

Сергій Фокін  
м. Київ, Україна

### Напівавтоматична вибірка синтаксичних паралелізмів з текстових корпусів за допомогою мови запитів CQL

The article focuses on general principles of building queries aimed at extracting syntactic parallelisms from textual corpora based on SketchEngine or NoSketchEngine interfaces. The indicated stylistic device is based on diverse kind of repetitions of either lexical stems or grammatical functions. These formal properties can be formulated by means of a machine understandable query language, such as CQL or similar. However, some queries do not yield an exhaustive selection of a stylistic device in question due to homonymy, part of speech ambiguity, tagging errors, considerable number of separators between stylistic device components.

*Keywords:* syntactic parallelism, query language, corpus, morphological annotation, stylistic device.

Паралелізм – колоритне, насичене різними шарами сенсу явище, сповнене глибокої експресії. Саме тому для корпусної лінгвістики стає справжнім викликом розгляд паралелізму під кутом зору його формальних властивостей, зрозумілих для опрацювання комп'ютером. Дослідники цього стилістичного явища у цілому сходяться щодо його формальних та змістових характеристик. Зокрема, Л.І. Мацько зазначає, що «паралелізм (від гр. *parallelissmos* – зіставлення, порівняння) – фігура мови з одного складного або кількох речень, побудована на паралельному використанні (повторі) однієї і тієї ж синтаксичної структури з тотожною модальністю (спонукальною, питальною чи оповідною), однаковим порядком слів та ідентичним інтонаційним малюнком. (...). Граматичні ознаки паралелізму виявляються в однаковій будові складних речень, а всередині складного – простих, в однаковій кількості членів речення і однотипності їх зв'язків та граматичних форм вираження (видо-часових форм дієслів, відмінкових форм іменників тощо)» [1, с. 459]. І справді, такі збіги не є рідкісним явищем у поезії, напр.:

From the bonny bells of heather  
They brewed a drink long-syne,  
*Was sweeter far than honey,*  
*Was stronger far than wine*

(Robert Louis Stevenson)

Л.І. Мацько додатково вказує, що «в основі паралелізму лежить семантико-синтаксичне зіставлення образних ситуацій зі сфер природи й людського життя, їхня взаємозумовленість як загального й часткового, вічножиттєвого і швидкоплинного» [1, с. 459]. У наведеному нижче прикладі з вірша П. Верлена завдяки використанню стилістичної фігури справді проводиться паралель між явищем природи і сентиментальним переживанням людини:

- Ah ! *les premières fleurs, qu'elles sont parfumées!*  
Et qu'il bruit avec un murmure charmant  
*Le premier oui qui sort de lèvres bien-aimées !*  
(Paul Verlain)

Автор відомого «Поетичного словника» А. Квятковський спостерігає, що «синтаксичний паралелізм є найбільш поширеним і полягає у тому, що у суміжних віршах спостерігається однакова структура речень, на відміну від строфічного паралелізму (який також характеризується інколи й лексичною будовою віршів), і ритмічного паралелізму [2, с. 193-194].

Звичайно, для автоматизованої вибірки синтаксичних паралелізмів доцільно скористатися корпусами, які є синтаксично анотованими (тобто, в яких за кожним слововживанням закріплено тег з його синтаксичною функцією). Іншими словами, синтаксичний текстовий корпус – це такий корпус, програмне забезпечення якого може виявляти тип синтаксичного зв'язку між словами, а також синтаксичну роль того чи іншого компонента речення. Однак на сьогодні синтаксично розмічених корпусів відносно небагато, у той час як значна кількість корпусів є морфологічно анотованими. Спеціальні мови запиту, такі як *CQL*, *CQP* та подібні, дають широкі можливості добування синтаксичних даних навіть без явної синтаксичної розмітки, за умов наявності морфологічної розмітки. Морфологічно розмічений корпус – це корпус, програмне забезпечення якого може співвідносити різні граматичні форми лексеми з її початковою формою, містити відомості щодо граматичних ознак та частиномовної характеристики словоформи. Отже, мета дослідження – створити запити для напів-автоматичної вибірки синтаксичних паралелізмів з анотованих корпусів за допомогою морфологічної розмітки засобами *CQL* (*Corpus Query Language*). Поза тим, дослідники помічають, що синтаксичний паралелізм завжди супроводжується морфологічним [3, с. 8]. Оскільки збіг синтаксичних функцій слів у двох реченнях може бути випадковим і не становити стилістичну фігуру, для оцінки експресивної

функції потрібен аналіз експерта. Отже, вибірка не може бути цілком автоматизованою.

