

Беларуская мова і камп'ютар

ПРАВАПІС СПАЛУЧЭННЯЎ ЕА/ЭА, ЕЯ/ЭЯ

СПРОБА КАМП'ЮТАРНА-ЛІНГВІСТЫЧНАГА АНАЛІЗУ

У красавіцкім нумары “Роднага слова” за 2004 г. мы паведамлялі пра распрацоўку і выхад у свет камп'ютарнай праграмы праверкі беларускай мовы (інтэрнэт-адрас: www.pravaris.tut.by). З мэтай пашырэння функцыянальных магчымасцяў прадукту мы ствараем электронны слоўнік беларускай мовы з паказам парадыматыкі кожнага слова, які будзе сумяшчацца з варыянтам праграмы праверкі арфаграфіі (спэл-чэкера). Аўтарам распрацавана электронная база беларускай мовы (фармат Access), якая ўтрымлівае больш за 180 тыс. запісаў (рээстравых слоў).

Дэталёвае апісанне прынцыпаў арганізацыі і структуры базы будзе прапанавана пазней. У якасці адметнай рысы адзначым, што з кожным рээстравым словам базы суаднесены ягоны граматычны код, які дазваляе пры дапамозе спецыяльных алгарытмаў паказаць парадыгму адпаведнага слова і пазіцыю націску ў ягоных словаформах.

Электронная база дадзеных (далей БД) дазваляе выконваць разнастайныя запыты і атрымліваць у якасці выніку выбаркі слоў, якія адпавядаюць азначаным умовам (напрыклад, назоўнікі з фіналямі *-эр* і *-эль* або ўсе 4-складовыя словы з пазіцыяй націску на 4-й літары з пачатку). Зварот да прадстаўнічых моўных БД – гэта эфектыўны шлях выяўлення тэндэнцый напісання, што паслужыць асновай для фармулявання прапаноў па ўдакладненні або змене правапісных нормаў.

Напісанне спалучэнняў *ea/ea*, *ea/ea* і іх варыянтаў пад націскам *eo/eo*, *eo/eo* не рэгламентуецца “Правіламі арфаграфіі і пунктуацыі” ў рэдакцыі 1959 г. Праект новай рэдакцыі “Правілаў арфаграфіі і пунктуацыі” (Мінск, 2004) таксама пакідае па-за ўвагай напісанне названых спалучэнняў. Між тым інтэнсіўнае развіццё навуковай і тэхнічнай сфер камунікацыі аб'ектыўна вядзе да неабходнасці ўводу ў мову значнай колькасці іншамоўных слоў, напісанне якіх павінна адбывацца на аснове выразна сфармуляваных правілаў або выключэнняў з іх.

Праведзеныя выбаркі з распрацаванай намі электроннай БД паказалі, што без уліку вузкаспецыялізаванай тэрміналагічнай лексікі масіў іншамоўных слоў з разглядаемымі спалучэннямі галосных налічвае 1375 слоў, з якіх 675 слоў улічаны ў СБМ. Для параўнання адзначым, што колькасць іншамоўных слоў з ненаціскнымі фіналямі *-эр* і *-эль*, у дачыненні да якіх былі сфармуляваны прапановы праекта “Правілаў арфаграфіі і пунктуацыі” па замене фіналямі *-ар*, *-аль*, складаюць, адпаведна, 163 і 9 выпадкаў, з якіх у СБМ улічаны, адпаведна, 81 і 6 слоў.

Аналіз тэндэнцый напісання запазычаных слоў з разглядаемымі спалучэннямі галосных у сучасных агульных і тэрміналагічных слоўніках і энцыклапедычных выданнях дазваляе вылучыць наступныя варыянты іх напісання.

1. Збег галосных *ea/ea* рэгулярна рэгіструецца ў складаных і простых словах іншамоўнага паходжання ў складзе наступных кампанентаў: *азеа-* (*азеатропны*), *акмеа-* (*акмеалагічны*), *алеа-* (*алеаграфія*), *археа-* (*археалогія*), *арэа-* (*арэапаг*, *арэал*), *астэа-* (*астэапароз*), *аўрэа-* (*аўрэаміцын*), *бальнеа-* (*бальнеатэрапія*), *відэа-* (*відэамантаж*), *гамеа-/гомеа-* (*гамеапат*, *гомеамарфізм*), *геа-* (*геалогія*), *гнасеа-* (*гнасеалогія*), *ідэа-* (*ідэалагічны*, *ідэал*, *ідэальны*), *лінеа-* (*лінеал*, *лінеарны*, *алінеатар*), *меа-* (*меандр*), *метэа-* (*метэастанцыя*), *мімеа-* (*мімеаграфічны*), *неа-* (*неалагізм*), *нуклеа-* (*нуклеафільны*), *палеа-* (*палеабатаніка*), *панкрэа-* (*панкрэас*, *панкрэатыт*), *плеа-* (*плеахраізм*), *рэа-* (*рэалогія*, *рэакцыя*), *спелеа-* (*спелеаэкспедыцыя*), *стэрэа-* (*стэрэафанія*), *тахеа-* (*тахеаметрычны*), *трахеа-* (*трахеатамія*), *тырэа-* (*тырэатаксікоз*), *тэа-* (*тэалогія*), *феа-* (*феадал*), *фразеа-* (*фразеалогія*), *харэа-* (*харэаграфія*). Пры гэтым у вышэйпрыведзеным напісанні гэтыя кампаненты не выступаюць у якасці самастойных слоў за выключэннем нескланяльных слоў *відэа* і *стэрэа*.

