УДК 811.161.2=222.1'373'374.7-13:39

МЕТОДИКА ВИЯВЛЕННЯ КЛЮЧОВИХ СЛІВ НА ОСНОВІ ЗІСТАВЛЕННЯ ЛЕКСИКИ ЧАСТОТНИХ СЛОВНИКІВ

Максимів Ольга Йосифівна,
викл.
Львівський національний університет імені Івана Франка

У статті запропоновано методику виявлення ключових слів мови на основі зіставлення даних частотних словників. Розроблено методику формування кореневих груп слів за даними частотних словників, зіставлення частотних списків цих груп і виявлення лексики, етноспецифічної для кожної із зіставлюваних мов та специфічної для галузі, на базі якої укладалися частотні словники.

Ключові слова: *частотні словники, етноспецифіка, ключові слова, кореневі групи слів, зіставлення частотних словників.*

У сучасній лінгвістиці активно і з різних позицій вивчаються ключові слова художніх [Семенюк 2002; Форманова 1999] та технічних [Алексеев 1974; Пиотровский 1976] текстів. Вихідним положенням є те, що кожна природна мова є кодовою системою передавання інформації, яка може породжувати тексти. Конкретний вигляд цих текстів визначається, по-перше, можливостями самої системи, по-друге, обмеженнями, які накладає на цю систему сучасна мовна норма і, по-третє, ситуацією, яку описує текст. Ключовими або домінантними, опорними словами вважаються інформаційно навантажені слова, які позначають поняття, безпосередньо пов'язані з ситуацією, котра описується в конкретному тексті. У технічних текстах це терміни, у художніх – "слова-символи. Позначаючи найактуальніші явища, реалії, вони … відіграють роль певних "опорних" точок у побудові та структуруванні концептуальної та мовної картин світу" [Семенюк 2002, 129]. Відповідно, матеріалом для виявлення цих слів слугують технічні та художні тексти.

Відзначається також, що функціональні стилі мови та ідіостилі письменників різняться в першу чергу мовними особливостями, зокрема лексичними [Баланюк 2002; Головин 1970; Перебийніс 1985; Шашкевич 1966 та ін.]. І наголошується при цьому не лише на відмінностях у вживанні якихось специфічно маркованих слів, а на тому, що "велика увага повинна приділятися саме стилістично нейтральному шару, який і складає лексичне багатство твору" [Перебийніс 1985, 160]. У межах однієї мови порівнюються твори різних авторів чи тексти різних функціональних стилів, застосовується складні статистичні підрахунки. Результатом є визначення ступеня близькості аналізованих текстів, ступеня складності та різноманітності лексики чи ступеня належності того чи іншого твору до даного стилю (у кількісному вираженні).

Для того, щоб перевірити і конкретизувати припущення про те, що "при всій істотній подібності етноузуальна структура в західних і східних мовах досить відмінна" [Тищенко 2007], спробуємо зіставити лексику частотних словників публіцистичного стилю західної (української) та східної (перської) мов. Будь-який стиль мовлення має універсальні й етноспецифічні риси. Виявивши ключові слова обох мов та згрупувавши їх за певними ознаками, можна буде визначити етноспецифіку кожної мови, яка полягатиме у комбінації і частотності загальновживаних лексем кожної мови.

Нам відоме лише одне дослідження [Мухин 2008], яке використовує зіставлення лексики частотних словників для виявлення особливостей індивідуального авторського слововживання. Знову ж таки, зіставляються художні тексти у межах однієї мови.

Метою даної статті є продемонструвати методику виявлення ключових слів різних мов на основі зіставлення лексики частотних словників одного функціонального стилю.

Джерелом дослідження слугують дані частотних словників публіцистичного стилю мовлення української [Дарчук 2009] (далі – ЧСУ) та перської [Максимів 2008] (далі – ЧСП) мов. Обидва словники укладені за матеріалами лексики публіцистичних текстів, тобто дотримано стилістичної однорідності [Андрющенко 1968], із центральних та обласних українських та іранських газет, тобто дотримано просторової єдності, за 2003–2004 (відповідно 1381–1382) роки, тобто дотримано хронологічної однорідності. Таким чином, дотримано один з основних для контрастивного дослідження принципів – принцип системності [Дубичинский 1998, 66; Кочерган 2006, 85–92], який передбачає перш за все ієрархічну впорядкованість вибору параметрів зіставлення: відповідно до частотної ієрархії рангів зіставлено лексеми перської та української мов.

