

## ТЕОРЕТИЧНІ ПИТАННЯ МОВИ

М. О. Шведова (Київ)

УДК 81'26+81'28

### ПРОБЛЕМА ЛІТЕРАТУРНОЇ НОРМИ УКРАЇНСЬКОЇ МОВИ З ОГЛЯДУ НА ЇЇ РЕГІОНАЛЬНІ ВАРІАНТИ (ЗА РЕГІОНАЛЬНО АНОТОВАНИМ КОРПУСОМ УКРАЇНСЬКОЇ МОВИ)

Літературна мовна норма – це сукупність мовних засобів і правил їх вибору й використання, яка історично склалася в певному суспільстві та зафіксована в нормативних словниках і граматиках. Сумніви щодо нормування того чи того мовного явища виникають там, де система мови нежорстка, допускає варіативність. На формування літературної мовної норми впливають як природній розвиток мови, її внутрішні процеси, так і нелінгвістичні, суспільні чинники. Укладачі літературної норми враховують частотність мовного явища, поширення його, уживаність у літературних текстах тощо. Але також існує деякий, інколи значний, суспільний, політичний тиск на нормування мови. Якщо в післявоєнні радянські часи діяла відома політика наближення української літературної мовної норми до російської, то в наш час навпаки, є тенденція до їх розмежування.

В українському мовознавстві, як і в російському, традиційно дотримувались думки про необхідність жорсткої стандартизації мови. Як зазначав І. Огієнко, „для одного народу – одна літературна мова, один правопис” [Єрмоленко 2000]. Літературні норми української мови розробляли О. Курило, Б. Антоненко-Давидович, М. Жовтобрюх, М. Пилинський, А. Коваль та багато інших вчених.

Сучасна лінгвістика приділяє більше уваги вивченню мови в її різноманітних проявах і варіантах, ніж власне нормуванню мови. У другій половині ХХ – ХХІ ст. вивчають особливості розмовної мови, мову різних соціальних спільнот, мову міст, мову діаспори, мову Інтернету як письмовий різновид усної мови, мовні помилки тощо. Корпусні дослідження дають можливість отримати найбільш об'єктивні відомості щодо мовних явищ, їх розвитку, уживаності, частотності.

Уважаємо, сучасні корпусні методи мають бути використані для нормативного опису української мови, виявлення наявних варіантів, зокрема регіональних, уточнення старих норм, які часом є дуже жорсткими й не завжди підтверджуються в разі перевірки за великим обсягом текстів.

Ми розробили корпус української мови, розмічений за регіонами. У ньому 6255 текстів загальним обсягом понад 100 мільйонів слівформ. Кожному тексту приписано таку інформацію: ім'я автора, назву, жанр, рік написання (для перекладів – рік створення українського перекладу),

якщо текст перекладний, також ім'я перекладача та мову оригіналу. Кожному авторові приписано один або кілька регіонів, де він народився, вчився чи жив тривалий час. Перекладному тексту в корпусі відповідають регіони, які приписано перекладачеві. Частка перекладів у корпусі – приблизно 30%. Корпус містить переклади з 38 мов, найбільше – з англійської (8,5%) та російської (7,4%).

За жанром найбільша частка – це художні тексти (майже 60%), решта (приблизно 40%) – це публіцистичні, наукові й навчальні тексти (з історії, права, філософії тощо).

За часом написання текстів корпус охоплює майже 200 років – від 1819 до 2017 р.

В основу розмітки корпусу за регіонами покладено сучасний адміністративний поділ України. За областями тексти розподіляються нерівномірно. Один текст може належати одночасно до кількох регіональних підкорпусів: якщо автор мешкав тривалий час у різних місцях, йому приписували два чи більше регіонів. Багато людей переїжджали з різних місць до Києва, тому київський підкорпус найбільший – 60% від загального обсягу текстів, але він найменш інформативний з погляду регіональної мовної інформації. Серед інших підкорпусів найбільший за обсягом підкорпус текстів Львівської області – майже 20%; 9% Полтавської; приблизно по 5% Івано-Франківської, Житомирської, Волинської, Київської, Дніпропетровської, Тернопільської та Сумської областей; менша частка інших. У корпусі представлені тексти з усіх областей України та з Криму.

