

УДК 003

Жижко Володимир Абрамович

незалежний експерт

Жижко Владимир Абрамович

независимый эксперт

Zhyzhko Volodymyr

Independent Expert

**ЩОДО УКРАЇНСЬКОГО ТРАНСЛІТУ
ОТНОСИТЕЛЬНО УКРАИНСКОГО ТРАНСЛИТА
ABOUT UKRAINIAN TRANSLIT**

Анотація. Виходячи з особливостей українського правопису, обґрунтовано оптимальну транслітерацію - правила написання українських текстів латинськими літерами.

Ключові слова: транслітерація, текст українською мовою, латинська абетка.

Аннотация. Исходя из особенностей украинского правописания, обоснованно оптимальную транслитерацию - правила написания украинских текстов латинскими буквами.

Ключевые слова: транслитерация, текст на украинском языке, латинский алфавит.

Summary. Coming from the features of the Ukrainian spelling, optimal transliteration - are rules of writing of Ukrainian text with Latin letters proved.

Key words: transliteration, text by Ukrainian, Latin alphabet.

Зміст проблеми. У 1999 році в [1] з'явилася стаття про доцільність використання латинської абетки для українського правопису. Це питання час від часу піднімають у публічному просторі. У пропонованому дописі викладена спроба обґрунтувати оптимальне рішення. Чому саме у комп'ютерному виданні було порушено це питання - зрозуміти неважко. Саме розробники та користувачі програм витрачають певні зусилля на те, щоб прочитати або вірно сформулювати текстовий результат. Ось кілька типових ситуацій:

- у переліку листів електронної пошти можна зустріти назву, відображену недоречною комбінацією символів;
- щось подібне трапляється інколи при спробі зайти на кириличний сайт;
- при конвертації з Excel у текст (Save As *.txt) інколи невірне перетворення символу "і". (Причина полягає у невдалому виборі кода символу "і" української абетки у Windows);
- неповна сумісність таблиць баз даних в DOS та Windows за наявності полів типу "текстовий" при використанні "локальних" (не латинських) абеток;
- неповна відповідність відображення тексту на екрані та в твердій копії (коли для якогось рядка встановлено, приміром, шрифт Arial замість ArialCyr).

Перелік схожих ситуацій можна подовжити.

У наведених ситуаціях йдеться про (не)відповідність типу коду, кодової сторінки та шрифту програмним засобам відображення текстів.

Подолати проблему можна різними способами, наприклад:

- 1) розробити досконалі програми конвертації текстів, Це складна задача, оскільки різні системи кодування не повністю узгоджені навіть по наборах символів. Цей шлях описаний в [2];

2) працювати в обмеженому ("стерильному") кодовому середовищі. Це цілком можливо та ефективно для закритої корпоративної спільноти. Типові обмеження - єдина ОС, єдина СУБД, єдині поштова програма та браузер Інтернету, 3 шрифти для документів. У звичайному "гетерогенному" середовищі програмісти і користувачі інколи й досі мають клопіт з цими дрібними докучливими проблемами;

3) обмежити набір символів латинським регістром клавіатури. Саме про це увесь подальший виклад. Перевага цього способу - практично усувається причина, яка породжує проблему, не потрібні додаткові програмні засоби, недолік - незвичність та велика кирилична "спадщина".

Багатьох проблем не існує для латинської абетки, оскільки латина розміщується в основній половині кодового простору (а будь-яка інша "локальна" абетка - у допоміжній половині), тобто для будь-якої мови код латини все одно присутній, а відповідність код-символ однакова для латини у всіх мовах (кодових сторінках). З часу появи згаданої статті [1] дещо змінилося, і питання стало ще актуальнішим.

Успішно реалізовані російські проекти Russkaja Latinica, Cifirica, Shtirlitz (подробиці на одноіменних сайтах .ru , див. також [3]). Більшість країн колишнього Союзу зробили кроки для переходу або вже перейшли з кириличного письма на латинське. Усі країни Європи на Захід від України користуються латиницею, за винятком Сербії та Болгарії, де й була створене слов'янське письмо - кирилиця, до речі, із значними запозиченнями з латиниці.

Наступна проблема - неоднозначність транслітерації. Відсутність єдиного стандарту відповідного правопису найбільш рельєфно видно при написанні власних імен та прізвищ - на візитівках, пластикових картках та міжнародних паспортах. У різних документах, що їх могла одержати від різних установ людина на прізвище ХИЖНЯК, ймовірно знайти такі варіанти написання Khizhnyak, Hyghnjac, Huzhniak і т. ін. в залежності ледве

не від особистих уподобань працівників різних установ. (Наочний приклад - прізвища на футболках гравців та у титрах). Це може спричинити не тільки незручності, але й правові колізії. На сайті [4] наведені аж 8 стандартів української латиниці. Усі вони виходять з однозначної заміни символів абетки незалежно від лексичного контексту.