Об'єктом дослідження стали спільні та етноспецифічні ключові слова частотних списків перської та української мов. Специфічними будуть як ситуації і факти дійсності, до яких частіше звертаються у пресі одного періоду різних країн, так і слова, які обирають для опису стандартних ситуацій чи вираження сприймань, думок, почуттів. Так звані безеквівалентні слова, які переважно виступають об'єктом етнолінгвістичних досліджень і відображають особливості національної культури, побуту народів, перебувають на периферіях лексичних систем [Манакин 2002, 268] і не мають високої частотності, тому вони не стали об'єктом нашого дослідження.

Одним із можливих способів виявлення інформаційно навантажених ключових слів тексту (переважно термінів) є обчислення емпіричних розподілів [Алексеев 1974, 212–213; Пиотровский 1976, 3–13]. З використанням частотних словників продуктивним, на нашу думку, є застосування закономірності Бредфорда: виокремити зону ядра, центральну зону та зону усічення [Хайтун 1983, 71–85; Чурсин 1982, 91–100]. До зони ядра, очевидно, у першу чергу, потраплять службові та малоінформативні загальні повнозначні слова. До центральної зони – ключові слова, а до зони усічення – лексика, яка визначає багатство словника кожної мови, кожного автора зокрема. Таку лексику не можна вважати етноспецифічною. Сюди також належить, з іншого боку, застаріла чи неусталена лексика [Тисенко 1977, 927–928].

Виділення частотних зон можливе за різними ознаками [Тулдава 1987, 65]. Його можна здійснювати залежно від цілей і завдань конкретного дослідження. Найдоречнішим є таке розмежування частотних зон, при якому кожна зона охоплювала б діапазон, який дорівнює одному порядку [Малаховский 1980, 101]. Пропонуємо лексику ЧСУ та ЧСП поділити порціями по 100 слів на основі рангів. Далі послідовно зіставити між собою дані групи слів. Будь-яке зіставлення мов у результаті приводить до встановлення трьох основних властивостей [Манакин 2002]: всезагальних, спільних та відмінних. Результати зіставлення ділимо на три групи:

1) слова, спільні для обох мов (сюди увійдуть і усезагальні, і спільні слова);

2) слова, частотніші для української мови;

3) слова, частотніші для перської мови.

Типологічно різні мови характеризуються різною структурою слів, і тому визначальні для виділення слів критерії в одних мовах стають незастосовними щодо ін-

ших мов [Вихованець 1988, 14–19]. Враховуючи типологічну різницю перської та української мов (напр., велику кількість складних дієслів перської мови чи різновидових дієслів української мови), частотні словники лексем не можуть бути предметом зіставлення. Виникає потреба шукати для цього інший рівень.

Для міжмовного зіставлення продуктивним вважається утворення лексичних групувань слів за різними принципами (див., напр. [Кочерган 2006, 296–297; Липатов 1981, 51–57, Тулдава 38–39]). Утворені групування розглядаються як побудови, які дозволяють моделювати структуру словника з певної (у нашому випадку це частотність) точки зору. Залежно від типу взаємозв'язку слів у системі лексики виділяються особливі субсистеми під загальною назвою "лексичні групи" (див. [Тулдава 1987, 39]). Вони мають свої різновиди на різних рівнях опису лексики. На підставі подібності фонетичного складу чи словотворення виділяють "лексико-формаційні групи" слів. Спільність граматичних значень є підставою для утворення "лексико-граматичних груп". На основі близькості чи однорідності семантики слів утворюються "лексико-семантичні групи". Тематичні групи формуються за сферами вживання слів, практично безвідносно до того, в яких відношеннях один до одного перебувають слова за їхнім значенням. Крім цих основних типів лексичних груп, можливі також групування слів на різних інших підставах (етимологічних, стилістичних та ін.). Можливе групування і за змішаними ознаками.