Сучасні корпуси з морфологічною і частиномовною розміткою дають змогу оцінити синтаксичну подібність речень. Зокрема, будемо виходити з того, що у слів синтаксично подібних речень будуть збігатися не лише синтаксичні функції, а й морфологічні та частиномовні характеристики слів.

Оскільки слова (та будь-які інші сполучення символів), з яких складається корпус, називаються токенами, надалі у статті використовуємо саме цей термін і «слово» як синоніми (докладніше зі значенням термінів, пов'язаних з використанням корпусів, можна ознайомитися тут [4]).

В інтерфейсі *NoSketchEngine*, *SketchEngine*, яким послуговується Генеральний регіонально анований корпус української мови (ГРАК) [5], морфологічні характеристики найчастіше містяться у тегах, і наступний запит дозволяє вибрати однослівні синтаксичні паралелізми:

```
(<s>1: [] 2: [] </s>)(<s> 3:[] 4:[] </s>) & 1.tag=3.tag & 2.tag=4.tag
```

Так, якщо обмежити цей запит пошуком у поетичних текстах, можна отримати конкорданс, частину якого показано на скріншоті:

Глого народження день ! Безсмертен смертний .	Салют ! Салют !	За вічний день народження ! </doc><doc>
ловом дивиться око , Шкляється холодні сльози .	Бог ? Сатана ?	— Не знаю . Кличеш і маниш до себе . Звс
алась душа в одчаю ; Кров'ю чорний і потом . . .	Бог ? Сатана ?	— Не знаю . . . Довго ще ? Надиш , зводиш
каменем лежить серед пустелі Нікчемне серце .	Випито . Розлито .	Виходь , мій принце . Над вінком Офелії Х
итись вже ніколи і нікому ? . . . Над чим боліти ?	Випито . Розлито .	* Зруйновано , сплюндровано нещадно Мі
в смерть , пірнути під удари ! Ні . ' Не вбіжу .	Повзу . Очунюю .	Мовчу . Дивлюсь . Мій край . Моє старе по
ерть , пірнути під удари ! Ні . ' Не вбіжу . Повзу .	Очунюю . Мовчу .	Дивлюсь . Мій край . Моє старе подвір'я . Е
кто він такий ? Напружено розгадую , — втікач ? Шаленець ? Злодій ? Дезертир ?	Чим гнаний він кризь цей жахли	?
Напружено розгадую , — втікач ? Шаленець ? Злодій ? Дезертир ?	Чим гнаний він кризь цей жахливий шир ?	?
ганьбу , в мій жах . — ... Ні , добреду . Не вмру .	Дійду . Долізу .	Страшніше перейшов , трудніше переліз .
йому кулю ввігнав ? Я все розповім тобі , Боже .	Все . Все .	Все . Цю землю , як шлюбне ложе , зненаг
кулю ввігнав ? Я все розповім тобі , Боже . Все .	Все . Все .	Цю землю , як шлюбне ложе , зненависть
зкритий , як чорна яма , вирита втомою в мурі .	Мури ! Мури !	Над його сном тяжіли поверхи каменю , і в
сонце біле . </doc><doc> ЖІНОЧЕ З річки ішла .	Паленіла . Кохала .	Сіяла мревом . Пахла світанком тіла , Ми

(вибрано з корпусу ГРАК-10 [5])

З конкордансу видно, що паралелізм часто є буквральним повтором. При цьому ми не демонструємо приклади, в яких повторено суто знаки пунктуації або інші, не власне морфологічні характеристики слів.

Пояснимо крок за кроком функції усіх компонентів запиту.

