

## КОРПУС ПЕРСЬКОЇ МОВИ ЯК ДЖЕРЕЛО МАТЕРІАЛУ ДЛЯ НАВЧАЛЬНИХ СЛОВНИКІВ-МІНІМУМІВ

Ольга МАКСИМІВ

*Львівський національний університет імені Івана Франка,  
вул. Університетська, 1, кімн. 239, 79000, Львів, Україна,  
тел.: (00380 32) 2394704, e-mail: persochka@yahoo.com*

Проаналізовано джерела, які використовували для відбору слів до навчальних словників та лексичних мінімумів, визначено місце принципу частотності під час цього відбору та обґрунтовано відповідність корпусу перської мови для укладання частотних словників різних типів і навчальних словників-мінімумів.

*Ключові слова:* викладання іноземних мов, джерела відбору лексики, частотність, корпус текстів, частотний словник, навчальний словник-мінімум.

На сучасному етапі викладання іноземних мов питання якісного і кількісного відбору навчального матеріалу актуальні викладачів, що працюють з іноземцями, методистів та лінгвістів. Надзвичайну увагу в системі вивчення іноземної мови приділяють лексиці, тому раціонального відбору потребує перш за все словниковий склад. Академік Л. В. Щерба [34, с. 290] зазначав, що для кожної іноземної мови необхідним є навчальний словник для початківців. Статистика ж свідчить, що перша сотня найчастотніших слів покриває приблизно половину середньостатистичного тексту. Однак дані частотного словника об'єктивно відображають лексику мови лише тією мірою, якою є об'єктивним відбір джерел для цього словника. Їхня надійність залежить від кількісної і якісної характеристики лексичного масиву для підрахунку. Тому на сучасному етапі першочергового значення набуває чітке формулювання головних принципів і способів кількісного аналізу фактів мови і мовлення.

Метою нашої розвідки є теоретичне обґрунтування корпусу перської мови як джерела для укладання навчальних словників-мінімумів.

Під час укладання частотних списків різних мов дослідники використовують різні джерела.

У процесі укладання навчального словника-мінімуму для студентів медичних вищих навчальних закладів використовували підручники зі спеціальних предметів, художні тексти, науково-популярні статті, записи телепрограм, дані анкет, словники (тлумачні, іншомовних слів, частотний та інші словники-мінімуми, короткі текстові словники, видані в 90-х роках, словник за редакцією професора В. В. Морковкіна) [див.: 8, с. 43]. Своєю чергою, словник за редакцією професора В. В. Морковкіна відібраний шляхом особливого суміщення даних восьми частотних словників російської мови [див.: 12, с. 9–10].

Як джерела для відбору лексики до єдиного лексичного мінімуму російської мови для національної середньої школи використовували [див.: 2]:

- узагальнений список слів частотних словників російської мови, складений на основі різноманітних матеріалів;
- лексичний матеріал великої кількості словників-мінімумів для неросійських учнів та коротких двомовних словників, на основі яких також створили узагальнений лексичний список;
- частотні списки лексики художніх творів, а також частотні списки, складені на основі послідовного розписування художніх та навчальних текстів хрестоматій з російської літератури.

Для лексичного підґрунтя курсу російської мови використовували програми з російської мови, посібники, словники, частотні списки [див.: 30, с. 38]. При укладанні “Словника найуживаніших слів німецької мови” – частотні словники і матеріали інших словників [див.: 17, с. 5]. Для укладання “Англо-українського навчального словника-мінімуму” бралися слова з частотних словників, підготованих із навчальною метою [див.: 23]. Це слова, які входять до більшості підручників з англійської мови як іноземної, виданих у США, Англії та Німеччині, а також до різних лексичних мінімумів. Тематично лексика відповідає програмам і літературі для позакласного читання.

