

# ФІЛОЛОГІЧНІ НАУКИ

УДК 81'33=161.2(075)

**Л. М. Гриднєва,**  
кандидат філологічних наук, доцент,  
науковий співробітник лабораторії комп'ютерної лінгвістики  
(Київський національний лінгвістичний університет);

**В. І. Перебийніс,**  
доктор філологічних наук, професор  
професор-консультант кафедри германської і фінської філології  
(Київський національний лінгвістичний університет)

## ФОРМУВАННЯ РЕЄСТРУ АНГЛО-УКРАЇНСЬКОГО НАВЧАЛЬНОГО СЛОВНИКА

*Статтю присвячено вивченню формування реєстру англо-українського навчального словника. Зазначається, що цей реєстр повинен відповідати програмам іноземної мови на певному етапі її викладання, включати обмежену кількість часто вживаних слів з високою словотвірною та сполучувальною потенцією. Зроблено висновок, що при його формуванні слід спиратися як на лексичні мінімуми, так і на частотний словник.*

Призначення навчальних словників – забезпечити навчальний процес вивчення іноземної мови. Це означає, що словники повинні відповідати програмі викладання її у школі чи вищому навчальному закладі. Зрозуміло, що програми з іноземної мови залежать від рівня викладання та спеціалізації навчального закладу. Отже, повинні створюватися серії навчальних словників для різних рівнів вивчення мови та різних видів [1].

Як правило, виділяють три рівні вивчення іноземної мови: початковий, середній і просунутий. Тому повинно бути принаймні три навчальних словники: для першого, другого та третього рівнів. Вони відрізняються один від одного перш за все обсягом реєстру, а також деталізацією словникової статті. У навчальній лексикографії визначено, що обсяг реєстру словника першого рівня – 2000-2500 слів, другого – до 5000 слів, третього – до 10 000 слів. Саме таким обсягом реєстру характеризуються навчальні словники англійської мови [2; 3]. Перші 2000 найчастотніших слів становлять реєстр словника першого рівня і потім відмічаються у словниках вищих рівнів. Тобто словники трьох рівнів будуються концентрично: у кожний словник наступного рівня включається реєстр словника попереднього рівня і додається відповідна кількість нових слів. Представлена стаття має на меті аналіз особливостей формування реєстру навчальних словників.

У лабораторії комп'ютерної лінгвістики Київського національного лінгвістичного університету укладено англо-українські навчальні словники першого [4] і другого [5] рівнів, які існують у паперовому і комп'ютерному форматі ([www.mova.info](http://www.mova.info)).

При формуванні реєстру словника першого рівня ми спиралися на аналіз лексичних мінімумів, підручників з англійської мови як іноземної, виданих у Великобританії, США та Німеччині, і на частотний словник, укладений з навчальною метою на навчальних текстах з різних дисциплін, що вивчаються в середній школі [6], обсяг вибірки – 5 млн словживань.

Використання лексичних мінімумів та підручників забезпечувало орієнтацію на програму початкового рівня вивчення мови, а частотний словник – відбір достатньою мірою частотної лексики: до реєстру включалися слова, частота яких становила 100 і більше вживань.

Якщо слово мало частоту не нижчу заданого порогу, але не відповідало програмі тематично, семантично чи мало стилістичне забарвлення, воно не включалося до реєстру, що і сталося, наприклад, зі словами **acid** кислота (частота 140), **adopt** усиновляти; засвоювати; привласнювати (частота 145), **aluminium** алюміній (частота 134), **altitude** висота над рівнем моря (частота 165), **ancestor** предок, праотець (частота 153), **banana** банан (частота 198), **buffalo** буйвол, американський бізон (частота 302) та ін.

Формування реєстру словника другого рівня передбачає доповнення реєстру словника першого рівня приблизно такою ж кількістю лексем. Для цього аналізувався реєстр навчального словника [5], відібраних з нього слова включалися в реєстр, якщо їх частота становила 50 і більше вживань у частотному словнику [6]. Отже, доповнення реєстру словника спиралося на поєднання тематичного та стилістичного підходів, що цілком виправдано, як показує використання цього словника.

Але доповнення реєстру для словника вищого рівня полягає не лише у збільшенні кількості словникових статей, а й у внесенні змін до будови словникової статті: збільшується кількість значень

© Гриднєва Л. М., Перебийніс В. І., 2009

ряду реєстрових слів і кількість наведених у статті словосполучень чи ілюстративних прикладів. Для ефективного розв'язання цієї задачі дуже корисними були б семантичні частотні словники, яких, на жаль, у нашому розпорядженні немає. Допомогали дані словників сполучуваності, особливо довідника найчастотніших англійських словосполучень [7], але він наводить дані для менш ніж 500 слів.

Зараз розпочато роботу по укладанню навчального англо-українського словника третього рівня, яка своїм першим етапом має доповнення реєстру словника другого рівня приблизно 5000 нових реєстрових слів.