У випадках, коли автор тривалий час жив за кордоном або емігрував, йому приписували також іншу країну. Отже, маємо чималий підкорпус текстів авторів діаспори. Тексти США, Канади, Польщі, Німеччини, Великої Британії, Франції здебільшого належать емігрантам 40-х років, менша частка – емігрантам 20-х рр. Обсяг корпусу української мови діаспори – понад 10 млн. словоформ.

Для пошуку за корпусом використовуємо пошукову систему diskMETA-Personal. Це програма для локального пошуку, яка шукає за лексемою, словоформою та точною фразою. За результатами такого пошуку можна рахувати частотність слів, словоформ і сталих виразів, порівнювати таку частотність у різних областях.

За корпусом було перевірено кілька відомих нормативних рекомендацій щодо української мови.

#### ***Добрий день vs. доброго дня***

Академічний словник української мови [СУМ] фіксує варіанти з називним і родовим відмінком: „Добрий ранок (день, вечір); Доброго ранку (дня, вечора) – традиційне взаємне дружнє вітання при зустрічі вранці (удень, увечері)”. Олександр Пономарів рекомендує інше: „Щодо формул вітання я вже неодноразово писав: *доброго ранку, добрий день (добридень), добрий вечір (добривечір), добраніч (на добраніч)* – це давно усталені формули. Гадаю, їх не варто міняти на штучні: *доброго дня,*

*доброго вечора*. Так можна казати, коли при цих сполуках маємо ще дієслово: *бажаю вам (тобі) доброго дня, доброго вечора*” [Пономарів 2014]. Деякі популярні видання, напр., газета „Факти” в рубриці „Уроки української”, поширюють рекомендацію О. Пономарева вже як правило: *Доброго дня!* – неправильно [Горская 2013]. Але пересічні носії мови висловлюють і протилежну думку, напр., з дискусії у Фейсбуці: „Коли я вітаюся „доброго дня” я ніби зичу людині доброго дня, так би мовити благословляю її, бажаю того, щоб день таким вдався”.

У текстах корпусу української мови знаходимо всі варіанти: *Добрий ранок (день, вечір)* та *Доброго ранку (дня, вечора)*. Для ранку й вечора основним є один варіант: *доброго ранку* (296 разів, *добрий ранок* тільки 41) та *добрий вечір* (257 разів, *доброго вечора* – 67). Різниця в кількості вживання між виразами *добрий день* і *доброго дня* менша, хоча переважає дійсно *добрий день* (301 і 121 раз). Співвідношення частоти вживання виразів *добрий день* і *доброго дня* в різних регіонах різне: у Полтавській, Дніпропетровській, Донецькій областях кількість знайдених виразів *добрий день* і *доброго дня* майже однакова (54% та 46%: у Полтавській – 17/14, Дніпропетровській – 12/10, Донецькій – 13/11); у Львівській, Івано-Франківській значно переважає вираз *добрий день* (77%, це більше, ніж середнє значення по корпусу; у Львівській обл. співвідношення 42/12, в Івано-Франківській – 17/5). Тобто норма, яку пропонують сучасні мовознавці, тяжіє до західного варіанта мови, а східний не беруть до уваги. Уважаємо, обидва варіанти можна вважати нормативними й рекомендувати до вживання.