1. Відкриваючий і закриваючий тег з літерою s (позначки в трикутних дужках) означає початок і кінець речення:

```
<s> </s>
```

Відповідно частина запиту

```
<s>1: [] 2: [] </s>
```

означає, що необхідно знайти речення з двох токенів (токен номер один і токен номер два). Для пошуку речення мінімальної довжини (однослівного речення) потрібні два токена, оскільки знак пунктуації у кінці речення теж є окремим токеном. Нумерація токенів довільна; номери токенів потрібні, щоб у запиті до них можна було повторно звернутися.

2. Аналогічно частина запиту

```
<s> 3:[] 4:[] </s>
```

означає, що необхідно знайти також речення, що складається з токена номер три і токена номер чотири.

```
<s>1: [] 2: [] </s><s> 3:[] 4:[] </s>
```

Таким чином, речення з двох токенів (зокрема, речення з токенів номер 1 і номер 2 та речення з токенів номер 3 і номер 4) стоять посліпль у тексті, одне за одним, як відбувається у більшості паралелізмів.

3. Далі необхідно проаналізувати характеристики раніше згаданих токенів. Для цього необхідно написати номер токена, а після крапки – атрибут *tag*. Частина запиту, що вводиться символом & означає «за умови, що»:

```
& 1.tag=3.tag & 2.tag=4.tag
```

А саме, за умови, що тег токена номер 1 є таким самим, як і тег токена номер 3 (морфологічні характеристики слова першого речення збігаються з морфологічними характеристиками слова другого речення), і тег токена номер 2 є таким самим, як і тег номер 4. Оскільки останнім токеном у речення буде знак пунктуації, це означає, що обидва речення мають закінчуватися однаковими знаками пунктуації.

4. Наступний запит дозволяє знайти мінімальні (однослівні) паралелізми (речення з подібною синтаксичною структурою), якщо припустити, що речення з рівнотегованими токенами мають подібну синтаксичну структуру:

Наступний запит схожий на перший, однак у ньому додано по ще одному токеноу до кожної речення, отже, він спрямований на вибірку двослівних паралелізмів:

```
<s>1: [tag!="punct"] 2: [] 3:[]</s><s> 4:[] 5:[] 6:[] </s> & 1.tag=4.tag &  
2.tag=5.tag & 3.tag=6.tag
```

Окремі його результати можна побачити на скріншоті:

ерить його звідси ! Заберіть його !	Заберіть його . Заберіть його .	Геть — геть — геть ! Вони виють перед
ть ! Вони виють перед усмішкою .	Перед зневагою . Перед упевненістю .	Перед мовчанням . Так познайомився
буття . Люди крові одної зі мною !	Захищайте життя ! Захищайте життя !	III Десь із півночі , звідти , 3-за полярнк
тегонне віття жене неситу снагу .	Не кличеш . Не молиш .	Лучиш — льодова ловитва в очах . В м
ас не може спинити , Ні продати .	Ні бити . Ні обдурити .	</doc> <doc> СІНБОНТЕ ' СПІБА НА Т
ви плоди плодили людям гречні .	Розум клали . Розум клали .	Розумом поїли . Боже милий , а ми , а
го він , мій бог , моя висока віра ?	Камінний ідол ? Глиняний бовван ?	Юнак прекрасний з головою звіра ? Чи
им виском чи зашморгом на шії !	Я жду . Я виглядаю .	Я стою . В гігантській , наче вибух , без
ом чи зашморгом на шії ! Я жду .	Я виглядаю . Я стою .	В гігантській , наче вибух , безнадії над
дезертир , чи звір він , чи мара ?	Я жду . Я виглядаю .	Я стою в дідівському розбитому раю ні
тцах розтягує , і смерть мовчить .	Ні поруху . Ні стогону .	Лиш трупи з дивною і впертою увагою
Хай буде так ! Навіки час заляк .	Навіки пустота . Навіки самота .	Тінь хреста . Буде так . Буде так : струх
вузкий розум Лічильних машин .	В пориві . В пориві .	В пориві Плач Африки . У наших руках
і . Спиніться ! Ніхто не звитяжив !	Досить маршів ! Досить овацій !	Скиньте з грудей моїх вражу Тріумфал
іг Так як оті , які за себе дбають ?	Не дбав . Не міг .	І ось , згорнувши сніг , У мерзлоті діру
не значить , що я хвора людина .	Це неправда . Це нісенітниця !	Я зовсім здоровий хлопець . Те , що в

(вибрано з корпусу ГРАК-10 [5])

Оскільки насправді паралелізми можуть складатися з більше, ніж одного та двох слів, немає сенсу для речень різних обсягів створювати окремий запит. Врахуємо також, що в паралелізмах синтаксичний малюнок повторюється не завжди ідеально: морфологічна характеристика двох-трьох слів може збігатися, в той час як в одне з речень можуть додаватися інші вкраплення.