Проаналізувавши, які джерела використовували для відбору слів до навчальних словників та лексичних мінімумів можна помітити, що це переважно частотні словники певної мови [див.: 11, 27, 29, 33]. Окрім того, словники-мінімуми, короткі словники, частотні списки лексики художніх творів, лексика з підручників, програми та посібники. Для укладання спеціалізованих словників ще залучали лексику науково-популярних статей та словника іншомовних слів. Також використовували записи усного мовлення [див.: 21, 22, 28].

Головна розбіжність методистів та авторів проаналізованих навчальних словників і словників-мінімумів полягає в тому, що вони по-різному ставляться до принципу частотності. Аналіз різних поглядів на цей принцип дає змогу говорити про декілька основних позицій авторів, а саме:

1. Визнання частотності за визначальний принцип та доповнення даних частотного словника тематично відібраними словами [див.: 10]. Це не може бути достатнім, скажімо, для українсько-перського навчального словника-мінімуму. По-перше, тому, що в українській лексикографії немає частотного словника, який би репрезентував усі (чи хоча б найважливіші) сфери мовленнєвої діяльності. По-друге, для вивчення мови замало мати лише набір частотних та тематично цінних лексем. Необхідно, щоб достатньою мірою репрезентувалася граматична структура мови.

2. Визнання частотності головним принципом та подальші операції над верхньою межею частотного (чи частотних) словника з урахуванням низки інших принципів [див.: 2–7, 18, 28]. Можна погодитися з тим, що “статистичний аналіз є важливішим для попереднього поділу словника мови на головне і другорядне, але переоцінювати цей критерій не можна” [20, с. 5]. Застосування поряд із частотністю інших принципів потребує детального їх аналізу.

3. Визнання частотності головним принципом у процесі відбору лексики для читання та додатковим у відборі лексики для усного мовлення (чи говоріння) та писання [див.: 15, 16, 22, 23]. Це є дещо штучним. Визначальним аргументом тут є те, що у першому випадку йдеться про пасивний словник учня, а в другому – про активний, і об’єм пасивного словника є більшим. Оскільки вся лексика навчального словника-мінімуму призначена для активного оволодіння, то й відбирати її треба на

основі різних критеріїв, одним із яких є частотність. Зауважимо, що причиною недовіри до частотних словників є відсутність задовільної методики їх укладання [див.: 14, с. 73].

4. Визнання частотності лише як “обмежувача” у відборі лексики за іншими принципами [див.: 8, 24, 25, 31]. Такий підхід вимагає аналізу всього словникового складу мови з погляду інших принципів, з подальшим додаванням деяких слів із частотного списку. На нашу думку, доцільніше робити навпаки. Адже з’ясовано, що запас у 2000 слів, які вживаються найчастіше, забезпечує розуміння звичайного (неспеціального) тексту на 85% і дає змогу вести бесіду на будь-яку (неспеціальну) тему [див.: 32, с. 330]. Ці дані повинні переконати кожного філолога у надзвичайній важливості принципу частотності.

5. Визнання частотності за один із складників поняття “вживаності слова у мовленні” [див.: 26]. Це положення компенсується застосуванням низки інших принципів та відповідним відбором джерел для статистичного аналізу. На користь цього свідчить і такий факт, що різні автори поняття “вживаності слова” трактують по-різному. Наприклад, професор В. А. Бухбіндер вважає, що головні чинники, які впливають на вживаність лексичної одиниці, полягають у її семантичній вартості, словотворчій здатності, багатозначності, сполучуваності, структурній здатності, частотності та стилістичній диференційованості. П. М. Алексеев зазначає, що під уживаністю слова варто розуміти здатність слова траплятися у певній кількості джерел (текстів) з певною частотою, тобто мається на увазі частотно-рангова характеристика слова.

Під час укладання частотного словника велике значення має тип джерел, які відбирають для статистичної обробки. Статистика є точним інструментом дослідження, однак вона дає об’єктивну інформацію лише про той матеріал, який аналізували [див.: 1, с. 7]. Суб’єктивні рішення у відборі матеріалу для аналізу неминуче відображаються і на результатах.

