

14. Хроменков П. Н. Анализ и оценка эффективности современных систем машинного перевода: дис... канд. филол. наук : 10.12.21 / Павел Николаевич Хроменков. – М., 2000. – 265 с.
15. Шульжук К. Ф. Синтаксис української мови / Каленик Федорович Шульжук. – К. : Видавничий центр «Академія», 2004. – 408 с.
16. Яковина В. Огляд та аналіз метрик оцінювання якості перекладу / В. Яковина, В. Масюкевич // Вісник Національного університету «Львівська політехніка». Комп'ютерні науки та інформаційні технології. – 2013. – № 771. – С. 101–107.

## ВІДОМОСТІ ПРО АВТОРА

**Юлія Стахмич** – кандидат філологічних наук, доцент кафедри філології та перекладу Івано-Франківського національного технічного університету нафти і газу.

*Наукові інтереси:* комп'ютерна лінгвістика, машинний переклад, штучний інтелект, проблеми перекладацької еквівалентності та адекватності.

УДК 81'33

## ЦИФРОВА ЛЕКСИКОГРАФІЧНА СИСТЕМА «С. ІВАНИЦЬКИЙ, Ф. ШУМЛЯНСЬКИЙ. РОСІЙСЬКО-УКРАЇНСЬКИЙ СЛОВНИК» ТА ЇЇ МОЖЛИВОСТІ

**Людмила ТОМІЛЕНКО, Олександр РАБУЛЕЦЬ (Київ, Україна)**

*У статті розглянуто процес створення цифрової версії «Російсько-українського словника» (1918 р.) Сергія Іваницького та Федора Шумлянського. Стисло описано структуру досліджуваної праці як основу для побудови інформаційної системи «С. Іваницький, Ф. Шумлянський. Російсько-український словник». Продемонстровано можливості згаданої системи.*

**Ключові слова:** лексикографія, перекладний словник, російсько-український словник, структура, словникова стаття.

*В статье рассматривается процесс создания цифровой версии «Русско-украинского словаря» (1918 г.) Сергея Иваницкого и Федора Шумлянського. Кратко описывается структура исследуемой работы как основы построения информационной системы «С. Иваницкий, Ф. Шумлянський. Русско-украинский словарь». Продемонстрированы возможности этой системы.*

**Ключевые слова:** лексикография, переводной словарь, русско-украинский словарь, структура, словарная статья.

*The article discusses the process of creating a digital version of “Russian-Ukrainian Dictionary” (1918) S. Ivanytskyi and F. Shymlianskyi. Briefly described the structure of the study work as the basis for constructing an information system of “S. Ivanytskyi, F. Shumlianskyi. Russian-Ukrainian Dictionary”. Demonstrated capabilities of said system.*

**Key words:** lexicography, translation dictionary, Russian-Ukrainian dictionary, structure, entry.

Помітну роль у сучасному мовознавстві під час вивчення лексико-семантичних, граматичних та інших особливостей української мови відіграють перекладні лексикографічні довідники післяреволюційних років. Найпопулярнішим і найдослідженішим сьогодні є, безумовно, академічний «Російсько-український словник» за ред. А. Кримського та С. Єфремова (праці Є. Карпіловської, О. Тищенко, О. Скопненка, Ю. Поздрань та ін.). Однак комплексного аналізу лексики цієї праці, як і інших словників постреволюційної доби та періоду українізації, ще немає.

«Російсько-український словник» (1918 р.) у 2-х томах С. Іваницького та Ф. Шумлянського – досить відомий свого часу перекладний словник, що після тривалого забуття знову побачив світ 2006 р. Як зазначено в передмові, перевидання цієї лексикографічної праці в незалежній Україні відповідає потребі повернення всього цінного з репресованої мовознавчої спадщини першої половини ХХ ст. Уведення словника в практично-ужитковий і науковий обіг сприятиме поверненню українській мові повноти лексичного складу, а українській лексикографії – вилученій ланки, необхідної для вивчення словникарства відповідної доби [1, с.5].