### ***Кілька vs. пара***

Б. Антоненко-Давидович пише: „«Позич мені на пару днів гроші»; «Зустрінемося через пару днів», – чуємо інколи в розмовах, а зрідка читаємо й на сторінках сучасних художніх творів. Так сказати по-українському не можна, а слід: «Позич мені на кілька день (днів) (на два-три дні, на які два дні) гроші»; «Зустрінемося через кілька день (днів) (через два-три дні, через які два дні)» [Антоненко-Давидович 2010: 156 – 157]. У текстах корпусу є слово *пара* у значенні ‘кілька’: „Якогось там коржа з білої муки, якихось розбити *пару* яєць...” (Улас Самчук „Юність Василя Шеремети”, 1946); „*На пару днів* поїду” (Олексій Коломієць „Дикий ангел”, 1978); „*На веранді, взагалі то, крім канати, магнітофона та ще столика з парою стільців, більше нічого й немає*” (Любка Дереш „Поклоніння ящірці”, 2000); „– *А скільки це триватиме? – питається Какао. – Та пару годин, може, трішки довше*” (Сергій Жадан „Депеш Мод”, 2004). Академічний словник української мови [СУМ] подає слово *пара* в значенні ‘кілька’ навіть без позначки *розм[овне]*: „9. Невелика кількість чого-небудь; декілька”. У прикладах, які наведено в словнику, ужито *пара день [днів]* (Гнат Хоткевич) і *пару літ* (Григорій Тютюнник); за ромбом подано вираз *пару слів (фраз) написати (сказати і т. ін.)* – коротко написати, сказати про кого-, що-небудь. Ми пошукали за регіональними підкорпусами вирази: „*пару*

днів”, „пару годин”, „пару місяців”, „пару тижнів”, „пару років”, „пару хвилин”, „пару разів” і „пару слів”. Таких було знайдено загалом 599, вони наявні в текстах усіх регіонів України, а також у текстах української діаспори. Це менше, ніж кількість виразів „кілька днів”, „кілька годин”, „кілька місяців”, „кілька тижнів”, „кілька років”, „кілька хвилин”, „кілька разів” і „кілька слів”, але різниця не дуже велика.

кілька днів	156	пару днів	129
кілька годин	155	пару годин	133
кілька місяців	110	пару місяців	89
кілька тижнів	75	пару тижнів	69
кілька років	224	пару років	159
кілька хвилин	169	пару хвилин	137
кілька разів	231	пару разів	163
кілька слів	146	пару слів	127

Уважаємо, обидва синоніми можна вважати нормативними й рекомендувати до вживання.

Отже, корпусні технології можуть бути застосовані до багатьох типів лінгвістичних завдань, зокрема до нормування мови, укладання словників і граматик. За допомогою корпусу можна в багатьох випадках уточнити наявні норми літературної мови та рекомендації з культури мови. Інколи вони виявляються застарілими, або такими, що орієнтовані на певний регіональний варіант мови й не враховують варіативного вживання на іншій великій території, або такими, що фактично не відповідають мовній практиці і потребують коригування.

#### Література

**Антоненко-Давидович 2010** – Антоненко-Давидович Б. Д. Як ми говоримо / Б. Д. Антоненко-Давидович. – К.: Книга, 2010. – 252 с.; **Горская 2013** – Горская Д. Доктор филологических наук Наталия Шумарова: Нині серед молоді стало популярним на прощання казати „спишемося” чи „зідзвонимося” // Факти. – 19.02.2013. – Режим доступу : <http://fakty.ua/158517-doktor-filologicheskikh-nauk-nataliya-shumarova-nini-sered-molodi-stalo-populyarnim-na-procshannya-kazati-spishemosya-chi-zidzvonimosya>; **Єрмоленко 2000** – Єрмоленко С. Я. Норма мовна // Українська мова : енциклопедія. – К.: Вид-во „Українська енциклопедія” ім. М. П. Бажана, 2000. – 752 с.; **Пономарів 2014** – Пономарів О. Блог професора Пономарева // BBC. Україна. – 2014. – Режим доступу : [http://www.bbc.com/ukrainian/blogs/2014/04/140430\\_ponomariv\\_blog17\\_ko](http://www.bbc.com/ukrainian/blogs/2014/04/140430_ponomariv_blog17_ko); **СУМ** – Словник української мови : в 11 т. – К., 1970 – 1980. – Режим доступу : <http://sum.in.ua/>.

