

new words, have already turned into semantic neologisms, as they have expanded their semantic structure by forming new lexical and semantic variants.

Keywords: vocabulary; neologism; semantics; secondary nomination.

I. NOSKO

SEMANTIC REORIENTATION OF LEXICS AS THE MAIN METHOD OF CREATING ENGLISH NEOLOGISMS

The article is dedicated to the problem of semantic reorientation of English lexics. The aim of the work is to describe English semantic neologisms and distinguish their role in the modern English language.

Linguists point out that an absolute progress of a language becomes apparent first of all in the development and replenishment of its vocabulary. Changes in society are reflected in the language, because being a dynamic system it quickly reacts to new technologies, developing science and technique, broadening the geography of business contacts by extension of its vocabulary. All media, including radio, television and newspapers, are in constant need of new terms, especially when new concepts are introduced to a certain culture.

The most recognized classification of neologisms is the classification of Lui Guilbert: new words can be divided into 4 groups according to the ways they are formed. They are subdivided into phonological neolo-

gisms, borrowings, semantic neologisms and syntactical neologisms. V. Zobotkina singles out proper neologisms, semantic neologisms and transnominations.

There is a certain correlation between the type of a nominative unit and the way of its coinage. From the point of view of their creating proper neologisms, units that combine new form with the new meaning are phonological, morphological neologisms and borrowings. Transnominations that combine new form with the meaning already existed in another form appear because of borrowing and morphological word formation. Semantic innovations that combine new meaning and the existed form appear because of the changes in their meaning.

It was proved that today the overwhelming majority of new words belong to the last group. L. Ganyak has estimated that the number of semantic neologisms exceeds 75% of the total number of neologisms. Yu. Zatsnyi claims they make up at least 25%. That is conditioned by the fact that besides creating the new meaning, there are numerous cases of appearing of new shades of meaning, new connotations and new combinatory variants. Such phenomena are much more difficult to notice and to register in the dictionaries.

It is necessary to point out that there is a tight connection between the semantic innovations and lexical, phraseological neologisms: acquisition of a new meaning is accompanied by new word and idiom collocations.

Стаття надійшла до редакції 10.05.2017

УДК 81'367:81'373.7: 81'32

СИТАР Ганна

АЛГОРИТМ РОЗПІЗНАВАННЯ СИНТАКСИЧНИХ ФРАЗЕОЛОГІЗМІВ (ДЛЯ ПОСЛІДОВНОСТІ ЧОТИРЬОХ СЛІВ)¹

У статті здійснено спробу побудови алгоритму розпізнавання синтаксичних фразеологізмів для випадків чотирьох слів, розташованих у тексті поспіль. Спираючись на результати попередніх теоретичних досліджень, виділено формальні і статистичні параметри, за якими можна відмежувати синтаксичні фразеологізми від інших мовних одиниць у корпусі українськомовних текстів.

Ключові слова: алгоритм; корпус текстів; модель речення; синтаксичний фразеологізм; українська мова; фразеологізоване речення.

Вступ. Синтаксичні фразеологізми (або фразеологізовані речення) — особливий тип речення, у якому стрижневий і змінний компоненти пов'язані ідіоматично і розташовані переважно фіксовано, граматичні зв'язки і прямі лексичні значення слів послаблені [2; 4; 11;

12; 13; 14], наприклад: *Що за день! Чим не подарунок?! От вам і відповідь. От же ж новина! Злочин є злочин. Треба так треба. Теж мені молодці! Яке там успішно! Де вже там встигли! Дружба вона і є дружба. Кохання і в Африці кохання* та ін. Такі речення належать до дієвих

¹ Дослідження виконане в межах програми фундаментальних досліджень Міністерства освіти і науки України (проект № 0115U000088 «Комунікативно-прагматична і дискурсивно-граматична лінгвоперсоналогія: структурування мовної особистості та її комп'ютерне моделювання»).

засобів експресивного синтаксису, за допомогою яких мовець виражає ставлення до певної реальності або ситуації дійсності, прагне досягнути впливу на почуття адресата мовлення, сформувані в нього потрібну думку або викликати відповідну дію.

Аналіз досліджень і публікацій. Грунтовний аналіз синтаксичної будови та семантики українських фразеологізованих речень та розмежування трьох ступенів фразеологізації речень здійснили М.І. Личук та В.Д. Шинкарук [8, с. 9]. Особливості окремих груп синтаксичних фразеологізмів висвітлено в межах досліджень відповідних семантичних або комунікативних категорій: ствердження і заперечення (праці М.П. Баган, О.П. Куш, А.Й. Паславської), згоди (дослідження Н.К. Войцехівської), комунікативної інтенції (праці С.Т. Шабат-Савки) та ін. [1; 3; 7; 10; 18]. Проте синтаксичні фразеологізмидосі не розглядали в аспекті автоматичного аналізу.