Для вітчизняної лінгвістики особливу цінність має українська частина згаданої праці як важливе джерело вивчення лексичного складу мови. Оскільки словник російсько-український, то пріоритетним завданням комплексного аналізу української лексики, є «витягнення» її з реєстру. Зрозуміло, що без застосування програмних методів зробити це було б украй важко. Ще однією суттєвою проблемою стало те, що у вільному доступі наявний тільки сканований варіант словника, а не його версія в текстовому форматі. Тому скановану копію було переведено в текстову за допомогою програмних засобів оптичного

розпізнавання символів (англ. Optical character recognition, OCR) і вручну виправлено помилки, що з’явилися в процесі розпізнавання тексту.

Щоб отримати базу українських слів досліджуваного джерела, необхідно описати структуру всіх різноманітних словникових статей. Подібний опис уже зроблено на матеріалі «Словника московсько-українського» В. Дубровського [3], структура якого є дещо простішою порівняно зі структурою аналізованої праці. Хоча лексикографія й виробила досить чітку схему будови словникової статті у тлумачному та перекладному словнику, систему розділових знаків, шрифтів, умовних позначок, які ділять словникову статтю на рубрики чи зони [2, с. 26–27], різні лексикони (особливо минулих років) усе-таки відрізняються своєю структурою. Отже, розглянемо основні типи словникових статей «Російсько-українського словника» С. Іваницького та Ф. Шумлянського.

1. Найпростіша словникова стаття складається з російського реєстрового слова (жирним шрифтом), після тире подано український відповідник (або відповідники через кому) світлим шрифтом. В українській частині реєстру зафіксовано також деякі слова з різними варіантами написання, напр.: **Авантюристь** – пройдис(ь)віт.



Рис. 1. Модель найпростішої словникової статті

2. Словникові статті з ремарками оформлено так: російський відповідник, кома, ремарка, латинська назва, тире, український відповідник (відповідники): **Волнуха**, гриб *Agaricustorminosus* – вовнянка). Ремарка або коментар може міститися ще в дужках після російського відповідника: **Ораторія** (муз.) – ораторіум. (У багатозначних словах ремарка з латинським відповідником міститься в українській частині реєстру: **Ворона** – 1) пт. *Corvuscornix* – ворона, гава, гава; 2) роззява, гава, гава (про чоловіка).)



Рис. 2. Модель словникових статей типу «Волнуха»

3. Різні значення полісемічних слів подано через крапку з комою із нумерацією: **Вооруженіє** – 1) узброєння, озброєння; 2) зброя. (В українській частині реєстру можуть бути коментарі в дужках після певних слів. Їх до уваги не беремо: **Воротца** – 1) ворітця, ворітки; 2) хвіртка, хвірточка (у воротях).)



Рис. 3. Модель словникових статей типу «Вооружение»

4. Перед низкою українських відповідників, які подано після основного значення, може бути в дужках ремарка або пояснення, що вказує на певні семантичні відтінки або пов'язані з реєстровим словом поняття. Оформлено це так: **Гвоздь** – 1) гвіздок, цв'ях, (для підків) – ухналь, (без головки) – шпень, шпеник, (для шин) – шиналь, (гонтовий) – гонталь, (великий і тонкий) – бринталь, (дерев'яний) – тибель, (у шевців) – шпилька, (набитий цв'яшками) – цв'яхований, цв'яшкований; 2) кілок, кілочок.