Наприклад, у цій дитячій поезії, що її використав Льюїс Керрол в «Алісі», збігаються морфологічні характеристики перших дев'яти tokenів; восьмий token збігається частково, оскільки в першому реченні він є неозначеним займенником, а в другому – означеним артиклем (але в багатьох корпусах ці різні частини мови будуть анотовані однаково як детермінативи):

*The Queen of Hearts,*  
*she made some tarts,*  
 All on a summer's day;  
*The knave of Hearts,*  
*he stole the tarts,*  
 And took them clean away  
 (Lewis Carroll)

к вас ? я все забуваю , як ваше ім'я й по- батькові .	Артемій Филипович . Артемій Филипович .	Хлестаков . Зробіть ласку , Артемію Филиповичу
еред бурєю . Чуті , ніби навіть буревісників клеїт .	Візьму ! ... Візьму ! ...	Ще крок ... Протягається рука , щоб зрвати покрив
і обох сторін . З боку українців : полк . Микитка сот .	Льобковець сот . Тавчер сот .	Вершні сот . Вознякі я . 42 З боку денкінської армі
з боку українців : полк . Микитка сот . Льобковець сот .	Тавчер сот . Вершні сот .	Вознякі я . 42 З боку денкінської армі ! ген . Шта
вни , а то незабаром на річці буде видно як удень	В човни ! В човни !	Не можна гаяти ні хвилини . Прощатись було ніко
гросові дали підсумок і пару пістолетів та кинджал .	Склали койки . Наставили гармати .	Зарядили мушкети . Приготували сокири та гаки .
ієркул » ; Катрі Гриневичевої — « Шестирилець » .	Митрополит Кирило . Єпископ Серапіон .	Згодом , як проминули перші потрясіння , також у
. Залпаканий Василько недого шльохав і хлипав .	Не втерпів ! Не всидів !	Обкрутив п'яту шкурлатом , стиснув у рученятах і
іть визначити , в якому напрямі їх несе . На північ ?	На південь ? На схід ?	На захід ? Хто знає , де вони опиняться завтра . /
брат ?.. Єдиний ?.. А може також уже одружений ?	З родиною ? З родиною ...	А як не знайде і звалиць , мов той Одисей по маң,
» . Професор міжзоряної авіації Нагірний . Прошу .	Будь ласка . Будь ласка ...	У нього був такий комічний вигляд , що навіть Вас
в на місці ока засохлу шаркалужину ... А де ж око ?	Боже мій , Боже мій !	Невже воно витікло ... І батько відмоуче теплою
юстору в хаті , де скрізь тільки : " Не руш ! Посидь !	Не гарцюй ! Не займай !	Розбісь ! Пореш !.. " І скрізь , що не візьме , куд
милій , мій дорогий , мій одинокий ! Куди ідеш ?..	На Варшаву ? На Варшаву ?	Господи ! І нащо тобі та Варшава ? Він отак , чуєт
іж людьми , не догнав ... Де там доженеш машину .	То сила ... То сила ...	То , чуєте , страшна сила ... Кинув торбу з хлібом
званий володар тягнувся багнюкою великої імперії .	На північ ! На північ !	В лиси ! Обернемо вісь бігуна і пустим планету на
о Тиллявки прибігли перші втікачі . — Палять ! Бда !	Ховайте нас ! Рятуйте нас !..	Пропали ми !.. — В ліс ! В хащі ! А димові хмари і

(вибрано з корпусу ГРАК-10 [5])