Для виявлення специфічності лексичної системи було прийнято рішення звести вихідні частотні словники лексем до частотних словників словотвірних гнізд, кореневих морфем, і вже після цього проводити зіставлення. Адже саме корінь і афікс, на противагу слову, виступають мінімальними (нечленованими далі) значущими одиницями, мінімальними двобічними одиницями, у яких за певним експонентом закріплений той або інший елемент змісту (див. [Вихованець 1988, 14–19]). І саме корінь несе лексичне значення. Перенісши увагу з ізольованої лексеми на групу однокореневих вдалося мінімізувати вплив типу мови на результати зіставлення, вплив граматичних розбіжностей зіставлюваних мов, і зосередитися на кореляції частотності вживання певних понять, узагальнених ідей, закріплених за тим чи іншим словом-вершиною групи однокореневих.

Під словотвірним гніздом маємо на увазі, слідом за Ю. Тулдавою [1987, 122] та І.А. Кіссеном [1964, 44–58] групу однокореневих слів, об'єднаних словотвірними відношеннями. Додавання абсолютних частот однокореневих слів суттєво змінило картину рангового розподілу слів. Наприклад, у частотному списку лексем '*виконання*' було у 7-й сотні, коренева група з однойменною вершиною виявилася у 2-й сотні, лексема زمین *zamin* "*земля*" була у 3-й сотні, група з такою вершиною виявилася у 1-й.

Під час формування частотних списків кореневих груп слів зіставлюваних мов дотримуємося такої ієрархії принципів:

1. Принцип частотності. Оскільки ми відштовхувалися від частотних словників двох мов, то вершиною групи стає найчастотніше у ній слово. Але, напр., багато дієслів перської мови складаються мінімум з двох компонентів і вважаються двочленними фразеологічними одиницами, компоненти яких зберігають словесну автономність, а не складними словами (див. [Рубинчик 1991, 114–116]): خبر دادن xabar dādan "повідомляти". Кожен компонент таких дієслів, відповідно, у частотному

словнику реєструється як окреме слово. Українською мовою складні дієслова переважно перекладаються одним словом. Для наочності і однотипності подання результатів дослідження у тексті оперуватимемо не словом-вершиною групи, а найчастотнішим простим іменником (часто віддієслівним) цієї групи. Наприклад, для групи *голос* вершиною (найчастотнішим) є віддієслівний іменник *голосування*, для групи جهانjahān "світ" – не іменник, а прикметник جهانیjahāni "світовий".

2. Морфемний принцип. Починаючи з вершини частотного списку для кожного слова знаходимо усі слова, які мають спільну з ним кореневу морфему. Так, до однієї групи, напр., в українському списку потрапляють слова *йти (іти), піти, вийти, виходити, прийти, проходити, приходити*, у перському رفتن raftan "іти", راهraveš "спосіб", رفتارraftār "поведінка, вчинок", پیشرفت pišraft "рух уперед, прогрес" та ін. Відповідно до цього принципу, найвище у частотному списку кореневих груп розташовуватимуться прості слова, далі – похідні і складні, хоча у частотному словнику лексем похідне слово може мати вищий ранг, напр. у групі *громада* найчастотнішим було слово *громадянин*, у групі نمودن namudan "виставка, показ" – نماینده namāyande "представник, депутат".

3. Перекладний принцип. Відомо, що семантичний об'єм слів, схожих у відношенні до поняття, яке вони виражають, у різних мовах найчастіше не збігається (див., напр., [Ярцева 1960, 3–14]). Оскільки кінцевою метою нашого дослідження є виявити специфічні, ключові слова, а не детально описати семантику цих слів, то було прийнято рішення під час зіставлення спиратися не на значення слів, а на їх взаємні переклади обома мовами. Менш частотне спільнокореневе слово вилучаємо зі списку кореневих груп слів у разі, якщо його можна перекласти іншою мовою з використанням слова, спільнокореневого до вихідного. Напр., у перському списку, включивши до групи (اصول) اصل asl (osul) "основа, суть" слова (اصلاً) اصلا aslan "зовсім, насправді" і اصلي asli "основний, суттєвий", перше ми залишили у списку кореневих груп слів і воно сформувало свою групу, а друге зі списку вилучили; в українському списку зі слів *видання* та *передавати*, які увійшли до групи *дати* з перськими еквівалентами دادن dādan, زدن zadan і گذاشتن gozāštan, перше, яке перською перекладається نشر našr, چاپ čāp у списку залишили, а друге, яке має переклад دادن dādan, رساندن rasāndan, вилучили. Предметна співвіднесеність слова дає можливість застосувати перекладний принцип, не порушуючи у загальних рисах цілісності семантичного ядра кожної групи в кожній мові. Таким чином, основою зіставлення стає "третій член" (про це див. дет. [Кочерган 2006, 79–85]) – узагальнене лексично-поняттєве значення кореневої морфеми кожної групи слів.