У комп'ютерній лінгвістиці вважається, що фразеологічні одиниці (у широкому розумінні цього терміна) становлять один вузол, відповідно їх вводять до електронних словників програм автоматичного аналізу тексту або автоматизованого перекладу в готовому вигляді, інколи із перерахуванням можливих варіантів. Вчинити так само із синтаксичними фразеологізмами досить проблематично через те, що будь-який синтаксичний фразеологізм складається з двох компонентів — стрижневого (незмінного) та змінного, при цьому лексичне наповнення і морфологічна варіантність останнього можуть бути настільки ємними, що її неможливо описати вичерпно та зафіксувати у словниках.

Розглянемо такі речення:

Горпина. *То й добре! Сідайте ж, будьте ласкаві!*

Бублейниця. *Еге! Де вже там сидіти! Мені аж танцювати хочеться, так весело було в су-сіді!* (І. Нечуй-Левицький. На кожум'яках)²;

Був він худенький і такий мерзлякуватий, що у найжаркіший день, тільки залізе у воду, через хвилину вже цокає зубами і, обхопивши себе руками навхрест за плечі, дрібно тремтить, як цуценя під дощем. Де вже там плавати, як просто у воді перебувати несила (В. Нестайко. Одиниця з обманом);

Софрон. *Ну, прощай Василю! (Обніма і цілує її). Не сумуй же!*

Отаман. *Де вже там не сумувать* (І. Карпенко-Карий. Підпанки);

...це було лице, яке Пігмаліон бачив на своїх колінах тієї літньої ночі у мерехтливих світанкових відблисках, яке він цілував і пестив, обличчя, про яке він казав: «Ти така гарна...» — а вона посміхалася: «Де вже там гарна? Довгоноса, пащекувата. Як моя мама каже — на любителя...» (В. Дрозд. Пігмаліон);

— *Що тобі, моя дитино! — упада вже коло Катрі мати. — Уступила до хати біла, мов білочок... що тобі, доню?*

— *Нічого, мамо, нічого!*

— *Де вже там нічого! Журиш ти мене, дочко!* (М. Вовчок Три долі);

...Єлені стало ясно — де вже там пробитися через ту пустелю й дістатися до свого справжнього давнього Олекси (Г. Хоткевич. Авірон).

У наведених реченнях стрижневим є компонент *де вже там*, а змінним — *сидіти, плавати, не сумувать, гарна, нічого, пробитися, дістатися* (із залежними компонентами *через ту пустелю, до свого справжнього давнього Олекси* відповідно).

На сьогодні наявні спроби створення алгоритмів розпізнавання стійких сполук у широкому розумінні, а не синтаксичних фразеологізмів як окремого типу. Зокрема, у праці В.Д. Гусева та Н.В. Саломатіної, виконаній на матеріалі російської мови, запропоновано алгоритм для біграм — сполучень двох слів у тексті. Визначення їхньої стійкості ґрунтується на кількісному параметрі, а саме на відношенні частоти біграми до частоти триграми — поєднання біграми і слова зліва та поєднання біграми і слова справа [5]. Дослідники пропонують вважати ланцюжок слів стійким за умови, «якщо кількість усіх можливих різних його контекстів (як ліво-, так і правобічних) є достатньо великою, наприклад, зіставляваною з частотою вживання цього ланцюжка», а зазначені відношення частот $\leq 0,5$ [5]³.

Спроба реалізації формалізованого підходу до ідентифікації синтаксичних фразеологізмів дозволяє перевірити, чи можуть виділені в теоретичних працях ознаки прислужитися для автоматичного встановлення цих одиниць

² Тут і далі наведено приклади, почерпнуті з Українського національного лінгвістичного корпусу (далі УНЛК) Українського мовно-інформаційного фонду Національної академії наук України.

³ У праці В.Д. Гусева та Н.В. Саломатіної термін «ланцюжок слів» позначає будь-які два або більше слова, що розташовані поруч у тексті й не перервані знаками пунктуації, при цьому не враховано наявності / відсутності синтаксичного зв'язку між цими словами [5].

у тексті. Відповідно такий алгоритм може бути використаний для вдосконалення процедур автоматичного опрацювання тексту та створення систем автоматизованого перекладу, що й мотивує актуальність цього дослідження. У побудові алгоритму для випадків чотирьох слів, розташованих у тексті поспіль, спираємось на попередній досвід з укладання алгоритму для послідовності трьох слів, які охоплюють синтаксичні одиниці із одно- та двочленним стрижневим компонентами [16]. На нашу думку, для створення алгоритму недостатньо тільки структурних або тільки кількісних параметрів аналізованих одиниць, тому ми пропонуємо застосовувати їх поєднання.