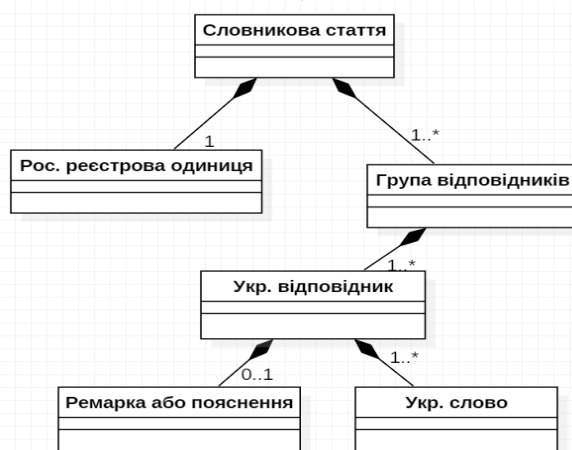


Рис. 4. Модель словникових статей типу «Гвоздь»

5. Біля іменників чоловічого роду в російській частині через кому подано квазіфлексію жіночого роду. В українській частині слова подано або повністю, або також лише квазіфлексії: **Воспитанникъ, ца** – вихованець, вихованка, годованець, годованка, плеканець і **Вредитель, ница** – шкодник, ця. Аналогічно подано дієслова на -ти й на -ся, а також прислівники, утворені від прикметників.

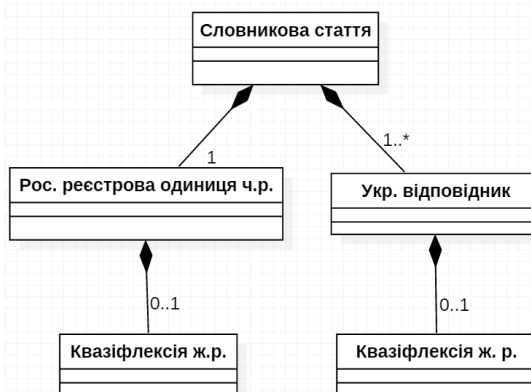


Рис. 5. Модель словникових статей типу «Воспитанникъ», «Вредитель»

6. Поширеними є також статті, що мають таку структуру: російський відповідник (з ремаркою або без) – український відповідник (відповідники), крапка й тире, російське словосполучення (з ремаркою або без неї) – українські відповідники: **Вишня**, рос. *Prunus cerasus* – вишня (дерево і ягоди), вишнина (одна деревина). – **Вишня б'ялая**, *Pavium c.* – черешня, (одна деревина) – черешнина. – **Вишня б'шєнная**, *Atropabelladonna* – вовчі ягоди. – **Вишня жидовская**, *Physalis Alkekengi* – жидівські груші, міхунка.

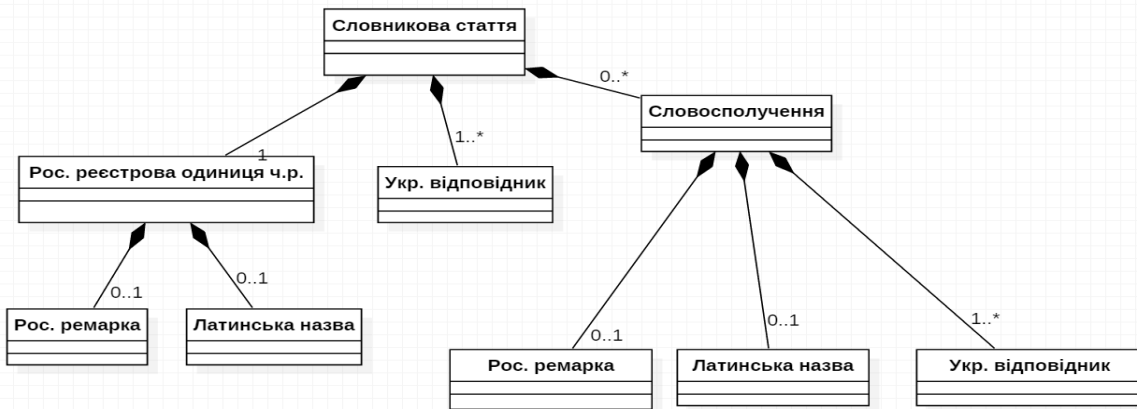


Рис. 6. Модель словникових статей типу «Вишня»

Унаслідок проведених робіт отримано цифрову лексикографічну систему «С. Іваницький, Ф. Шумлянський. Російсько-український словник» із двома реєстрами (російським та українським) (див. рис. 7, 8).