Перше завдання, яке стоїть перед упорядниками навчального словника, – відібрати лексичний оптимум, обумовлений програмами: робити це треба так, щоб забезпечити включення до словника найбільш потрібних для кожної теми слів. Формування реєстру розбивається на дві задачі: 1) підбір достатньо великого списку слів і 2) перевірку їх на живаність і частотність.

Як і в роботі з уже виданими й упорядкованими словниками, добір нових лексем до реєстру відбувається з урахуванням досягнень лексикографів англійських країн, а саме Великобританії та Сполучених Штатів Америки. На першому етапі було використано три словники середнього та вищого рівнів: Longman. Wordwise Dictionary for pre-intermediate to intermediate Learners [3]; Macmillan. Essential Dictionary for learners of English [8] та Oxford Student's Dictionary of English [9]. Аналіз відібраних лексем показав, що лише 58% від загальної кількості наявні в усіх трьох словниках, а переважна більшість (98,5%) – у двох останніх словниках [8; 9]. Це дає підставу для того, щоб у подальшому працювати з одним із них, наприклад Оксфордським словником [9].

Обмеження, чи оптимізація, кількості реєстрових слів проводиться за допомогою Частотного словника англійської мови Дж. Керролла, Р. Девіса, Б. Річмана [8], який використовувався й раніше. Використання частотного словника дуже важливе, по-перше, тому, що він також був створений з навчальною метою; по-друге, ще й тому, що дозволяє відібрати ті лексичні одиниці, що відповідають заданому порогу частоти, яким вибрана частота у 10 слововживань, тобто до реєстру вносилися слова, частота яких була 10 і більше.

Щоправда, використаний частотний словник є словником словоформ, а не слів; у ньому констатується частота словоформи, а не лемми, до того ж без урахування частиномовної приналежності. Тому виникає проблема лематизації та ідентифікації омонімів (омоформ): як бути у випадку омонімії іменника і дієслова; іменника, дієслова і прикметника; дієслівної форми *past participle* і прикметника; дієслова та прикметника; іменника та прикметника і т.ін.?

У кожному окремому випадку рішення приймається залежно від наявності та частотності різних словоформ. При великій загальній частоті словоформ, як правило, усі омоніми вносяться до реєстру. Так, наприклад, іменник *buzz* *дзижчання; гудіння; бриніння* і дієслово *buzz* *дзижчати; гудіти; дзинчати* представлені у частотному словнику словоформами *buzz* (частота 67), *buzzed* (21), *buzzes* (6), *buzzing* (48); отже доцільно ввести до реєстру обидві лексеми: *buzz*<sub>1</sub> *n* і *buzz*<sub>2</sub> *v*, тому що однозначно дієслівних словоформ 69 (48,6% від загальної частоти), а тих, які можна трактувати і як іменник, і як дієслово, – 78. Навіть, якщо вважати, що 50% цієї частоти належить дієсловам, то частота іменних форм становить 36 і втричі перевищує поріг частоти для включення в реєстр.

Так само вирішується проблема з омонімами *burrow* *n* *нора* і *burrow* *v* *рити нору* (*burrow* – 49, *burrowed* – 6, *burrowing* – 19, *burrows* – 26). Обидва омоніми *breeze*-іменник, який раніше не був внесений у реєстр попереднього словника, оскільки не відповідав програмам тих рівнів викладання мови, та *breeze*-дієслово попадають у новий реєстр на підставі високих частот зафіксованих словоформ: *breeze* – 147, *breezes* – 32, *breezed* – 2; останні дві словоформи дієслівні; навряд чи назва вітру вживається у множині. Так само вирішується з омонімами *braid* *n* *коса*, *заплетене волосся* і *braid* *v* *заплітати косу; зав'язувати волосся*, які мають достатні частоти словоформ для ідентифікації обох лексем (*braid* – 20, *braided* – 17, *braiding* – 6, *braids* – 26).

Іноді один із омонімів уже був внесений у реєстр словника нижчого рівня, а другий додається на наступному етапі вивчення мови. Наприклад, іменник *butter* *масло* наявний у всіх згаданих вище словниках, а дієслово *butter* *намазати маслом* додається тільки в словник вищого рівня, тому що має для цього достатню загальну частоту дієслівних словоформ (*buttered* – 20 та *buttering* – 2) і відповідає програмі. Так само іменник *branch* *гілка; галузь; відгалуження; філіал, відділення* наявний у попередньому реєстрі; а дієслово *branch* зі значенням *розгалужуватися; галузитися, гілкуватися; розширюватися* дописується у словник вищого рівня: сумарна частота і частоти окремих словоформ достатньо високі для того, щоб і дієслівна лексема була представлена у реєстрі (*branch* – 243, *branched* – 8, *branching* – 16, *branches* – 345). Це ж стосується дієслова *bubble* *пузиритися, киніти; бити джерелом (струменем)* (*bubble* – 60, *bubbled* – 10, *bubbles* – 98, *bubbling* – 24), також іменника *burn* *опік* (*burn* – 194, *burns* – 118), який додається в реєстр до уже наявного там дієслова *burn* *палити, спалювати, обпалювати; горіти, згорати, палати*. Омоформи *broken* як *past participle* від *break*, наявний у попередніх словниках, і *broken* як прикметник зі значенням *розбитий, зруйнований*;

ламана, калічена (про мову); нестійкий, несталий мають дуже велику сумарну частоту – 538; і **broken** прикметник, цілком реальний у програмах вищого рівня, теж додається до списку нових слів.