Які умови треба виконати, аби український правопис латинськими літерами (скорочено UL) став конкурентним або й привабливішим за кириличний? Мабуть, основні умови такі:

- зберегти існуючу простоту правопису - в українській мові практично існує однозначна відповідність звуків та літер, тобто відсутня транскрипція;
- середня довжина UL тексту не повинна перевищувати цей показник для кирилиці;
- простота та зручність перетворення та вводу текстів;
- стандартний набір символів латинської абетки (без надчерків і т. ін.).

Критерій оптимальності транслітерації можна визначити так:

- мінімум випадків порушення локальної відповідності символів кириличної та латинської абеток, отже, мінімум ситуацій контекстно-залежних заміन (транскрипцій) символів кирилиці на символи латиниці;
- мінімум використання спецсимволів;
- мінімум виключень з правил;
- "оригінальних", використання традиційних для латинських мов лексичних конструкцій (сполучень символів).

Відповідність кирилиця - латиниця

А тепер конкретно про заміну набору кириличних символів на латинські. Хоча може здаватися, що перехід від кириличного набору до латиниці майже однозначний, насправді це не так. Мабуть, кожен програміст був змушений писати (або з власної цікавості писав) уривки

текстів у "польській" нотації. При цьому можна спостерігати велике розмаїття варіантів відповідності кирилиця - латиниця. Встановити певний стандарт відповідності кирилиця - латиниця можна було б на основі логічного обґрунтування. Неоднозначні рішення виникають, в основному, при спробі встановити відповідності для голосних та шиплячих. Будемо шукати рішення, яке задовольнить певні умови, а саме:

- набір латинських символів відомий, збільшувати його не можна, хіба що на один-два символи, присутні на стандартній клавіатурі;
- написання слів має бути зручним і коротким, ніякий символ кирилиці не слід зображувати більш ніж двома символами латиниці (нагадаємо - символів в українській кириличній абетці майже на третину більше за латинську абетку);
- слід використовувати особливості української мови для скорочення письма;
- правила відповідності кирилиця-латиниця мають бути простими.

Ці умови здаються чимсь само собою зрозумілим лише на перший погляд. Почнемо з голосних. Скористаємося відомим у фонетичному аналізі розкладом звуків на первісні складові. Приміром, слово "ягня" можна зобразити у вигляді "йагня", де первісною складовою є звук "а", а допоміжними "й" та "ь". Отже, повний набір голосних утворюється з глухих голосних шляхом йотування та пом'якшення. Побудуємо таблицю для усіх голосних

Таблиця 1

Транслітерація голосних

1	2	3	4	5	6	7	8
а	йа	я	ья	я	а	ја	`а
е	йе	є	ье	є	е	је	`е
И	йи	ї	ьи	і	у	ји	і
О	йо	ю	ьо	ю	о	јо	`о
У	йу	ю	бу	ю	и	ји	`и

Поля таблиці:

- 1) первісна кирилична літера (глухий звук);
- 2) йотована кирилична;
- 3) зображення йотування українською лексикою;
- 4) пом'якшена кирилична;
- 5) зображення пом'якшення українською лексикою;
- 6) латинська відповідність первісній кириличній літері;
- 7) транслітерація йотованої кириличної;
- 8) транслітерація пом'якшеної кириличної.

Вибравши для йотування в латиниці, природньо, символ "j", а як еквівалент м'якого знаку - апостроф, отримуємо просту відповідність символів кирилиці та латиниці – права частина таблиці 1. (У часи Нечуя-Левицького та Франка вживалося, зокрема, ja). При явному використанні "j" для йотування голосних немає більше потреби вживати апостроф для відокремлення при вимові голосного звуку від приголосної.

Для скорочення запису слів можна використати таку особливість української мови - у величезній більшості слів, де зустрічаються поруч дві голосні, друга є йотованою. Це дозволяє опускаєти "j" на письмі, додаючи цей звук при вимові. Приклади:

*Kyiv ta kyany, Ukraina, Azia, Rosia, maetok pana maora z kamjanou
bramou*

Виключення становлять слова іншомовного походження та ті, де префікс закінчується на голосну, та корінь слова починається з голосної. Аби уникнути неоднозначності, у цих випадках слід явно відокремлювати на письмі ці голосні. Одним з варіантів може бути такий - використати апостроф. Приклади: na`uka ne`obachno

Для шиплячих можна скористатися звичною англійською відповідністю:

(ж => zh) (ч => ch) (ш => sh) (щ => sch)

яка ґрунтується на фонетичній спорідненості. А саме, можна простежити таку послідовність утворення базових шиплячих звуків s=>(s c z) => (sh ch zh).

