в именах существительных при словоизменении.

Список литературы: 1. *Русская грамматика*. — М.: Наука, 1980. — Т. 1, 783 с. 2. *Шабанов-Кушнарченко Ю. П.* Об алгебре конечных предикатов. — АСУ и приборы автоматики, 1977, вып. 52, с. 21—28. 3. *Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П.* О математическом описании естественного языка. — Пробл. бионики, 1981, вып. 27, с. 9—13.

Поступила в редколлегию 03.01.83.