Постановка мети і завдання. Мета цього дослідження — побудувати алгоритм розпізнавання синтаксичних фразеологізмів (для послідовності чотирьох слів у тексті). Відповідно до аналізу залучені речення із дво- та тричленним стрижневим компонентом. Поставлена метаумовлює виконання трьох завдань: 1) виділити ознаки аналізованих синтаксичних фразеологізмів, придатні для формалізації; 2) запропонувати алгоритм розпізнавання в тексті синтаксичних фразеологізмів з урахуванням виділених ознак та різних типів цих одиниць; 3) подати пояснення кроків запропонованого алгоритму.

Формальні критерії аналізу синтаксичних фразеологізмів для випадків чотирьох слів, розташованих поспіль. Аналіз фактичного матеріалу, дібраного з текстів української художньої літератури кінця ХІХ — початку ХХІ ст., періодичних видань і українськомовних інтернет-ресурсів, обсягом більше 6000 контекстів уживання синтаксичних фразеологізмів дає змогу виділити кілька формальних ознак.

Наявність повтору в реченні. Повтор лексем у реченні виступає своєрідним організаційним чинником, який «цементує» речення, забезпечує його фразеологізацію (про типи повторів у межах фразеологізованих речень див. у праці [17]). Порівняно із синтаксичними фразеологізмами з одночленним стрижневим компонентом серед обстежених конструкцій таких моделей зафіксовано небагато, проте саме повтор є тією специфічною формальною ознакою, яку можна використати для побудови алгоритму. З-поміж таких речень можна виділити дві групи:

а) речення із повтором тієї самої форми слова — наявний повний збіг словоформ, між якими вживається двочленний стрижневий

компонент *є ... і/ї/та: З «друзями», як СНД, не потрібні вороги. Друзі — Польща, Словаччина, Чеська Республіка, ... хочуть, щоб Україна була в ЄС, для кращого життя українського народу, друзі — Німеччина продає Україні дешевий газ. Друзі — Грузія хоче, щоб Україна вступила в НАТО, є друзі і «друзі»* (www.unian.ua/comments/535333);

б) речення з повтором одного слова (у нашому матеріалі — іменника або дієслова) у різних формах. Якщо перше і четверте слово є різними словоформами слова, то між ними в нашому матеріалі зафіксовано займенник/іменник у непрямому відмінку та частку *не: Ну, ти хоч і старий, а обдурити нас не обдуриш* (Г. Тютюнник. Вир) і под. Якщо різні форми одного слова виступають третім і четвертим компонентом, то до них уживається поєднання прийменника *з/із/зі* та займенника *усіх/всіх: Новий рік — з усіх свят свято* (Л. Нечай-Горбачук. Читанка для Калинки і Тарасика).

2. На підставі аналізу речень без повторів укладено реєстр моделей синтаксичних фразеологізмів із тричленним стрижневим компонентом, який складається з поєднання службових і повнозначних слів — типово займенників та прислівників, для яких характерне семантичне спустошення або семантичний зсув: *де вже там, куди вже мені, от тобі й, чому б не, як тут не, який же там* і под.: *Я сказав йому по щирості свою думку про його поезії, підніс їх добрі прикмети, та zarazом звернув його увагу на конечність їх обшліфування перед друком. Він усміхнувся сумно.*

— *Куди вже мені шліфувати їх!* — говорив, і в його голосі тремтіло щось, мов здавлені сльози (І. Франко);

— *Горнефта нема вдома, — продовжувала каліченою грецькою мовою, додаючи єгипетських слів. — Він повернеться — може, ввечері... а може, тільки завтра.*

— *От тобі й радуйся!* — з досадою вигукнула Гелене (Н. Королева. Предок);

Який же там з нього лицар, коли він усього так боїться... (Ю. Покальчук. Аве Марія).

На сьогодні реєстр синтаксичних фразеологізмів із тричленним стрижневим компонентом охоплює 26 моделей. Із урахуванням можливих родових та числових варіантів відповідних лексем (наприклад, *який же це / яка ж це / яке ж це / які ж це*), чергування голосних (наприклад, *от тобі і / от тобі й*), наявності загальнолітературного інваріанта і розмовного варіанта (зокрема, *чому б не / чом би не, чому*

ж не / чом же не) реєстр стрижневих тричленних компонентів включає 108 складених одиниць (див. табл. 1).