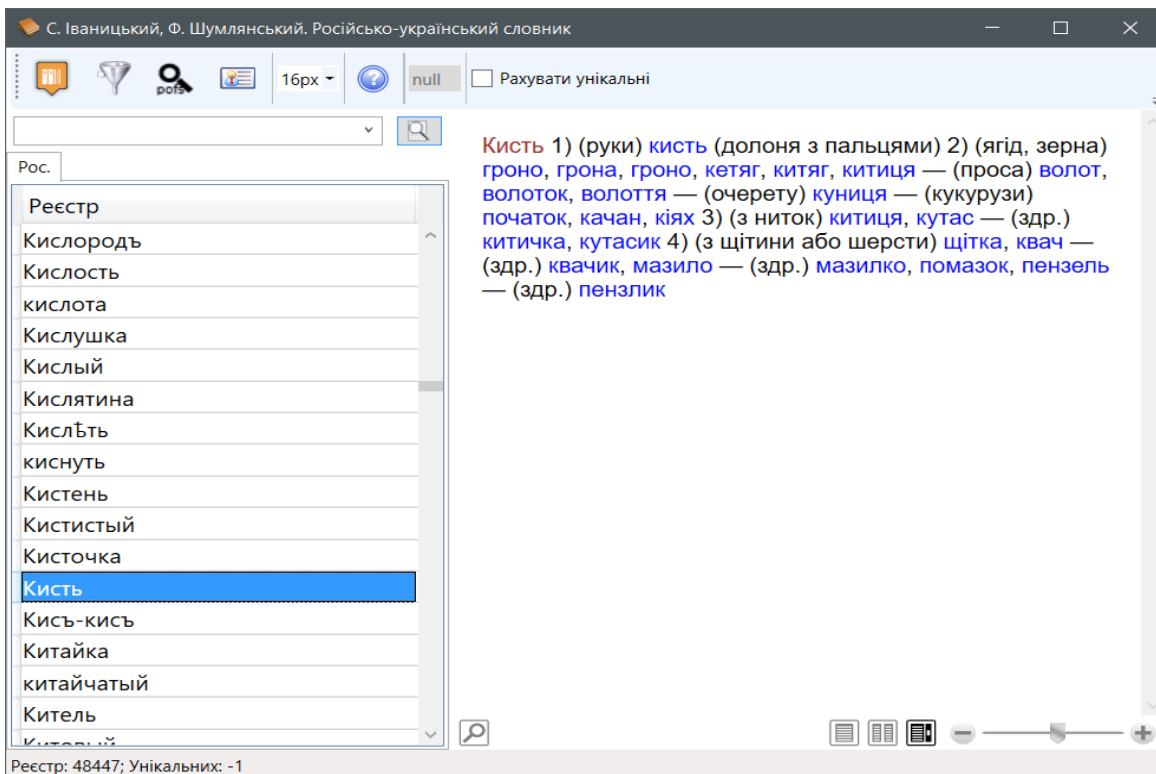


Рис. 7. Російський реєстр словника

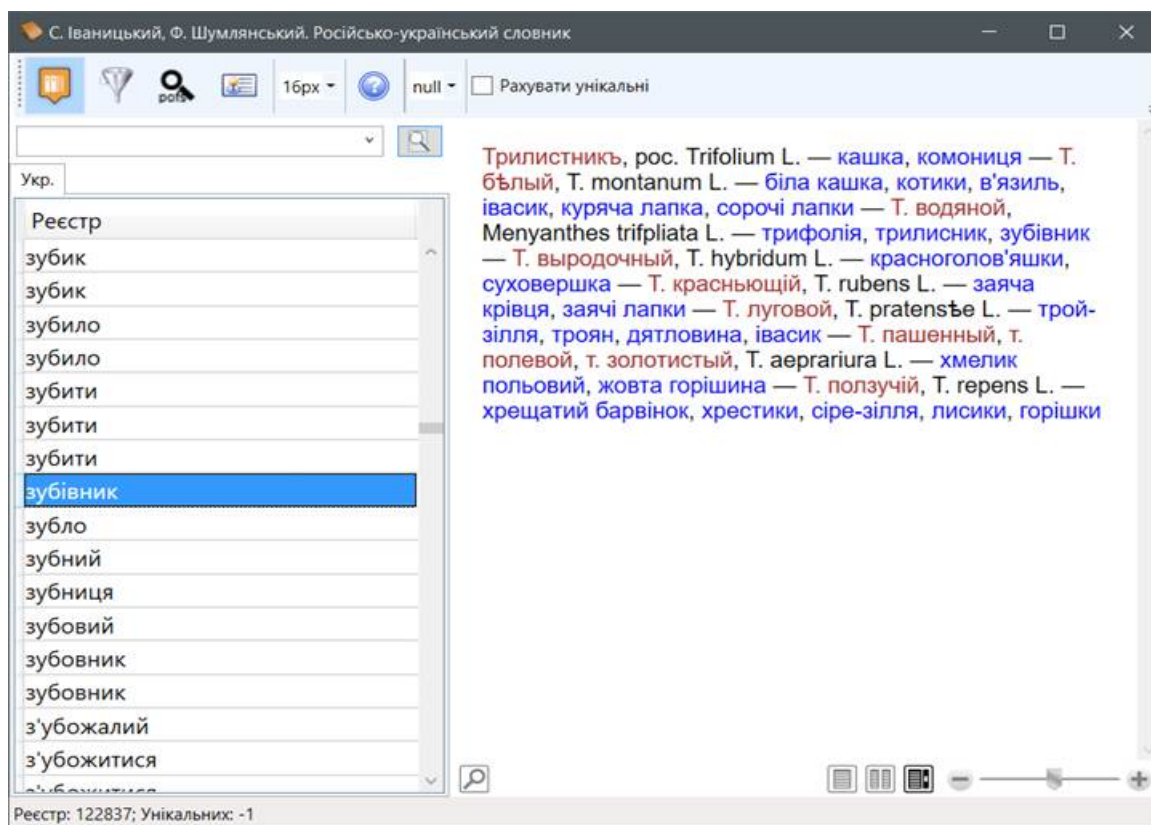


Рис. 8. Український реєстр словника

Система дозволяє виконувати фільтрацію реєстрових слів за частинами, ремарками тощо (див. рис. 9).

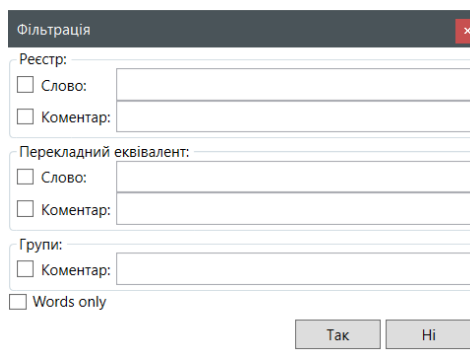


Рис. 9. Фільтрація реєстрових слів за частинами, ремарками тощо

До системи інтегровано граматичний словник, що дає змогу фільтрувати слова за частинами мови, а також уручну виконувати граматичне маркування неідентифікованої лексики (див. рис. 10, 11).