Формуючи списки кореневих груп слів ми дотримувалися таких правил:

1. Оскільки під час укладання ЧСП омонімія не усувалася, ЧСУ довелося відкоригувати, об'єднавши розмежовані там омоніми, напр., *жаль* як іменник чоловічого роду та прислівник, *засуджений* як іменник чоловічого роду та прикметник, *звичайно* як прислівник та частку. Об'єднали також слова, які різнилися лише великою/малою першою літерою, напр., *Автозаз* і *автозаз*.

2. Слова, які у своєму складі містили нелітерні символи (цифри, розділові знаки, слова, написані не перською (чи відповідно кириличною) графікою тощо), але складалися ще й з відповідних літер також, додаємо до відповідної кореневої групи і вилучаємо із загального списку. Напр., перське *EDI*نرم чи українське *PR-акція*.

3. Абревіатури та скорочення, більше властиві українській мові, аніж перській, по можливості розшифровували і додавали до відповідних груп слів, напр., слово *Укрзалізниця* увійшло до груп *Україна*, *залізниця* і *залізний*, слово *ЄС* – до груп *Європа* і *союз*.

4. Усі числівники розглядалися окремо, і складені до своїх компонентів не зводилися.

5. Якщо у тлумачному чи перекладному словнику для певного слова подається тлумачення (чи переклад) з використанням відсилання до іншого слова, або воно тлумачиться з використанням спільнокореневого, то це слово зводимо до того, до якого є відсилання і зі списку кореневих груп вилучаємо менш частотне з них. Напр., خفتن xoftan → خوابیدن xābidan "спати", تماشاچي tamāšāči → تماشاگر tamāšāgar "глядач, очевидець"; 'удатися' → 'вдатися', 'зіткнутися' → 'стикатися' та ін.

6. Щоб уникнути повторення груп слів зі схожими вершинами, вирішувалося питання про кожне слово, вже включене у певну кореневу групу. Тут було вироблено такі правила:

а) якщо слово складається з однієї кореневої морфеми і афікса, який не має власного лексичного значення, то воно зводиться до вихідного слова і зі списку вилучається зовсім. Напр., в українському словнику, під час формування групи *український*, до неї було включено слова *українець*, *українізація*, *по-українськи (українському)*, *українство*, *українськість* і з частотного списку їх вилучили. У перському словнику до групи مردم mardom "люди, народ" включили слова مردمي mardomi "людський, народний", نامردمي nāmardomi "нелюдяність, жорстокість" і також вилучили їх з частотного списку;

б) якщо слово складається з декількох морфем, кожна з яких має власне лексичне значення, то воно додається до вихідного слова, але зі списку не вилучається і може стати вершиною іншої кореневої групи. Напр., в українському словнику, під час формування тієї ж групи *український*, до неї було включено слова *україномовний*, *західноукраїнський* але з частотного списку їх не вилучали, вони увійшли також до груп *мова* та *західний* відповідно. У перському словнику до групи اقتصاد eqtesād "економіка" увійшли слова اقتصاددان eqtesāddān "економіст", قتصاددانی eqtesāddāni "професія економіста" та سراقتصاددان sareqtesāddān "головний економіст" які, у свою чергу, сформували групу з вершиною اقتصاددان eqtesāddān "економіст" і також стали членами групи دانستن dānestan "знання";

в) відповідно до перекладного принципу, зі списку кореневих груп слів не вилучалося слово, утворене за допомогою одного кореня і афікса, переклад якого іншою мовою не можливий з використанням кореня-вершини групи. Напр., перське آینده āyande "майбутнє", похідне від дієслова آمدن āmadan "приходити", було включено до групи آمدن āmadan "прихід, прибуття", але зі списку не вилучалося і сформувало свою групу آینده āyande "майбутнє"; українське *голосування* як похідне увійшло до групи *голос* з перськими відповідниками صدا sedā, آواز āvāz і сформувало свою групу з перським перекладом رای گیری rāygiri. Таким чином одне слово (переважно складне) одночасно може бути вершиною кореневої групи і входити до декількох інших груп як похідне від їхніх вершин.