Статистичні критерії аналізу синтаксичних фразеологізмів (для послідовності чотирьох слів у тексті). Статистичним критерієм визначення синтаксичного фразеологізму у тексті став показник асоціації МІ (від англ. mutual information — англ. *взаємна, спільна, повна інформація*) — коефіцієнт, який відбиває не випадковість (залежність) певної послідовності слів у тексті [20]. Оскільки Кеннет Ворд Чарч і Патрік Хенкс ввели формулу МІ для поєднання двох слів, для багатоконпонентних конструкцій застосовуємо модифіковану формулу (1) [19, с. 586], у якій передбачено необхідну кількість складників:

$$MI(c_1, c_2, \dots, c_i) = \log_2 \frac{f(c_1, c_2, \dots, c_i) \times N^{(i-1)}}{f(c_1) \times f(c_2) \times \dots \times f(c_i)}, (1)$$

де i — кількість компонентів конструкції;

c_1 — перша лексична одиниця;

c_2 — друга лексична одиниця;

c_i — i -а лексична одиниця;

$f(c_1, c_2, \dots, c_i)$ — абсолютна частота вживання конструкції c_1, c_2, \dots, c_i в корпусі (з урахуванням порядку одиниць усередині конструкції);

$f(c_1)$ — абсолютна частота c_1 в корпусі;

$f(c_2)$ — абсолютна частота c_2 в корпусі;

$f(c_i)$ — абсолютна частота c_i в корпусі;

N — загальна кількість словоформ у корпусі;

\log_2 — логарифм числа за основою 2.

Здійснені обчислення демонструють, що для синтаксичних фразеологізмів із тричленним незмінним компонентом коефіцієнт МІ перебуває в межах від 19,87 (*куди вже їм*) до 28,4 (*де вже там*). Саме тому одним із кроків алгоритму є визначення МІ для послідовностей трьох слів, які претендують на статус стрижневого компонента синтаксичних фразеологізмів. До останніх зараховуємо ті конструкції, для яких обчислене $MI \geq 19,8$.

Алгоритм розпізнавання синтаксичних фразеологізмів. У запропонованому алгоритмі використано умовні позначки і правила виконання алгоритмів, викладені в Державному стандарті 19.701-90 [6]. Зокрема, термінатор застосовано на позначення початку й кінця роботи алгоритму; прямокутник — для процесу, тобто виконання однієї або кількох операцій; ромб — для рішення або функції перемикального типу з одним входом і двома альтернативними виходами, з яких тільки один може бути активований після виконання вказаних умов.

Таблиця 1

Реєстр тричленних стрижневих компонентів синтаксичних фразеологізмів в українській мові

№ з/п	Стрижневий компонент синтаксичного фразеологізму
1	2
1.	<i>де вже вам</i>
2.	<i>де вже йому</i>
3.	<i>де вже їй</i>
4.	<i>де вже їм</i>
5.	<i>де вже мені</i>
6.	<i>де вже нам</i>
7.	<i>де вже там</i>
8.	<i>де вже тобі</i>
9.	<i>до чого ж</i>
10.	<i>куди вже вам</i>
11.	<i>куди вже йому</i>
12.	<i>куди вже їй</i>
13.	<i>куди вже їм</i>
14.	<i>куди вже мені</i>
15.	<i>куди вже нам</i>
16.	<i>куди вже тобі</i>
17.	<i>ну така вже</i>
18.	<i>ну така ж</i>
19.	<i>ну таке вже</i>
20.	<i>ну таке ж</i>
21.	<i>ну такі вже</i>
22.	<i>ну такі ж</i>
23.	<i>ну такий вже</i>
24.	<i>ну такий же</i>
25.	<i>ну чим не</i>
26.	<i>ну який вже</i>
27.	<i>ну який уже</i>
28.	<i>ну який же</i>
29.	<i>ну яка ж</i>
30.	<i>ну яка вже</i>
31.	<i>ну яке вже</i>
32.	<i>ну яке ж</i>
33.	<i>ну які вже</i>
34.	<i>ну які ж</i>
35.	<i>ну яка там</i>
36.	<i>ну яке там</i>
37.	<i>ну який там</i>
38.	<i>ну які там</i>
39.	<i>он який вже</i>
40.	<i>он який уже</i>
41.	<i>он яка вже</i>
42.	<i>он яка уже</i>
43.	<i>он яке вже</i>
44.	<i>он яке уже</i>
45.	<i>он які вже</i>
46.	<i>он які уже</i>