**Шведова М. О. Проблема літературної норми української мови з огляду на її регіональні варіанти (за регіонально анотованим корпусом української мови)**

У статті окреслено завдання верифікації нормативних тверджень, що пов'язані з вибором між тими чи тими синонімічними засобами вираження, на матеріалі літературного узусу за допомогою корпусних методів. Авторкою укладено корпус текстів української мови, який охоплює 200-річний період. Корпус розмічено за ознакою територіального походження авторів (у випадках залучення до корпусу перекладних текстів – перекладачів). У статті розглянуто кілька конкретних сюжетів (*добрий день* vs. *доброго дня*, *кілька* vs. *пара*), де перший синонім певні авторитетні посібники визнають нормативним, а другий таким, що з тих чи тих причин перебуває за межами літературної норми. За результатами квантитативного аналізу корпусу виявлено, що частотність уживання цих варіантів у літературних текстах фактично порівнянна. Для першої пари виявлено певні регіональні переваги. Висловлено думку, що нормативні рекомендації, пов'язані з цими словами, можуть бути скореговані.

*Ключові слова:* українська мова, норма, корпусна лінгвістика, регіоналізми, варіанти, синоніми, лексика.

**Шведова М. А. Проблема литературной нормы украинского языка с точки зрения его региональных вариантов (по регионально аннотированному корпусу украинского языка)**

В статье ставится задача верификации нормативных утверждений, связанных с выбором между теми или иными синонимичными средствами выражения, на материале литературного узуса при помощи корпусных методов. Автором подготовлен корпус текстов украинского языка, охватывающий 200-летний период и размеченный по признаку территориального происхождения авторов (в случае включения в корпус переводных текстов – переводчиков). В статье рассмотрен ряд конкретных сюжетов (*добрий день* vs. *доброго дня*, *кілька* vs. *пара*), где первый синоним признается некоторыми авторитетными пособиями нормативным, в то время как второй выводится за рамки украинской литературной нормы по тем или иным основаниям. На основании квантитативного анализа корпуса выясняется, что частотность употребления этих вариантов в литературных текстах фактически сопоставима. Для первой пары выявляются определенные региональные предпочтения. Представляется, что нормативные рекомендации в связи с этими словами могут быть скорректированы.

*Ключевые слова:* украинский язык, норма, корпусная лингвистика, регионализмы, варианты, синонимы, лексика.

**Shvedova M. O. The problem of the normative Standard Ukrainian in the context of its regional variation (based on a regionally annotated corpus of Ukrainian)**

The paper's aim is to test against the benchmark of a regionally annotated corpus Ukrainian standard texts some normative claims made by renowned authors of normative handbooks. The norms are thus compared to the observed use of synonymous forms. The author has prepared a comprehensive corpus of (Standard) Ukrainian literary texts, encompassing a 200-year period since the beginning of the Modern Ukrainian language, including both texts written originally in Ukrainian or translated from other languages. These texts are annotated by the origin of the authors (their region of birth or literary activity) or the translators. The paper discusses some case studies, concerning lexical items (*dobryj den'* or *dobroho dnja* 'good afternoon! [greeting]', *kil'ka* vs. *para* 'some, a few'). Some authors deem only the first element of each pair to be an element of the standard language while totally discarding the other on various reasons. Our quantitative study shows that the frequencies of both elements of the pairs in the standard literary texts are in fact comparable. Moreover, the first synonymous pair has rather clear tendencies of distribution between different regions of Ukraine, viz. *dobryj den'* is an element of the Western norm of Ukrainian language (with *dobroho dnja* more frequent in the East).

*Key words:* Ukrainian language, language norm, corpus linguistics, regional variants, synonyms, lexicon.

*Стаття надійшла до редакції 10.10.2017 р.*

*Прийнято до друку 11.10.2017 р.*

*Рецензент – канд. філол. н., доц. Шутова Л. І.*