7. Для перської мови, окрім формально фонетично однакових, спільнокореневими також вважалися:

а) слова, утворені від різних основ (теперішнього та минулого часу) того самого дієслова: напр., до групи داشتن dāštan "мати, володіти" увійшли як слова بازداشتگاه bāzdāštgāh "місце попереднього ув'язнення", بزرگداشت bozorgdāšt "повага", утворені від основи минулого часу дієслова, так і جاندار jāndār "істота, живий", دارا dārā "той, хто має", утворені від основи теперішнього часу.

б) слова-арабізми, утворені за допомогою внутрішніх флексій від одного арабського кореня, якщо тлумачні словники перської мови не фіксують різкого зсуву значення чи переклади яких українською мовою містять той самий корінь. Напр., до групи (حدود) حد hadd (hodud) "межа, край" було включено слова محدود mahdud "обмежений, межуючий" та (حدوداً) حدودا hodudan "приблизно, у межах", до групи مسکن maskan "житло" слово مسکوني maskuni "житловий".

Оскільки під час формування лексико-семантичних груп, як правило, враховується синонімія і слова-синоніми включаються до груп, ми також зробили спробу включити до кореневих груп слів синоніми, зафіксовані у частотному словнику. Проте згодом відмовилися від цього, бо група розросталася до надто великих розмірів, залученим виявлявся цілий ряд інших високочастотних груп, втрачалася формальна єдність і можливість зіставляти дані словників пари мов, базуючись на частотності. Абсолютних синонімів дуже мало, різні слова у різних контекстах мають різні синоніми, що фіксують відповідні словники, а переклади цих синонімів іншою мовою не обов'язково є синонімами між собою, групи виявляються різноспрямованими. Це інший план мови, опис не суто мовної, а мовно-ментальної системи. Необхідний ряд додаткових досліджень, щоб виробити критерії включання синонімів до однієї групи, визначити межі таких груп.

Процедура зіставлення мала такий вигляд:

1. Списки кореневих груп перської та української мов було розділено на порції по 100 слів у кожній.

2. Зафіксовано усі словникові варіанти перекладу кожного слова-вершини групи, а також слів, які увійшли до групи і були вилучені зі списку кореневих груп слів. Враховано і омонімію, і полісемію.

3. Пошук був спрямований згори кожного списку на предмет виявлення еквівалента до кожного слова-вершини групи у списку перекладів іншої мови. Групи вважалися еквівалентними, якщо у сукупності можливих перекладів був хоча б один із зафіксованих на попередньому етапі. Усі подальші висновки базуються на парах еквівалентних слів.

4. Якщо еквівалент знаходили у межах даної сотні, то зараховували групу до спільних для обох мов. Якщо ж еквівалент виявлявся у іншій сотні, то група одержувала статус специфічної для тієї мови, у якій мала вищий ранг. Зауважимо, що специфічним слово-вершина групи вважалося лише за умови, що різниця рангів у списках була більшою, ніж 100. Формально близькі слова, які опинилися на межі, у різних сотнях, фіксувалися окремо.

Ми розуміємо, що інформацію перекладних словників слід використовувати критично, бо у них відповідники наводяться з певною мірою приблизності і до певного слова вихідної мови подаються всі можливі варіанти другої мови [Кочерган 2006, 300]. Звернення до перекладних словників дозволяє апроксиматизувати суку-

пність значень слів-членів кореневих груп. Наголошуємо, що кінцевою метою зіставлення є виявлення ключових слів певних мов (перської та української) та субмови (публіцистики), а не суто семантичних особливостей лексики цих мов.