Закінчення табл. 1

1	2
47.	<i>ось тобі і</i>
48.	<i>ось тобі й</i>
49.	<i>ось вам і</i>
50.	<i>ось вам й</i>
51.	<i>от вам і</i>
52.	<i>от вам й</i>
53.	<i>от тобі і</i>
54.	<i>от тобі й</i>
55.	<i>ото вам і</i>
56.	<i>ото вам й</i>
57.	<i>ото тобі і</i>
58.	<i>ото тобі й</i>
59.	<i>от же ж</i>
60.	<i>от який вже</i>
61.	<i>от який уже</i>
62.	<i>оце вже ж</i>
63.	<i>оце тобі і</i>
64.	<i>оце тобі й</i>
65.	<i>оце вам і</i>
66.	<i>оце вам й</i>
67.	<i>хто, як не/ хто як не</i>
68.	<i>чи ж не</i>
69.	<i>чи тобі не</i>
70.	<i>чи вам не</i>
71.	<i>чим же не</i>
72.	<i>чому б не</i>
73.	<i>чом би не</i>
74.	<i>чому ж не</i>
75.	<i>чом же не</i>
76.	<i>що воно за</i>
77.	<i>що ж за</i>
78.	<i>що це за</i>
79.	<i>що то за</i>
80.	<i>що, як не</i>
81.	<i>як вам не</i>
82.	<i>як же не</i>
83.	<i>як їй не</i>
84.	<i>як їм не</i>
85.	<i>як йому не</i>
86.	<i>як мені не</i>
87.	<i>як нам не</i>
88.	<i>як тобі не</i>
89.	<i>як тут не</i>
90.	<i>яке вже там</i>
91.	<i>яке ж там</i>
92.	<i>який же це</i>
93.	<i>який же то</i>
94.	<i>яка ж це</i>
95.	<i>яка ж то</i>
96.	<i>яке ж це</i>

1	2
97.	<i>яке ж то</i>
98.	<i>які ж то</i>
99.	<i>які ж це</i>
100.	<i>яка вже там</i>
101.	<i>яка ж там</i>
102.	<i>яка уже там</i>
103.	<i>який вже там</i>
104.	<i>який же там</i>
105.	<i>який уже там</i>
106.	<i>які вже там</i>
107.	<i>які ж там</i>
108.	<i>які уже там</i>

Графічно створений алгоритм подано у вигляді блок-схеми (див. схему 1). Він має досить складну структуру та включає всі можливі типи базових структур алгоритмів — слідування, розгалуження та повторення. Відповідно цей алгоритм поєднує в собі ознаки лінійного, розгалуженого та циклічного алгоритмів. Пояснимо кроки цього алгоритму.

На вході алгоритму маємо корпус текстів, з якого вичленовуємо речення. Під реченнями в суто формальному плані розуміємо послідовності кількох слів, відокремлені одна від одної такими знаками пунктуації, як крапка, знак оклику, знак питання, знак питання і знак оклику, три крапки або крапка з комою.

Оскільки синтаксичні фразеологізми аналізованого типу є багатокомпонентними одиницями, що містять чотири і більше компоненти, у межах кожного речення аналізуємо по чотири слова поспіль (усі можливі лінійно розташовані поєднання чотирьох слів).

Далі перевіряємо вичленовані чотири слова на наявність повторів. Спочатку з'ясовуємо, чи реалізовано повний збіг другої і четвертої лексеми.

Якщо так, то чи є першим компонентом дієслівна форма *є*, а третім — сполучник *і*. Якщо ця умова виконана, то речення вважаємо синтаксичним фразеологізмом і переходимо до наступних чотирьох слів у реченні. Якщо у реченні більше немає слів, то беремо наступне речення і починаємо цикл перевірки спочатку.

Якщо умову щодо другого і четвертого компонента не виконано, перевіряємо чи є перше і четверте слово словоформами одного слова. В УНЛК це можна перевірити за збігом квазіоснов. Якщо так, то визначаємо, чи є другим компонентом займенник або іменник,