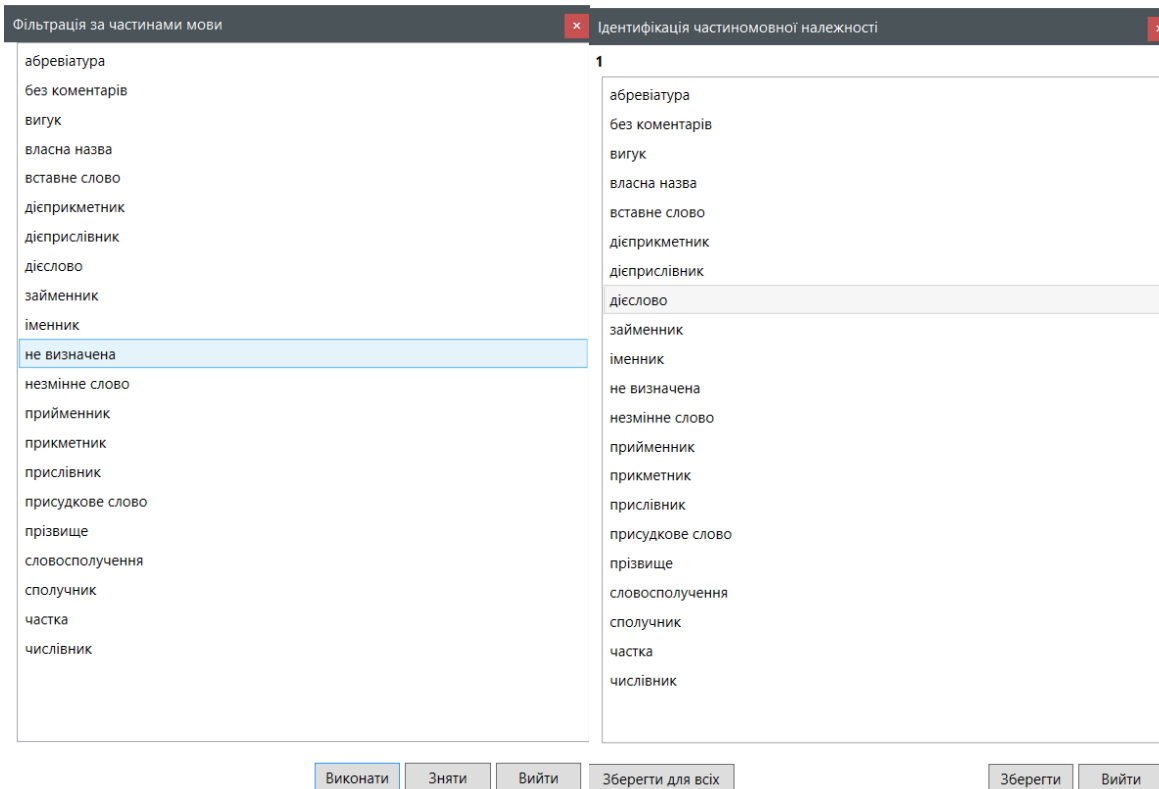


Рис. 10, 11. Фільтрація реєстрових слів за частинами мови та граматичне маркування неідентифікованої лексики

Отже, за допомогою створеної інформаційної системи можна здійснити комплексний аналіз лексики української частини реєстру словника, порівняти її з одиницями, що містяться в лексикографічних джерелах цього ж періоду та в сучасних лексиконах.

#### БІБЛІОГРАФІЯ

1. Масенко Л. Передмова / Лариса Масенко // Російсько-український словник / Укл. С. Іваницький, Ф. Шумлянський. – К., 2006. – С. 4–6.
2. Перебийніс В. І. Традиційна та комп'ютерна лексикографія: [навчальний посібник] / Перебийніс В. І., Сорокін В. М. – К.: Вид. центр КНЛУ, 2009. – 218 с.
3. Томіленко Л. М. «Словник московсько-український» В. Дубровського: склад і структура / Л. М. Томіленко, О. Г. Рабулець // Мовознавство. – 2015. – № 4. – С. 89–95.

#### ВІДОМОСТІ ПРО АВТОРІВ

**Людмила Томіленко** – кандидат філологічних наук, старший науковий співробітник відділу лінгвістики Українського мовно-інформаційного фонду НАН України.

*Наукові інтереси:* лексикологія, лексикографія, термінознавство.

**Олександр Рабулець** – кандидат технічних наук, перший заступник керівника УКС ВР України.

*Наукові інтереси:* інформаційні технології, лінгвістичні технології.