

Вакуленко Максим Олегович

канд. фіз.-мат. наук, ст. наук. співроб.,
завідувач відділом науково-дослідної роботи,
Державна науково-технічна бібліотека України,
м. Київ, Україна

УКРАЇНСЬКА ЛАТИНИЦЯ В СУЧАСНИХ ІНФОРМАЦІЙНИХ І ЛІНГВІСТИЧНИХ ТЕХНОЛОГІЯХ

Важливим аспектом цифровізації України є розвиток інформаційних і лінгвістичних технологій перетворення та передачі текстової інформації. Така інформація, створена в українському просторі, базується на кириличному письмі, в той час як світовий інформаційний простір є переважно латинографічним. Для потреб міжнародного спілкування держави з нелатинським письмом (Японія, Китай, Ізраїль, арабські країни, Росія, Монголія тощо) виробили відповідні національні стандарти транслітерації зі свого алфавіту на латиницю, які забезпечують виконання вимоги еквівалентності вихідного й транслітерованого записів. Прикладом такої залежності є відповідність між гаєвицею та вуковицею в країнах колишньої Югославії.

Зараз світ увійшов у добу інформаційної сингулярності, коли інтенсивність обміну інформацією зростає прискореними темпами. Це, своєю чергою, зумовлює інтенсивний розвиток різноманітних лінгвістичних технологій опрацювання мов, які використовують надвеликі текстові корпуси. Сучасні багатомовні текстові корпуси, які налічують мільйони одиниць, базуються на латиниці, а опрацювання цих корпусів відбувається в автоматичному режимі з застосуванням відповідного програмного забезпечення. Скрипти програмного забезпечення, яким користуються світові автоматичні системи опрацювання текстових даних, також використовують виключно літери латинського алфавіту. Усе це ще більше загострює потребу у відповідному транслітераційному стандарті.

Втім, Україна лишається єдиною державою з нелатинським письмом, яка не має відповідного національного стандарту транслітерації, що забезпечував би взаємнооднозначну відповідність між кириличною та латиничною формами представлення текстової інформації.

Відсутність такого стандарту вже має відчутний негативний вплив на розвиток і престиж української науки. Зокрема, автоматичні системи визначення індекса Хірша подають різні його значення для кириличного та

латиничного форм запису прізвища того самого науковця. Унаслідок цього українські науковці мають у світовій наукометрії занижені показники. А як відомо, цей індекс є ключовим міжнародним критерієм визначення рейтингу науковця, що впливає, зокрема, на розподіл світових грантових коштів.

Затверджена Постановою №55 Кабінету Міністрів України таблиця українсько-англійського транскодування порушує основну zásadu транслітерації – взаємнооднозначну відповідність між знаками кирилиці й латиниці. Це призводить до нееквівалентності вихідного й латинізованого текстів і відтак унеможливує відновлення вихідного запису. Наприклад, у такій системі Гальченко штучно “ототожнюється” з *Галченком*, Тронко – з *Троньком*, Банкова – з *Баньковою*, Паньківська – з *Панківською*, Польова – з *Половою*, Лялько – з *Ліалком*, Ліана – з *Ляною*, Маріан – із *Мар’яном*, Медіана – з *Медяною*, Возіанов – із *Возяновим*, Гундеріан – із *Гундеряном*, Клаузюс – із *Клазюсом*, Пії – з *Пієм*, Лар’їн – із *Ларіним*, Левитський – із *Левицьким*, Тоцька – з *Тотською*, Чернятський – з *Черняцьким* (див. <http://ukrlit.org/>). Через це така система неприйнятна ані для стандартизації, ані для сучасного спілкування, ані для застосування в сучасних банках даних, що передбачає автоматизовану латинізацію великих текстових корпусів і повне збереження інформації, яка в них міститься.

Українська «Київ» є штучно придумана форма «Kyiv» для позначення назви столиці України. Подібне поєднання фонем і відповідних графем є неприродним не тільки для української мови, а й для більшості світових мов. Саме штучність цього витвору і є одним із вагомих чинників, які зупиняють поширення українських графічних форм у міжнародному спілкуванні. Щобільше, передача літери «ї» латинською «i» передбачає для назви нашої держави запис “Ukraina” – що відповідає не українській, а російській латинографічній нормі. Навряд чи це сприяє зміцненню іміджу України як незалежної держави.