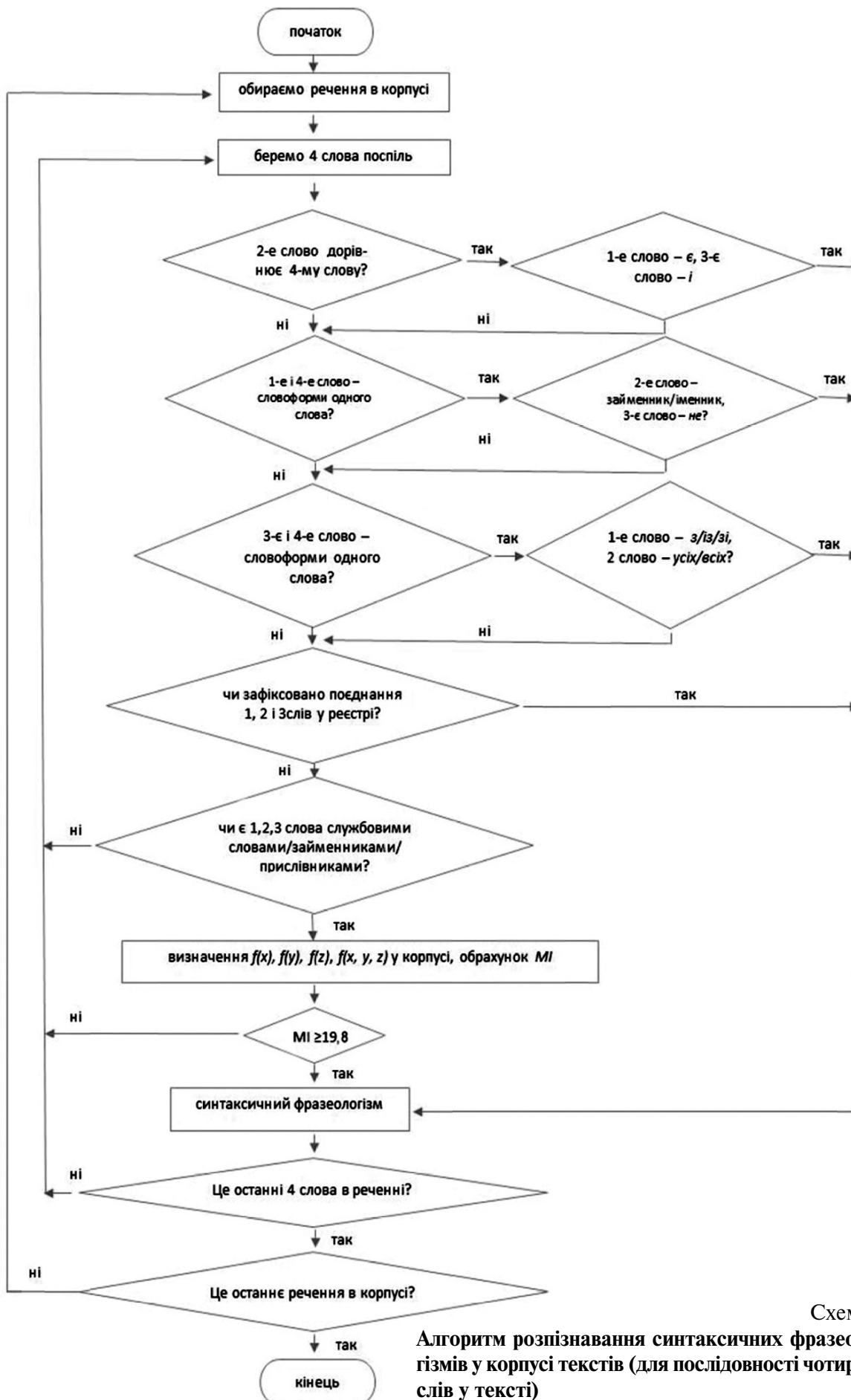


Схема 1
 Алгоритм розпізнавання синтаксичних фразеологізмів у корпусі текстів (для послідовності чотирьох слів у тексті)

а третім — частка *не*. За виконання цієї умови кваліфікуємо аналізовану послідовність чотирьох слів як синтаксичний фразеологізм.

Якщо ні, то зіставляємо третє і четверте слово і встановлюємо, чи належать вони до форм одного слова. Якщо цю умову дотримано, з'ясуємо, чи є перше слово прийменником з (або його варіантами *із/зі*), а друге слово — займенником *усіх/всіх*. Якщо так, то вважаємо, що це синтаксичний фразеологізм.

Далі переходимо до аналізу речень без повторів. Передусім перевіряємо, чи зафіксоване поєднання трьох слів в укладеному нами реєстрі тричленних стрижневих компонентів синтаксичних фразеологізмів (див. табл. 1). Якщо так, то маємо синтаксичний фразеологізм.

Якщо такого поєднання в реєстрі поки що не зафіксовано, встановлюємо частини мовну належність аналізованих слів. В УНЛК це можна зробити за допомогою квазіфлексій. Якщо ці слова належать до службових частин мови, займенників або прислівників, то є потреба в наступному кроці алгоритму.

Для незафіксованого в реєстрі поєднання слів виконуємо таку послідовність дій: визначаємо в корпусі текстів абсолютну частоту першого слова $f(x)$, абсолютну частоту другого слова $f(y)$, абсолютну частоту третього слова $f(z)$ та абсолютну частоту конструкції загалом $f(x, y, z)$, підставляємо отримані числа до формули (1) та обраховуємо коефіцієнт MI .

Якщо отриманий результат $MI \geq 19,8$, аналізовану конструкцію кваліфікуємо як синтак-

сичний фразеологізм і переходимо до аналізу наступних чотирьох слів або наступного речення за їхньої наявності.