У цьому запиті

$\langle s \rangle 1: [\text{tag!}=".*\text{punct.}*" ] \{0,1\} 2: \square \{0,1\} 3: \square \{0,4\} \langle s \rangle 4: \square \{0,1\} 5: \square \{0,1\} 6: \square \{0,4\} \langle /s \rangle \& 1.\text{tag}=4.\text{tag} \& 2.\text{tag}=5.\text{tag} \& 3.\text{tag}=6.\text{tag}$

також враховано, що паралелізми не обов'язково передбачають ідентичні структури речень, а лише подібні, тому й однаковими тегами мають характеризуватися не всі токени, а переважна частина з них. Зокрема, обов'язковою умовою у цьому запиті є збіг тегів, щонайменше, трьох tokenів (номер 1 і номер 4, номер 2 і номер 5, номер 3 і номер 6). Між токенами, що явно збігаються в обох реченнях, можуть розташовуватися й інші токени, що відрізняються. Так, у мові *CQL* [6] пара порожніх квадратних дужок  $\square$  означає будь-який токен, а числа у фігурних дужках означають діапазон кількості повторів попередньо вказаного токена. Зокрема, ця частина запиту

$\square \{0,4\}$

означає послідовність від 0 до 4 будь-яких tokenів. Врахування діапазону, починаючи від нуля, дозволяє врахувати й ті випадки, коли слова, що переривають паралелізм, відсутні. Але якщо вони є, то номери tokenів у запиті відповідатимуть їхнім позиціям у реченнях. На скріншоті видно початок конкордансу з такими паралелізмами, серед яких трапляються структури, в яких не однакова кількість слів, при цьому однаковий або схожий синтаксичний малюнок.

Тяжко мені жить ! Тяжко мені плакати !  
 Чому не п'є з козаками ? Чому не співає ?  
 І холод . І сонце . І буде дощ .  
 Серце моє ! Серце моє !  
 Быть беде . Быть беде .  
 Доле , де ти ! Доле , де ти ?  
 Комунізм — глибина , Як найглибші морські глибини . Комунізм — вишина , Як найвищі гірські верши  
 Це Дніпро ! Це Дніпро !  
 Заберіть його ! Заберіть його .  
 Заберіть його . Заберіть його .  
 Перед зневагою . Перед упевненістю .  
 Ми з нею тут заприсяглись В любові незрадливій . Ми з нею вперше обнялись На колосистій ниві  
 Гай-гай , єдиного ! Гей-гей , єдиного !  
 Гай-гай , єдиного ! Гей-гей , єдиного !  
 Захищайте життя ! Захищайте життя !

(вибрано з корпусу ГРАК-10 [5])

Отже, збіг морфологічних тегів трьох токенів видається цілком достатньою умовою для напіваавтоматичної вибірки паралелізмів. Як видно зі згенерованих конкордансів, окрім збігу синтаксичних і морфологічних характеристик слів сусідніх речень, впадає в око дуже частий лексичний повтор, – риса, що заслуговує на окреме детальне дослідження.

Оскільки паралелізми можуть спостерігатися не обов'язково в сусідніх реченнях, а й бути дистактними (перериватися іншим реченням), здійснимо запит, що враховує наявність додаткового речення, яке перериває паралелізм. Зокрема, речення (від початку до кінця) у мові *CQL* позначатиметься так: `<s></s>`, а весь запит з поправкою на нього матиме таку форму:

```
<s>1: [tag!="punct"] []{0,1}2: [] []{0,1} 3:[] []{0,4}</s></s><s>
4:[] []{0,1} 5:[] []{0,1} 6:[] []{0,4} </s> & 1.tag=4.tag & 2.tag=5.tag &
3.tag=6.tag
```

Згенерований конкорданс показує, що справді дистактно розташовані речення мають паралельні синтаксичні структури:

Мені чудово . Якщо навіть я втрачу мову і не зможу кричати про це . Мені чудово .  
 или З мою рідною сестрою з джунглів , Чому вона до мене не приходить Уранці , Зволожена у сивому тумані , З покірливістю і з двома синами  
 О дивне циганське місто ! Жандармерія відходить крізь темний тунель мовчання , поки ти доліваєш . О дивне циганське місто !  
 СПІД « Item : вона вміє шити » . ЛАНС Виходить , коло неї буде з чого жити . СПІД « Item : вона вміє плести » .  
 їсь навколішки , Ставай навкарачки , Збирай павутиння , Сміття й харкотиння Чи в пишний вітальні , Чи в пральні великій Повзи повід крисло , Пі  
 Лише давні пісні ? І це все , що морю відомо ? Лиш гарні давні пісні ?  
 Я не прошу пощади . Я викриваю диктатора — мерця , що править життям живих і хоче бачити голову мою насадженою на кіл . Я не прошу по  
 Добрий день , поля зелені ! Добрий день , бородачі ! Добрий день , скеляста Сьєрро !  
 Добрий день , блакитна мріє ! Добрий день , мій одяг , рис ! Добрий день вам , коси-змі !  
 Добрий день , біійці залізни ! Добрий день вам , добрі дні ! Добрий день , сеньйор , сеньйор !  
 мерть ! Ні , не померти — побачити , як народ перемаже , щоб жить , миром , любов'ю землю напоє , з квітів , добра і свободи світу надіне вінок  
 Що він простий ? То скільки є простих ! Що він величний ?

(вибрано з корпусу ГРАК-10 [5])

Використані запити не є буквальною інструкцією щодо вибірки паралелізму, а радше є ілюстрацією загальних принципів такої вибірки. Це пов'язано з тим, що в різних корпусах, усе ж, запити необхідно адаптувати до атрибутів, наявних у корпусі. Оскільки мова *CQL* передбачає можливість змінної кількості атрибутів та свободи їхніх найменувань, заздалегідь не відомо, які саме атрибути використали розробники, і які можливі значення цих атрибутів. Приміром, в одному корпусі іменник може позначатися як "NOM", в іншому – як "noun", а деяких корпусах – двома різними тегами: "PN" (для власних імен) і "CN" (для загальних назв), та й власне назва самого атрибута варіюється: у корпусі ГРАК морфологічні й частиномовні ознаки містяться в атрибуті "tag", а в деяких інших корпусах – в атрибуті "pos" тощо. Приміром, багато корпусів різних мов світу є на сайті Лейпцизького університету [7].

Оскільки існує доволі багато корпусів, побудованих на інтерфейсах *SketchEngine/NoSketchEngine* з можливістю використання мови *CQL*, проілюструємо, як видозміниться запит, спрямований на вибірку паралелізму, у корпусі English web-au\_eng\_2013 . Оскільки у цьому корпусі немає атрибута *tag*, натомість є атрибут *pos*, запит виглядатиме так:

```
(<s>1: [] []{0,1}2: [] []{0,1} 3:[] []{0,4}</s>)<s> 4:[] []{0,1} 5:[] []{0,1}
6:[] []{0,4} </s>) & 1.pos=4.pos & 2.pos=5.pos & 3.pos=6.pos
```

hought in dismay .	Not as he had loved his mother . Not as he had loved Stephen .	Not the chance of an
l she come back ?	Willow followed after the crying teenager . Willow read the cover .	Willow was wearing a
n , " he grumbled .	" The giant ? " he said uncertainly . " The groundskeeper , " he clarified .	The incident yesterda
aping strangers . "	" Why can't I wear normal clothes ? " " Why don't we try this then ? "	" Why don't you take
hat do you say ? "	What is your estate ? " What is your title ?	What is your universit
not just one heart .	It is not , of course . It is not rigid .	It is often helpful for tl
Society of Jesus ..	This is the heart of Ignatian spirituality . This is the office of consoler .	This is the vision of th
his cellphone on .	At age 12 , success is having friends . At age 20 , success is having sex .	At age 35 , success is
is having friends .	At age 20 , success is having sex . At age 35 , success is making money .	At age 3 , success is
ing in your pants .	At age 60 , success is having sex . At age 70 , success is having friends .	At age 80 , success is
ing through time .	Death of his father . Death of his mother .	Death of his sister Cc

Маємо відзначити, що в корпусі English web-au\_eng\_2013 відсутня морфологічна розмітка, і для співвіднесення структур двох речень ми маємо обмежитись порівнянням лише частиномовних характеристик. Але навіть на тлі таких обмежених даних вибірка достатньо задовільна. Це дозволяє зробити важливий висновок про те, що не стільки точна граматична копія речення, скільки аналогічний