У світовій практиці існує явище орфографічної трансплантації [1, с. 25] – коли назва з однієї мови, з’являючись в оточенні іншої мови, зберігає свою автентичну форму. Таким чином, в англійській мові іншомовні назви *Hercules Poirot*, *Ajax*, *Juventus*, *Johannesburg*, *Jakobson*, *Jensen*, *Sarajevo*, *Ljubljana* тощо не змінюють свого написання.

Отже, згадана проанглійська система вступає в суперечність із загальноприйнятою практикою орфографічної трансплантації і порушує основну zásadu транслітерації – вимогу взаємнооднозначної відповідності між графемами оригіналу і знаками латиниці. Тому за своєю суттю вона не є транслітераційною.

Щоб відвернути увагу від численних порушень транслітераційного принципу, цю систему було запроваджено під приводом «легкого читання» – що, своєю чергою, також суперечить транслітераційному підходу: при транслітерації не ставлять завданням досягти «правильного» прочитання. Разом з тим, виникають реальні труднощі при прочитанні і впровадженні штучно створених форм (зокрема, «Київ»). З іншого боку, як показує світова практика, орфографічно трансплантовані слов'янські назви *Czech, Sarajevo, Ljubljana, Jelinek, Jablonec* помітних труднощів у прочитанні не викликають.

Хибною є орієнтація на англійську мову як «посередника» для латинізації українських назв. З одного боку, це аналог «старшого брата», що не сприяє зміцненню авторитету України як незалежної держави. З іншого боку, за кількістю носіїв мови (людей, для яких дана мова є рідною) англійська мова посідає четверте місце в світі – після китайської (діалект мандарин), іспанської та хінді. Таким чином, спотворення мовної картини світу на англійський лад не має також і достатніх соціолінгвістичних підстав. Китайці, іспанці та латиноамериканці підтримують свою, а не чужу мову, і тому не прагнуть записувати свої назви таким чином, щоб полегшити їх прочитання за правилами англійської мови: *Xinhua, Xi'an, Guadalajara, Mexico, San José*.

Зазначені вади псевдотранслітераційної таблиці, запровадженої Постановою №55 КМУ, добре помічають іноземні партнери України. І вказують на них. Наприклад, у 2013, 2014 і 2015 рр. європейська програма Erasmus Mundus давала авторові гранти на розвиток і впровадження української латиниці, яка цих вад не має.

Взаємнооднозначну систему української латиниці, яка містить три транслітераційні таблиці, враховує історичний розвиток української мови і має комп'ютерну реалізацію, було розроблено ще в 90-х роках минулого століття [2, с. 234-254]. 18 жовтня 1995 року Держстандарт України затвердив принципи транслітерації, на яких повинен базуватися національний транслітераційний стандарт. Ці принципи забезпечують взаємнооднозначну відповідність вихідного і транслітерованого текстів, а також гармонізацію з прийнятим у Європі міжнародним транслітераційним стандартом ISO: 9. 16 листопада 2000 року цю систему схвалила Транслітераційна комісія, створена розпорядженням Президента НАН України Б. Патона і складена з фахівців Національної академії наук і Академії наук вищої школи України. Цю систему вже успішно використовує ряд ресурсів: домен *ua* (перший стандарт транслітерації), сайт *translit.kh.ua* (влаштований стандарт), сервіс *Google Code* на <http://code.google>.

com/p/cyr2url/, онлайн-словники *sum.in.ua*, *rytu.in.ua* тощо. Цю систему необхідно впровадити для використання в сучасних інформаційних і лінгвістичних технологіях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Суперанская А. В. Теоретические основы практической транскрипции. – М. : Наука, 1978. – 284 с.
2. Вакуленко М. О. Українська термінологія: комплексний лінгвістичний аналіз : [монографія]. – Івано-Франківськ : Фоліант. – 2015. – 361 с., іл.