На виході алгоритму маємо файл з усіма контекстами вживання синтаксичних фразеологізмів, отриманими на підставі аналізу корпусу текстів.

Висновки. Отже, створений алгоритм розпізнавання синтаксичних фразеологізмів об'єднує дві частини. Спочатку послідовності чотирьох слів перевіряються на наявність повторів слів (з урахуванням двох можливих варіантів — збіг слів повністю або наявність різних форм одного слова) та відповідність визначеному на попередніх етапах дослідження переліку стрижневих компонентів (*е ... і/ї/ма; з/із/зі усіх/всіх, не*).

Друга частина алгоритму призначена для ідентифікації іншого типу синтаксичних фразеологізмів — із тричленним стрижневим компонентом. Вона враховує такі властивості аналізованих одиниць: наявність поєднань службових частин мови, службової частини мови із займенником або прислівником, займенника та прислівника; відповідність поєднань слів укладеному реєстру стрижневих компонентів синтаксичних фразеологізмів; показник асоціації *mutual information* $\geq 19,8$.

На подальшому етапі дослідження плануємо об'єднати дві створені частини — для послідовностей трьох і чотирьох слів у тексті — у єдиний алгоритм і створити робочу версію програми автоматичного розпізнавання синтаксичних фразеологізмів на основі УНЛК.

Список використаної літератури

1. Баган Мирослава. Категорія ствердження — заперечення: нові аспекти дослідження // Науковий вісник Херсонського державного університету. Серія: «Лінгвістика». — Херсон, 2005. — Вип. 2. — С. 171–174.

2. Величко А.В. Фразеологизированные структуры русского предложения // Книга о грамматике: Русский язык как иностранный: Учеб. пособие / Под ред. А.В. Величко. — 2-е изд., испр. и доп. — М.: Изд-во Моск. ун-та, 2004. — С. 37–54.

3. Войцехівська Н.К. Категорія згоди в українському літературному діалогічному дискурсі (на матеріалі художньої літератури ХХ — початку ХХІ століть): автореф. дис... канд. філол. наук : 10.02.01 / Войцехівська Наталія Костянтинівна / НАН України; Інститут мовознавства ім. О.О. Потебні. — К., 2009. — 20 с.

4. Всеволодова М.В., Лим Су Ён. Принципы лингвистического описания синтаксических фра-

зеологизмов: На материале синтаксических фразеологизмов со значением оценки / Майя Владимировна Всеволодова, Лим Су Ён. — М.: МАКС Пресс, 2002. — 164 с.

5. Гусев В. Д., Саломатина Н.В. Алгоритм выявления устойчивых словосочетаний с учетом их вариативности (морфологической и комбинаторной) / Владимир Гусев, Наталья Саломатина // Труды междунар. конф. Диалог-2004, Верхневолжский, 2–7 июня 2004. — М.: Наука, 2004. — С. 530–535.

6. ГОСТ 19.701-90. Схемы алгоритмов, программ, данных и систем. Условные обозначения и правила выполнения [Электронный ресурс]. — Режим доступа: <http://cert.obninsk.ru/gost/282/282.html>

7. Куш О. П. Категорії ствердження і заперечення в українській мові : автореф. дис... канд. філол. наук : 10.02.01 / Ольга Петрівна Куш; Дніпропетр. нац. ун-т. — Дніпропетровськ, 2002. — 16 с.

8. Личук М.І. Ступені фразеологізації речень: Автореф. дис... канд. філол. наук: 10.02.01 / Марія

Іванівна Личук / НАН України; Інститут української мови. — К., 2001. — 16 с.

9. Личук М.І., Шинкарук В.Д. Ступені фразеологізації речень / Марія Іванівна Личук, Василь Дмитрович Шинкарук. — Чернівці: Рута, 2001. — 136 с.

10. Паславська А.Й. Заперечення як мовна універсалія: принципи, параметри, функціонування / Алла Йосипівна Паславська / Львівський національний ун-т ім. Івана Франка, 2005. — 289 с.

11. Русская грамматика : В 2-х т. — Т. 2. Синтаксис / Под. ред. Н. Ю. Шведовой. — М.: Наука, 1980. — 709 с.

12. Шмелёв Д.Н. Синтаксическая членность высказывания в современном русском языке / Дмитрий Николаевич Шмелёв. — Изд. 2-е, стереотипное. — М.: КомКнига, 2006. — 152 с.

13. Шведова Н.Ю. Очерки по синтаксису русской разговорной речи / Наталия Юльевна Шведова. — М.: Изд-во АН СССР, 1960. — 377 с.