На сучасному етапі розвитку мовознавства активно розвивається корпусна лінгвістика, укладаються корпуси різних мов. Їхні розробники намагаються представити користувачам функціонування описуваної мови якомога повніше і різноманітніше [див.: 9, с. 74].

Доступний в Інтернеті [35] корпус текстів перської мови, укладений на основі сучасної перської мови – мови останніх 75 років [див.: 36]. Містить як писемну мову, так і розмовну, мову різних стилів і різні підмови. За жанровим розподілом приблизно половину текстів становить художня література. А „твори художньої літератури найчастіше мають своїм ядром весь ступінь повсякденного спілкування” [19, с. 356]. Решта текстів належать до публіцистичного та наукового стилів, також є зразки розмовного стилю.

Дослідник має змогу укласти частотний список, використовуючи зразки як розмовної (виклад новин), так і писемної мови (прози і поезії), різних стилів (художнього, розмовного, публіцистичного) і різних підмов (медицини, спорту, мистецтва і т. ін.). Можна вибрати усі доступні опції, а можна укласти частотний список за підкорпусами одного типу. Частотний словник може бути за спадом або за зростанням частот, у прямому або зворотному алфавітному порядку. Також можна обмежити кількість слів (це актуально, перш за все, для лінгводидактичних потреб).

Збалансованість складу джерел будь-якого корпусу є надзвичайно серйозною і болючою проблемою [див.: 13]. Укладачі корпусу перської мови використали по

можливості усі типи письмових текстів, презентованих у мові, усні тексти, правда, дещо обмежено, також представлені. Вимогу випадковості вибірки для частотного словника забезпечили укладачі корпусу, відібравши до нього різнотипні тексти. Якщо цього не достатньо, дослідник може самостійно, методом випадкових чисел, відібрати відрізки однакової довжини з кожного або з деяких (випадково обраних) доступних текстів (усього 1500 повних текстів), що займе значно менше часу, аніж самому формувати такі вибірки вручну.

Під час укладання частотних списків подається уся необхідна статистична інформація: кількість слів у тексті, кількість словоформ, відносна і абсолютна частота слова, його ранг. Інші статистичні дані можна вирахувати на підставі поданих. Отож, тут дотримано відповідності між одиницею вимірювання довжини підкорпусу і одиницею статистичного списку явищ. Складні дієслова тут подано як два (чи більше) окремі слова. Власні назви залучено на загальних підставах як окремі слова. При бажанні їх майже завжди можна відокремити і вилучити.

Одержання доступу до іншої за тематичним спрямуванням версії корпусу допоможе укласти базові списки не лише для навчальних словників-мінімумів з перської мови як спеціальності, а також і для інших спеціальностей, оскільки укладачі задекларували, що „джерелами відбору даних до Корпусу є достатні зразки загальних і спеціальних видань з різних сфер і галузей” [37].

Отже, доступний в Інтернеті корпус текстів перської мови навіть у його демо-версії може слугувати джерелом для укладання:

- 1) корпусу текстів для статистичної обробки;
- 2) частотних списків різних типів;
- 3) частотних словників різних мовних стилів чи підстилів;
- 4) базових списків для навчальних словників-мінімумів для різних потреб.

Можна укласти такі словники для осіб, які вивчають перську мову:

- а) як спеціальність;
- б) як допоміжну для іншої спеціальності (журналістика, філософія, економіка).