граматичний потенціал його слів дає змогу для вибірки синтаксичного паралелізму. Проаналізуємо докладніше четвертий приклад з конкордансу: “Why can’t I wear normal clothes? – Why don’t we try this then?” Дієслова вжито в різних особах і числах, і, якби в запиті були враховані морфологічні характеристики, ці дієслівні форми були б різними за тегами, хоча й за частиномовною характеристикою вони збігаються. Отже, менш строгі фільтри пошуку не погіршують результати, а навпаки, поповнюють їх вартісними прикладами. І справді, що менше умов пошуку, то більше шансів віднайти такі менш явні з синтаксичного погляду паралелізми, як цей:

*Утомилась завірюха,  
Де-де позіхає;  
Ще б плакала Катерина,  
Та сліз більш немає.*

(Т. Шевченко)

Таким чином, приклади синтаксичного паралелізму як стилістичної фігури, що піддається формалізації, можна вибирати з текстових корпусів напівавтоматичної вибірки. Синтаксичний паралелізм ґрунтується на подібності синтаксичної структури суміжних або майже суміжних речень. Однак за відсутності в корпусі відповідної синтаксичної розмітки, морфологічні характеристики (а часом і лексичні) теж можуть цілком прислужитися як маркери подібності синтаксичного малюнку речень: у синтаксичних паралелізмах слова сусідніх речень мають однаковий або схожий набір тегів. У тегах як правило міститься морфологічна та частиномовна характеристика слова. Схожість або однаковість тегів може спостерігатися не в абсолютно всіх словах сусідніх речень: між морфологічно схожими словами можуть бути й інші слова-роздільники (наприклад, вставлені слова), що не мають граматичного корелята в іншому реченні. Первинні спостереження показують, що умова збігу тегів у трьох слів у сусідніх реченнях є загалом достатньою умовою для того, щоб приклад був паралелізмом. Оскільки синтаксичних паралелізм – явище, притаманне переважно поезії, перед початком вибірки необхідно застосувати фільтр пошуку саме в поетичних текстах. Слід робити поправки на типи атрибутів у том й їхні можливі значення, що в кожному корпусі влаштовані по-своєму. Тому в різних корпусах запити виглядатимуть по-різному, а спільними є лише принципи використання.

На практиці не всі вибрані приклади є стилістичними фігурами: у деяких з них просто наявний випадковий збіг маркерів фігури, часом виявлені псевдопаралелізми є наслідком друкарських помилок у

тексті, омонімії (зокрема, міжчастиномовної). Для оцінки наявності художньої функції прикладів необхідна участь людського експерта. Спостереження над невелико вибраними прикладами, опис їхніх формальних і функціональних параметрів дозволить удосконалювати раніше сформульовані запити, підвищувати точність результатів.

Перспективно поширити отриманий досвід і на вибірку інших стилістичних фігур, заснованих на формальних повторах, а саме: анадиплоз, епаналепсис, анафоричний повтор, хіазм.

### **Література**

1. Мацько Л.І. Паралелізм. *Енциклопедія. Українська мова*. К.: Українська енциклопедія ім. М.П. Бажана, 2000. С. 424.
2. Квятковский А.П. *Поэтический словарь*. М.: Советская Энциклопедия, 1966. 376 с.
3. Новикова Т.В. Синтаксический параллелизм как фигура экспрессивного синтаксиса (на материале современной французской литературы): автореф. дис. канд. филол.наук Текст. Смоленск, 1992. 21 с.
4. Корпус ГРАК: Інструкція до пошуку. Глосарій термінів <http://uasocorpus.org/>
5. Пошук у корпусі ГРАК. [Електронний ресурс]. Режим доступу: <http://www.parasolcorpus.org/bonito/run.cgi/>
6. Мова корпусних запитів SQL. [Електронний ресурс]. Режим доступу: <https://drive.google.com/file/d/1T13VGpFYbzXRYO0tPVZoyGT343e9jiWl/view>.
7. Leipzig Corpora Collection: [http://cql.corpora.uni-leipzig.de/bonito/run.cgi/first\\_form](http://cql.corpora.uni-leipzig.de/bonito/run.cgi/first_form).