Написання попарних комбінацій базових шиплячих можна скоротити, опускаючи першу літеру "h". Оскільки щ походить від шч, то звідси і відповідність щ => sch. Приклади слів із парними сполученнями інших шиплячих: nuzche blyzche Donecchyna Zaporizzh'a

Тепер можна подати повну таблицю латинського еквіваленту кириличної абетки:

Таблиця 2

Відповідність символів транслітерації

a	б	в	г	г'	д	е	є	ж	з	ь
a	b	v	g	q	d	e	je 'e	zh	z	'
i	ї	й	к	л	м	н	о	п	р	'
i	ji	j	k	l	m	n	o	p	r	j
с	т	у	ф	х	ц	ч	ш	щ	ю	я
s	t	u	f	h	c	ch	sh	sch	ju 'u	ja 'a

Отже, на пом'якшення голосних вказує апостроф перед цією голосною, а літера "j" відображує йотування або просте "й" залежно від лексичної ситуації. Після детального аналізу були відкинуті як недостатньо обґрунтовані деякі відповідності (табл.3), які можна зустріти у практичному трансліті (прізвища спортсменів у титрах та на футболках, таблиця Western Union, картки Visa, тощо):

Таблиця 3

Відхилені відповідності

г	г'	є	ж	й	с	х	щ	ю	я
		ie							
h	g	ye	gh	y	ts	x	xh	iu	ya

З боку канонічної латиниці, не використаними ("зайвими" для української латинської абетки) залишилися літери W X. Їх можна використати при оригінальному написанні іншомовних слів. Ясна річ, що лексикографічна упорядкованість в транслітному тексті буде визначена як латинська. Подальший виклад ведеться українською латиницею.

Mozhlyvosti skorochnn'a latyns'kogo zapysu

Na osnovi statystychnogo analizu movnyh sytu'acij mozhna zaproponuvaty sche kil'ka dzherel skorochnn'a pys'ma:

- 1) vvazhaty literu "l" pochatkovo mjakou: dla bilshosti plam, - pisma tvrdoj "l" stavty apostrof: bul`a zol`ota pal`ka;
- 2) opuskaty apostrof (mjakyj znak), zberigauchy mjaku vymovu:
 - u sufiksah "ськ","цьк","зьк" - pysaty "sk","ck","zk"ж;
 - pisma sufiksa "ist" (napr., jakist, vartist, kmitlyvist);
 - zvorotnu chastku "ся" u dieslovah pysaty "sa";
 - opuskaty na pys`mi "j" mizh golosnyh, - opuskaty na pys`mi "h" mizh shyplachymy, jak zhadano u p.2.

Cej perelik ne je vycherpnyh, ta blyzkyj do optymального.

V ukrainskij movi nemaє sliv, jaki pochynautsa z mjakyh golosnyh (je jotovani),a takozh nema spoluchen', de mjaka golosna bula b nastupnou za inshou golosnou (tobto preduvaty "mjakij" golosnij mozhe tilky prygolosna). Dla usih potreb mozhna obmezhytys' vzhyvannam odnogo specsymvola, bo navedeni nyzche movni sytu'acii mozhna vvazhaty vykluchnymy:

- aby pysaty inshomovni slova movou oryghinalu, dosyt' pochynaty ci slova z apostrofa. Podibnu sumish mozhna nazvaty cyvilizovanym surzhykom. Pryklad: Dla rozvytku c`ogo napr`amku stvoreno `site www.ukrlat.ua (Ce zajve dla sliv, scho mist`at' litery W X - Windows, text);

- aby vidrizn`aty u kilkoh des`atkah sliv spoluchenna "sh" vid shyplachykh mozhna mizh (s c z) ta h takozh pysaty apostrof:

s`hotiv s`hud s`hvalno S`hid s`hema

- v abrevi`aturah, skladyenyh velykymy literamy, usi litery slid vvazhaty okremymy, Pryklady:

ZMI SShA SRSR UNIAN

- sche odne za`uvazhenna. Z tablyci 2 vydno, scho proponovana vidpovidnist mae rozbizhnosti z orientovanymy na anglijsku transkrypciu pravylamy vidobrazhenn`a latynskymy literamy ukrainskyh vlasnyh imen ta ge`ographichnyh nazv. Porivn`alnyj pryklad:

Khmelnitsky - Hmelnyckyj

Povertauchys do vlasnyh imen, navedemo napysanna prizvyscha Хижняк - Hyzhn`ak.