- 
1. *Алексеев П. К* вопросу о типологии частотных словарей / П. Алексеев // Актуальные вопросы преподавания русского языка и литературы : Тезисы докладов и выступлений. – М., 1969. – С. 7–8.
  2. *Даунене З. П.* Специфика единого лексического минимума русского языка для национальной средней школы / З. П. Даунене // Проблемы учебной лексикографии и обучения лексике : [сб. ст. под ред. П. Н. Денисова и В. В. Морковкина]. – М., 1978. – С. 132–137.
  3. *Денисов П. Н.* Лексика русского языка и принципы ее описания / П. Н. Денисов. – М., 1980.
  4. *Денисов П. Н.* Лексические минимумы русского языка / П. Н. Денисов // Лексические минимумы русского языка / [под ред. П. Н. Денисова]. – М., 1972. – С. 3–15.
  5. *Денисов П. Н.* От редактора // Лексические минимумы русского языка / П. Н. Денисов / [под ред. П. Н. Денисова]. – М., 1972. – С. 122.
  6. *Денисов П. Н.* Очерки по русской лексикологии и учебной лексикографии / П. Н. Денисов. – М., 1974.
  7. *Денисов П.* Принципы отбора лексического материала / П. Денисов // Актуальные вопросы преподавания русского языка и литературы : Тезисы докладов и выступлений. – М., 1969. – С. 116–119.

8. *Журомская Л. И.* О проекте учебного словаря-минимума для студентов медицинских вузов / Л. И. Журомская // Филологические студии. – 2001. – №2. – С. 41–45.
9. *Карпіловська Є. А.* Вступ до комп'ютерної лінгвістики / Є. А. Карпіловська. – Донецьк : ТОВ „Юго-Восток, Лтд”, 2003. – 184 с.
10. *Кедайтене Е. И.* Об одном типе словаря для начального обучения / Е. И. Кедайтене // Актуальные проблемы учебной лексикографии / [сост. В. А. Редькин]. – М., 1977. – С. 173–179.
11. *Китайгородская Г. А.* Методика интенсивного обучения иностранным языкам / Г. А. Китайгородская. – М., 1986.
12. Комплексный учебный словарь “Лексическая основа русского языка” и его построение // Лексическая основа русского языка : Комплексный учебный словарь / [В. В. Морковкин, Н. О. Бёме, И. А. Дорогонова и др.]; под ред. В. В. Морковкина. – М., 1984. – С. 9–35.
13. *Копотев М. В.* Национальный корпус русского языка / М. В. Копотев, Л. Янда // Вопросы языкознания. – 2006. – № 5. – С. 149–155.
14. *Коршунова Я. Б.* Статистические данные и отбор языкового материала (Обзор) / Я. Б. Коршунова // Проблемы отбора учебного материала / [под ред. М. Н. Вятютнева]. – М., 1971. – С. 71–79.
15. *Лapidус Б. А.* Об учебном словаре “русский язык – родной язык” / Б. А. Лapidус // Русский язык за рубежом. – 1972. – № 2. – С. 47–51.
16. *Лapidус Б. А.* Проблемы содержания обучения языку в языковом вузе / Б. А. Лapidус. – М., 1986.
17. *Любимова З. М.* Предисловие / З. М. Любимова // Словарь наиболее употребительных слов немецкого языка. – М., 1990. – С. 5–6.
18. *Мэки В.* Отбор // Проблемы отбора учебного материала / [под ред. М. Н. Вятютнева]. – М., 1971. – С. 27–70.
19. *Мюллер Г.* Лингвистическая статистика и структура языковых полей / Г. Мюллер // Методика преподавания иностранных языков за рубежом. – М., 1967. – С. 351–361.
20. *Новиков Л. А.* Учебная лексикография и ее задачи / Л. А. Новиков // Вопросы учебной лексикографии / [под ред. П. Н. Денисова и Л. А. Новикова]. – М., 1969. – С. 3–11.
21. Очерки методики обучения устной речи на иностранных языках / [руковод. кол. авт. В. А. Бухбиндер]. – К., 1980.
22. Очерки методики обучения чтению на иностранных языках / [под ред. проф. И. М. Бермана и проф. В. А. Бухбиндера]. – К., 1977.
23. *Передмова // Перебийніс В. І.* Англо-український навчальний словник-мінімум / В. І. Перебийніс, Е. П. Рукіна, Хідекель. – К., 1995. – С. 5–6.
24. *Рахманов И. В.* Предисловие // Словарь наиболее употребительных слов немецкого языка / [под ред. проф. И. Рахманова]. – М., 1967. – С. 5–36.
25. *Рахманов И. В.* Рецептивное и репродуктивное усвоение иностранного языка в средней школе / И. В. Рахманов // Иностранные языки в школе. – 1952. – № 4. – С. 42–56.
26. *Салистра И. Д.* Вопросы отбора лексики иностранного языка для средней школы / И. Д. Салистра // Вопросы методики обучения иностранным языкам в средней школе. – М., 1956. – С. 7–32.
27. *Скопина М. К.* К вопросу о выделении существительных для основного словаря русского языка / М. К. Скопина // Международная конференция преподавателей русского языка и литературы. Актуальные вопросы преподавания русского языка и литературы : [тезисы докладов и выступлений]. – М., 1969. – С. 317–319.