Іноді до реєстру вноситься лише один із омонімів, а другий відкидається з різних причин: або через недостатню частоту словоформ, або через неможливість вирішити проблему без контексту. У випадку з парою слів **buckle** *n* *пряжка* і **buckle** *v* *гнути, згинати; згинатися* при частотах словоформ **buckle** – 24, **buckles** – 24, **buckling** – 3 неможливо вирішити щось стосовно дієслова; до реєстру вноситься тільки іменник. Слово **bust** може бути іменником зі значенням *бюст; жіночі груди* і дієсловом зі значенням *збанкрутувати*, проте, судячи зі словоформ, зафіксованих у частотному словнику (**bust** – 20, **busted** – 3), **bust** швидше за все функціонує у першому значенні; частота дієслівної словоформи незначна.

Звичайно, певною мірою рішення щодо відбору нових реєстрових слів носять суб'єктивний характер. Це зумовлюється кількома причинами. Перша з них – недосконалість частотних словників, у яких (зокрема і у Дж. Керролла [8]) частота дається для словоформ, а частоти слів як суми частот його словоформ відсутні, при тому що без контексту чи граматичної інформації у словниковій статті неможливо розрізнити, наприклад, множину іменника і третю особу дієслова в теперішньому часі (Present Indefinite), інгову форму дієслова і прикметник, іменник і прикметник, прикметник і прислівник та ін.

Отже, паралельно з упорядкуванням навчальних перекладних словників і словників сполучуваності необхідно створювати інші частотні словники: семантичні словники з акцентом на частотності різних значень полісемічних слів у текстах різної тематичної та стильової приналежності, частотні словники словосполучень, словники омонімів з фіксацією частотності усіх можливих омоформ.

#### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ТА ЛІТЕРАТУРИ

1. Перебийніс В. І. Лексикографічне забезпечення навчального процесу з іноземної мови // Вісник Київського лінгвістичного університету. – Т. 2. – № 1. – 1999. – С. 12-19.
2. Hornby A. S. Oxford Advanced Learners Dictionary of Current English. – Oxford University Press, 1989.– 1580 p.
3. Longman. Wordwise Dictionary for pre-intermediate to intermediate Learners. – Barcelona: Pearson Education Limited, 2001. – 722 p.
4. Англо-український словник-мінімум. – К.: Наукова думка, 1995. – 270 с.
5. Англо-український навчальний словник з методичними коментарями та граматичними таблицями / В. І. Перебийніс, Е. П. Рукіна, С. С. Хидекель; За ред. дійсного члена АПН України Г. І. Артемчука. – К.: Вежа, 2002. – VIII, 424 с.
6. John B. Carroll, Peter Davies, Barry Richman. Word Frequency Book. – New York: American Heritage Publishing Co., Inc., 1971. – 856 p.
7. Справочник наиболее употребительных английских словосочетаний: Пособие для учителя / В. И. Перебийнос, С. С. Хидекель, Л. Р. Вайнер и др. – М.: Просвещение, 1986. – 320 с.
8. Macmillan. Essential Dictionary for learners of English. + CD=ROM, 2003.
9. Oxford Student's Dictionary of English. – Oxford University Press, 2008. – 775 p.

Матеріал надійшов до редакції 05.03. 2009 р.

#### **Гриднева Л. М., Перебийнос В. И. Формирование реестра англо-украинского учебного словаря.**

*Статья посвящена изучению формирования реестра англо-украинского учебного словаря. Отмечается, что этот реестр должен соответствовать программам по иностранному языку на определенном этапе его преподавания, включать ограниченное количество часто употребляемых слов с высокой словообразовательной потенцией и сочетаемостью. Сделан вывод, что при его формировании следует основываться как на лексических минимумах, так и на частотных словарях.*

#### **Grydnieva L. M., Perebeynos V.I. Forming of the Glossary of the English-Ukrainian Learners Dictionary.**

*The article is dedicated to the study of English-Ukrainian learners' glossary formation. It is pointed out that this glossary must correspond to the program of teaching a language at a certain level of its study, includes a restricted number of highly frequent words having a high word building and combinatorial potential. The conclusion is made that while forming it both lexical minimums and frequency dictionaries should be used.*