Выключэнне складаюць словы *пляяда*, *эякуляцыя* і іх вытворныя, напісанне якіх запазычана без зменаў з рускай мовы.

Адзінкавыя адхіленні ад азначанага напісання ў тэкстах і нават асобных слоўнікавых выданнях, відавочна, памылковыя. Так, слова *алінеатар* у СБМ і БРС падаецца як “алінеатар”. У той самы час у іншых выданнях (СІС, РБПС) словы з кампанентам *лінеа-* падаюцца карэктна (*калінеарны*, *лінеарызацыя*).

2. Калі націск падае на другую галосную спалучэння, можа мець месца трансфармацыя ў *eo/ eo*: *геадэзія* – *геоід*, *ідэалогія* – *ідэолаг*, *метэарыт* – *метэор* і г. д. Выпадак адхілення ад дадзенага правіла ў межах даследаваных намі тэкставых масіваў не выяўлена за выключэннем відавочна памылковай падачы слова “меэз” у СБМ замест карэктнага *міэз* [рус. “миоз”], аналагічна напісанню *міёма*, *міёлаг* (СБМ, БРС).

3. У фіналях слоў іншамоўнага паходжання спалучэнні *ea / ea* ўскланяльных словах трансфармуецца ў *ea / ea*: *ідэя*, *трахея*, *Гея* (міф.), *Псіхэя* (міф.) і г. д.

У нескланяльных словах у залежнасці ад пазіцыі націску могуць быць напісанні *ea* або *eo / eo*: *відэа*, *стэрэа*, але *Барнео*, *радэо*.

Напісанне назвы горада *Ноўітэа*, пададзенай у РБС як “Нумеа”, з улікам вышэйзгаданага мэтазгодна замацаваць як *Нумея* (скланяльная форма). Заўважым, што ў сучасным рускамоўным друку такое напісанне мае тэндэнцыю да пашырэння.

4. Выпадкі ётацыі другога галоснага ў спалучэнні *ea(o)/ea(o)* у пачатку і ў сярэдзіне слова маюць месца толькі ў словах *эякуляцыя* (і вытворных ад яго) і *пляяда*, напісанне якіх адпавядае напісанню, прынятаму ў рускай мове.

5. Трэба адрозніваць выпадкі напісання складаных слоў, першым кампанентам якіх з’яўляюцца словы з фіналямі *-ея (-ей)/-эя (трахея, ідэя, алей)*, і складаных слоў з кампанентамі, якія не ўтвараюць самастойных лексем (*трахеа-, алеа-, ідэа-*). Сказанае можа быць праілюстравана наступнымі прыкладамі: *трахеяпадобны*, але *трахеатамія*; *алеяносны*, але *алеаграфія, алеафобны; ідэяносьбіт*, але *ідэалогія, ідэаграфічны* і г. д.

Такім чынам, на падставе аналізу тэкстаў і слоўнікавых матэрыялаў можна зрабіць выснову, што ў цэлым замацавалася аднастайнае напісанне слоў іншамоўнага паходжання са спалучэннямі галосных *е(э)а/е(э)я* і іх націскнымі варыянтамі *е(э)о/е(э)ё*.

З улікам значнага патэнцыялу пашырэння гэтага класа слоў за кошт утварэння новых тэрмінаў у сувязі з развіццём навукі і тэхнікі мэтазгодна замацаваць выяўленыя формы напісання ў новых правілах арфаграфіі.

Скарачэнні

БРС – Беларуская-рускі слоўнік: У 3 т. Выд. 3 / Пад рэд. А. Падлужнага. – Мінск: БелЭн, 2003; **СІС** – Булька А. Слоўнік іншамоўных слоў: У 2 т. – Мінск: БелЭн, 1999; **СБМ** – Слоўнік беларускай мовы: Арфаграфія. Арфаэпія. Акцэнттуацыя. Словазмяненне / Пад рэд. М. Бірылы. – Мінск: БелСЭ, 1987; **РБПС** – Русско-белорусский политехнический словарь: В 2 т. – Минск: Бел. наука, 1997 – 1998; **РБС** – Русско-белорусский словарь: В 3 т. – Минск: БелЭн, 2002.

Фёдар ПСКУНОЎ,
навуковы супрацоўнік Інстытута электронікі
НАН Беларусі.