28. *Степащенко С. М.* Про відбір словника-мінімуму для інтенсивного навчання іноземної мови на початковому етапі мовного вузу / С. М. Степащенко // *Методика викладання іноземних мов.* – 1990. – Вип. 19. – С. 56–60.
29. *Фалькович М. М.* Предисловие / М. Фалькович // *Лексический минимум по английскому языку (для языковых вузов).* – М., 1972. – С. 3–8.
30. *Фокина Г. В.* Лексическая основа курса русского языка / Г. Фокина // *Из опыта преподавания русского языка нерусским.* – М., 1965. – С. 36–59.
31. *Фридман Х. Х.* О некоторых критериях отбора словаря минимума (СМ) для курса практики устной и письменной речи / Х. Х. Фридман // *Вопросы научной организации преподавания иностранных языков в специальном вузе. Пути внедрения результатов научных исследований в практику обучения* : [тезисы докладов и сообщений научно-теоретической конференции]. – Горький, 1970. – С. 144–146.
32. *Хойпель К.* Частотность и структуры в современной теории преподавания иностранных языков / К. Хойпель // *Методика преподавания иностранных языков за рубежом.* – М., 1967. – С. 330–332.
33. *Штейнфельдт Э. А.* Частотный словарь современного русского литературного языка. 2500 наиболее употребительных слов / Э. А. Штейнфельдт. – Таллин, 1963.
34. *Щерба Л. В.* Языковая система и речевая деятельность / Л. В. Щерба. – Ленинград, 1974.
35. <http://pldb.iics.ac.ir>  
[/http://pldb.iics.ac.ir/Files/PLDB-REP.pdf](http://pldb.iics.ac.ir/Files/PLDB-REP.pdf) / اینترنت در فارسی های زبان داده پایگاه 36  
عاصی مصطفی. پایگاه داده های زبان فارسی: منبعی مستند برای پژوهش های زبانی // مجموعه مقالات سومین کنفرانس 37  
عاصی مصطفی. پایگاه داده های زبان فارسی: منبعی مستند برای پژوهش های زبانی // مجموعه مقالات سومین کنفرانس  
زبانشناسی، تهران، دانشگاه علامه طباطبائی و پژوهشگاه علوم انسانی و مطالعات فرهنگی. – تهران، 1376. – صص  
. 211–205

## PERSIAN CORPORA AS SOURCE OF MATERIAL FOR EDUCATIONAL DICTIONARIES-MINIMUMS

**Olha MAKSYMIV**

*Ivan Franko National University of Lviv,  
Universytetska str., 1, room 239, 79000, Lviv, Ukraine,  
tel.: (00380 32) 2394704, e-mail: persochka@yahoo.com*

The article deals with the sources of selection of words to the educational dictionaries and lexical minimums. The attention is concentrated on the place of principle of frequency during this selection and accordance of the Persian Corpora for the preparation different types of frequency dictionaries and educational dictionaries-minimums.

*Key words:* foreign language teaching, sources of word selection, word frequency, text corpus, frequency dictionary, educational dictionary-minimum.

Стаття надійшла до редколегії 27.09.2007

Прийнята до друку 03.03.2008