Detalnishe pro specsymvolovy. Na latynskij chastyni klaviatury prysutni dva vydy specsymvoliv:

- ti, scho potribni dla zvyhajnoi gramatyky - " () ! ? . , :

- ti, scho potribni dla inshyh gramatyk (kompjuternyh)

~ @ # \$ % ^ & * ' = + [] { } | \ / < >

Perelichymo usi vzhyvanna apostrofa:

- jak mjakyj znak: kin`, sv`ato;

- mizh golosnymy (ne dodavaty pry vymovi "j"): na`uka zo`ologia;

- mizh prygolosnymy (c s z) ta "h", aby vidriznyty vid shyplachyh s`hid, s`hud, roz`hytav, zap`hnuv;

- pislta tverdoi "l": mal`a kula;

- na pochatku inshomovnyh sliv v oryiginalnomu napysanni: `life.

Na klavi`aturi je dva apostrofy - pravyyj (na odnij klvishi z symvolom ") i livyyj (- z symvolom ~). Mabut`, docilno oby dva apostrofy vvazhaty rivnoznachnymy.

Inshi aspekty translitu

Osoblyva rol litery "h".

Litera "h" moze buty jak samostijnou fonetychnou odnyceu, tak i utvor`uvaty skladni zvuky u spoluchenni z inshymy prygholosnymy. Samostijnomu vzhyvannu vidpovidae zvuk "x": eh, usih, nathnenn`a.

U spoluchennah iz prygholosnymy litera "h" moze vzhyvats` dla utvorenn`a novyh zvukiv: ch – ч, sh – ш, zh – ж, kh - x jak danyna isnuuchym shtampam – Kharkiv, ph - ф u slovah latynskogo pohodzhenna - photo, pharmacia, graphologia, ieroglyph. Vynykae pevna dvoznachnist pravopysu: text - tekst, _pharmacia - farmacia, bo podobni formy je phonetychno odnakovi. Mabut`, slid vvazhaty dozvolenymy obydvya napysanna sliv, za vyn`atkom vlasnyh imen ta nazv. Blyz`ke do c`ogo pytanna - pravylo perenosu: ne slid rozryvaty perenosom lexemy ja je jo ju ji ch sh zh sch

Bilshist jevropejskyh mov vykorystovue latynsku abetku ne u chystomu vygladi, a z pevnymy domishkamy naci`onalnogo zabarvlenna (dosyt` zgadaty umlauty nimeckoi, jusy polskoi, toscho). Otozh, teper ukrainske latynske pys`mo stae u c`omu vidnoshenni na riven` anglijskoi chy italijskoi mov, oderzhuuchy pevnu perevagu nad reshtou jevropejskyh mov. Hocha malo scho zavazhae i tij reshti krain vykonaty kosmetychni reformy pravopysu z metou pryvedenna svoih abetok do kanonichnoi latynyci. Napryklad, u nimeckij movi isnie alternatyvnyj pravopys – zamist' "a", "o", "u" z dvoma krapkami zverhu (umlautiv) mozhna pysaty vidpovidno "ae", "oe", "ue".

Nedaleko vid pochatkovoї tezy pro vytoky problemy je misce i takym mirkuvannam. Pisle perehodu na latynyc`u sproshuetsa lokalizacia program dla kompjuiteriv ta aparatyry z mikroprocesornym keruvannan - prynteriv, kopiuvachiv, mobilnyh ta ofisnyh telephoniv, usiei skladnoi pobutovoї tehniky, menshe klopotu zi shryftamy, programamy rozpiznavanna textiv.

Algorytm transliteracii kyrylichnyh textiv ne je skladnyj, a z vidomyh program najbilsh blyz`kou je programa Shtirlitz, de z dopomogou nalashtovuvanyh s`hem transliteracii WIN -> LAT chy DOS-> LAT v kiLka stadij mozhna oderzhaty bazhanyj rezultat. Sprosheno ce vygladae jak

poslidovnist peretvoren` na zrazok z'ednatysya -> zjednatysya -> zjednatysa.
Zvorotne peretvorenn'a za skladnist`u malo vidrizn`aetsa vid pr`amogo.

Висновок. Практично існує кілька сфер вжитку українського трансліту. Можна побудувати чимало варіантів таблиць трансліту. У цьому дописі запропоновано стандарт українського трансліту на основі наведених критеріїв оптимальності.

Література

1. Блоха М. Використання латинської абетки для українського правопису, ж-л ComputerWorld Київ, №8, 1999.
2. Пилипенко О. Кодовая суматоха, ж-л Chip № 01, 2003.
3. Ж-л Chip № 07, 2002.
4. [Електронний ресурс]. – Режим доступу: translit.kh.ua/#gost779