Під час зіставлення статистичних характеристик словників словоформ обох мов [Максимів 2009] виявилося, що обидва словники мають приблизно однаковий обсяг – близько 60 тисяч словоформ. Роблячи попередні висновки, припускаємо, що окрема субмова (у даному випадку публіцистичний стиль мовлення) оперує певною відносно обмеженою кількістю слів – загальних для всієї мови і ключових для даної субмови, тобто виходимо на спеціальний лексикон галузі (публіцистики).

Таким чином, ми одержали три списки:

1) слова, частотніші для перської мови, їх можна вважати ключовими для перської лексичної системи, напр., اسلامیeslāmi "ісламський", آمریکا āmrikā "Америка", نفت naft "нафта", ایران irān "Іран", جهان jahān "світ";

2) слова, частотніші для української мови, ключові для української лексичної системи, напр., *молодий, влада, український, Росія, Янукович*;

3) слова, частотність яких є співвідносною в обох мовах (т. зв. "спільні"), серед них – ключові для лексики публіцистичного стилю мовлення, напр., رئیس ra'is "президент", دولت dowlat "держава", اطلاع ettelā' "повідомлення", سیاست siyāsat "політика", قانون qānun "закон".

Визначальним є те, що наведені приклади – це результат зіставлення перших двох сотень кореневих груп слів. Тобто вже навіть серед найчастотнішої лексики, переважно службових та загальних слів, виявилися етноспецифічні елементи.

Для подальших досліджень варто згрупувати виявлені списки специфічних лексем за певними (очевидно, семантичними) ознаками з метою конкретизації і формалізації відмінностей етноузуальної структури східної (перської) та західної (української) мов. Напр., виокремити відмінності, зумовлені структурою мови, географічним розташуванням країни, економікою, політикою, культурою тощо. Необхідно також розглянути словотвірний потенціал етно- та галузево-специфічних груп слів. Зокрема, цікаво визначити структуру гнізд з еквівалентними вершинами (наприклад, کشور kešvar "країна",حق haqq "право", گفتن goftan "слово").

В статье предлагается методика определения ключевых слов языка на основе сопоставления данных частотных словарей. Разработана методика формирования коренных групп слов по данным частотных словарей. Сопоставляются частотные списки этих групп и определяется лексика, которая является этноспецифической для каждого из сопоставляемых языков и специфической для отрасли, на базе которой составлялись частотные словари.

Ключевые слова: *частотные словари, этноспецифика, ключевые слова, коренные группы слов, сопоставление частотных словарей.*

The methodics of finding out the new words of the language on the base of comparison of the frequency dictionaries' data is offered. The methodics of forming the root groups according to the data of the frequency dictionaries, comparison of frequency lists of these groups and finding out the vocabulary, ethnospecified to each of the compared languages and specific for the area on the base of which the frequency dictionaries were composed.

Key words: *frequency dictionaries, ethnospecifity, key words, root groups of words, comparison of frequency lists.*

Література:

1. *Алексеев П.М.* Основы статистической оптимизации преподавания иностранных языков / Алексеев П.М., Герман-Прозорова Л.П., Пиотровский Р.Г., Щепетова О.П. // Статистика речи и автоматический анализ текста. – Л.: Наука, 1974. – С. 195–234.
2. *Андрющенко В.М.* Новые работы в области статистической лексикографии / В.М. Андрющенко // Вопросы языкознания. – 1968. – № 5. – С. 137.
3. *Баланюк С.С.* Контрастивно-зіставний аналіз текстів різних функціональних стилів / С.С. Баланюк // Культура народов Причерноморья. – № 32. – 2002. – С. 253–256.
4. *Вихованець І.Р.* Частини мови в семантико-гаратичному аспекті / І.Р. Вихованець. – К., 1988. – С. 14–19.
5. *Головин Б.Н.* Язык и статистика / Борис Николаевич Головин. – М.: Просвещение, 1970. – 190 с.
6. *Дарчук Н.П.* (ред.) Частотний словник українського публіцистичного стилю // Лінгвістичний портал MOVA.info. – 2003-2007. – [Цит. 24 лютого 2009]: Режим доступу http://www.mova.info/Page2.aspx?11=178.
7. *Дубичинский В.В.* Теоретическая и практическая лексикография / В.В. Дубичинский. – Вена; Харьков, 1998. – 160 с.
8. *Киссен И.А.* Опыт статистического исследования частотности лексики передовых статей газеты "Кизил Узбекистон" // Научные труды Ташкентского государственного университета им. В.И. Ленина. Филологические науки. – Ташкент, 1964. – Вып. 247. Иранская, тюркская, арабская филология. Лингвостатистика. – С. 44–58.
9. *Кочерган М.П.* Основи зіставного мовознавства / Михайло Петрович Кочерган. – К.: Вид. центр "Академія", 2006. – 424 с.
10. *Липатов А.Т.* Лексико-семантические группы слов и моносемные поля синонимов / А.Т. Липатов // Филологические науки. – 1981. – № 2. – С. 51–57.
11. *Максимів О.Й.* Принципи укладання частотного словника газетної лексики сучасної перської мови / О.Й. Максимів // Вісник КНУ ім. Т.Шевченка, 2008 (у друці).
12. *Максимів О.Й.* Специфіка частотних словників публіцистичного стилю сучасних перської та української мов // Вісник Львівського університету. Серія філологічна, 2009 (у друці).
13. *Малаховский Л.В.* Принципы частотной стратификации словарного состава языка / Л.В. Малаховский // Статистика речи и автоматический анализ текста. – Л.: Наука, 1980. – С. 99–105.
14. *Манакин В.Н.* Языковые картины мира в перспективах контрастивной лексикологи / В.Н. Манакин // Культура народов Причерноморья. – 2002. – № 32. – С. 265–268.
15. *Мухин М.Ю.* Частотный спектр как вид представления концептуальной системы автора / М.Ю. Мухин // MegaLing'2008 Горизонти прикладної лінгвістики та лінгвістичних технологій. – Сімферополь: ДИАЙПИ, 2008. – С. 73-74.
16. *Перебийніс В.С.* Частотні словники та їх використання / Перебийніс В.С., Муравицька М.П., Дарчук Н.П. – К.: Наукова думка, 1985. – 202 с.
17. *Пиотровский Р.Г.* От редактора / Р.Г. Пиотровский // Частотный немецко-русский словарь-минимум газетной лексики. [Сост. А.С.Ротарь, В.А.Чижаковский]. – М.: Воениздат, 1976. – С. 3–13.
18. *Рубинчик Ю.А.* Лексикография персидского язика / Ю.А. Рубинчик. – М.: Наука, 1991. – 224 с.
19. *Семенюк О.А.* Ключевые слова в картине мира и языке епохи / О.А. Семенюк // Культура народов Причерноморья. – № 32. – 2002. – С. 129–132.
20. *Тисенко Э.В.* Статистические параметры словаря // Частотный словарь русского языка [Под ред. Л.Н.Засориной]. – М.: Русский язык, 1977. – С. 922–933.
21. *Тищенко К.* Східні мови і метатеорія мовознавства / Костянтин Тищенко // http://www.ruthenia.info/txt/tyschenkok/me.html (скопійовано 17.10.2007).
22. *Тулдава Ю.* Проблемы и методы квантитативно-системного исследования лексики / Ю. Тулдава. – Таллинн: Валгус, 1987. – 204 с.
23. *Форманова С.В.* Ключові слова у мовній картині світу Михайла Коцюбинського: автореф. дис. на здоб. наук. ступ. канд. філол. наук: 10.02.01 "Українська мова" / С.В. Форманова. – К., 1999. – 17 с.
24. *Хайтун С.Д.* Наукометрия. Состояние и перспективы / С.Д. Хайтун. – М.: Наука, 1983. – 344 с.
25. *Чурсин Н.Н.* Популярная информатика / Н.Н. Чурсин. – К.: Техніка, 1982. - 158 с.
26. *Шайкевич А.Я.* Опыт статистического выделения стилей // Межвузовская конференция по вопросам частотных словарей и автоматизации лингвостатистических работ. Тезисы докладов и сообщений. – Л.: Изд-во Ленингр. ун-та, 1966 – С. 78.
27. *Ярцева В.Н.* О сопоставительном методе изучения языков / В.Н. Ярцева // Филологические науки. – 1960. – № 1. – С. 3–14.