Навука і камп’ютар

ЭЛЕКТРОННАЯ БАЗА ДАДЗЕННЫХ “БЕЛАРУСКАЕ МОВАЗНАЎСТВА”

Адна з важных умоў паспяховага развіцця навукі – інфармацыйнае забеспячэнне навуковых даследаванняў. Асноўныя носьбіты інфармацыйных рэсурсаў у Рэспубліцы Беларусь – Нацыянальная бібліятэка Беларусі, Цэнтральная навуковая бібліятэка імя Якуба Коласа Нацыянальнай Акадэміі навук Беларусі (ЦНБ імя Якуба Коласа) і шэраг рэспубліканскіх галіновых бібліятэк. Акрамя выдачы карыстальнікам першасных крыніц інфармацыі – кніг, часопісаў, навуковых бібліятэкі працуюць над стварэннем і папулярызацыяй інфармацыйна-бібліяграфічных рэсурсаў па пэўных галінах навукі. У ЦНБ імя Я. Коласа гэтую працу выконвае аддзел інфармацыйна-аналітычнага забеспячэння навукова-даследчай работы (НДР), які быў створаны ў 1952 г. і называўся аддзелам навуковай бібліяграфіі. Адзін з кірункаў яго дзейнасці – бібліяграфаванне літаратуры па беларускім мовазнаўстве.

У друкаваным выглядзе бібліяграфія беларускага мовазнаўства ахоплівае перыяд з 1825 да 1991 г. (паказальнікі 1960, 1967, 1980, 1993 і 2003 гадоў выдання). Інфармацыя пра навуковыя працы, створаныя ў наступныя гады, захоўваецца ў электроннай форме ў выглядзе бібліяграфічнай базы дадзеных (БД) “Беларускае мовазнаўства” ў адзеле інфармацыйна-аналітычнага забеспячэння НДР ЦНБ імя Якуба Коласа. База дадзеных складаецца з трох частак, абмежаваных храналагічнымі рамкамі: 1986 – 1991 гг., 1992 – 1999 гг. і ад 2000 г. да цяперашняга часу. На дадзены момант колькасць запісаў за перыяд 1992 – 1999 гг. складае 9546 адзінак, за перыяд ад 2000 г. да цяперашняга часу – 7900 адзінак. Кожны дзень база папаўняецца новымі запісамі з адлюстраваннем усіх бібліяграфічных элементаў і кароткага зместу кожнага дакумента. Агульная колькасць такіх запісаў у БД “Беларускае мовазнаўства” – больш за 20 000 (10 Мб). Мэта стварэння БД – своечасовае інфармацыйнае забеспячэнне навуковай працы ў галіне беларускага мовазнаўства.

Рэсурсы БД могуць быць карысныя для даследчыкаў мовы: навукоўцаў, выкладчыкаў, аспірантаў, студэнтаў і інш.

Магчымасці праграмнага забеспячэння дазваляюць здзяйсняць пошук неабходнай інфармацыі па шэрагу бібліяграфічных параметраў (палёў): аўтар, назва, год выдання і інш. Найбольш запатрабаваны сярод карыстальнікаў – пошук па ключавых словах (КС). Вылучэнне КС – найбольш важны і адказны этап у працэсе стварэння бібліяграфічнага запісу. Ён патрабуе ад бібліяграфа пэўнай лінгвістычнай падрыхтоўкі. Пры праглядзе дакумента на *de visu* (бегла) у поле “Ключавыя словы” ўносіцца ўся інфармацыя, якая дапамагае раскрыць яго змест. Абавязкова ўказваецца назва раздзела, да якога належыць дакумент, лінгвістычныя і аўтарскія тэрміны, мовазнаўчыя паняцці ад больш шырокіх да больш вузкіх. У якасці КС выкарыстоўваюцца назоўнікі (*сінтаксіс, перыфраза*), словазлучэнні (*семантыка лексічная, назоўнікі адушаўленыя*). Пры немагчымасці звесці КС да назоўніка ці словазлучэння выкарыстоўваюцца больш складаныя спалучэнні тыпу *твор мастацкі – мова, назвы геаграфічныя – паходжанне*. Пошук патрэбных дакументаў можа адбывацца па цэлым слове або па яго частцы.

Метадыка вылучэння КС накіравана на тое, каб карыстальнік не толькі знаходзіў канкрэтную інфармацыю, якая цікавіць яго ў дадзены момант, але і сутыкаўся з новымі паняццямі, тэрмінамі, што павышае зурывыстную каштоўнасць базы дадзеных. Гэта можа быць асабліва карысна для студэнтаў і маладых спецыялістаў у галіне беларускага мовазнаўства.

Лакальны доступ да рэсурсаў БД “Беларускае мовазнаўства” – у чытальнай зале № 5 ЦНБ імя Якуба Коласа.

Ірына СМІРНОВА,
малодшы навуковы супрацоўнік
аддзела інфармацыйна-аналітычнага забеспячэння
НДР ЦНБ імя Якуба Коласа.