14. Ситар Ганна. Статус синтаксичних фразеологізмів у системі фразеологічних одиниць / Ганна Ситар // Вісник Донецького національного університету. Серія Б. Гуманітарні науки. — Донецьк: ДонНУ, 2011. — № 2. — С. 66—74.

15. Ситар Ганна. Статистичні критерії аналізу синтаксичних фразеологізмів / Ганна Ситар // Вісник Донецького національного університету. Серія Б. Гуманітарні науки. — Вінниця: ДонНУ, 2015. — № 1—2. — С. 245—256.

А. СИТАРЬ

АЛГОРИТМ РАСПОЗНАВАНИЯ СИНТАКСИЧЕСКИХ ФРАЗЕОЛОГИЗМОВ (ДЛЯ ПОСЛЕДОВАТЕЛЬНОСТИ ЧЕТЫРЁХ СЛОВ)

В статье предпринята попытка построения алгоритма распознавания синтаксических фразеологизмов для четырёх слов, расположенных подряд в тексте. Опираясь на результаты предшествующих теоретических исследований, выделены формальные и статистические параметры, позволяющие разграничить синтаксические фразеологизмы и другие языковые единицы в корпусе украиноязычных текстов.

Ключевые слова: алгоритм; корпус текстов; модель предложения; синтаксический фразеологизм; украинский язык; фразеологизированное предложение.

H. SYTAR

ALGORITHM OF SYNTACTIC IDIOMS RECOGNITION IN THE CORPUS OF TEXTS (FOR FOUR WORDS IN THE SEQUENCE)

In the article it is made an attempt to construct the syntactic idioms recognition algorithm for four words text sequence. Based on the results of previous theoret-

ical studies we highlighted formal and statistical criteria that enable to distinguish syntactic idioms from other language units in the corpus of Ukrainian-language texts.

Keywords: algorithm; corpus of texts; model of sentence; syntactic idiom; sentence with phraseological structure; the Ukrainian language.

H. SYTAR
ALGORITHM OF SYNTACTIC IDIOMS RECOGNITION IN THE CORPUS OF TEXTS (FOR FOUR WORDS IN THE SEQUENCE)

Background: Attention of national and foreign researchers was focused so far on structural and semantic features of syntactic idioms. Automatic analysis of these peculiar units that are on the verge of syntax and phraseology still was not carried out in the scientific literature. This issue requires a theoretical understanding and practical implementation. In the construction of algorithm for four words in the sequence in the text, we lean against previous experience from the conclusion of algorithm for three words in the sequence that embrace for the sentences with one-term and sentences with two-term core component [Ситар 2016].

16. Ситар Ганна. Алгоритм розпізнавання синтаксичних фразеологізмів у корпусі текстів: спроба створення / Ганна Ситар // Лінгвістичні студії / Linguistic Studies : зб. наук. праць / Донецький нац. ун-т ім. Василя Стуса; гол. ред. А. П. Загнітко. — Вінниця: ДонНУ ім. Василя Стуса, 2016. — Вип. 32. — С. 100—108.

17. Ситарь Анна. Модели синтаксических фразеологизмов с повторами в украинском языке: попытка классификации / Анна Ситарь // Язык, сознание, коммуникация: Сб. статей / Редкол. М.Л. Ремнёва, Е.Л. Бархударова, А.И. Изотов, В.В. Красных, Ф.И. Панков. — М.: МАКС Пресс, 2013. — Вып. 47. — С. 485—493.

18. Шабат-Савка С.Т. Категорія комунікативної інтенції в українській мові: [монографія] / Світлана Тарасівна Шабат-Савка. — Чернівці: «Букрек», 2014. — 412 с.

19. Ягунова Елена, Пивоварова Лидия. «От коллокаций к конструкциям» / Елена Ягунова, Лидия Пивоварова // ACTA LINGUISTICA PETROPOLITANA. Труды Института лингвистических исследований РАН. — Т. X. — Ч. 2. Русский язык: грамматика конструкций и лексико-семантические подходы / С.С. Сай, М.А. Овсянникова, С.А. Оскольская (ред.). — СПб.: Наука, 2014. — С. 568—617.

20. Church Kenneth Ward, Hanks Patrick. Word Association Norms, Mutual Information, and Lexicography / Kenneth Ward Church, Patrick Hanks // Computational Linguistics. —1990. — 16(1). — P. 22—29.

21. Ситар Ганна. Алгоритм розпізнавання синтаксичних фразеологізмів у корпусі текстів: спроба створення / Ганна Ситар // Лінгвістичні студії / Linguistic Studies : зб. наук. праць / Донецький нац. ун-т ім. Василя Стуса; гол. ред. А. П. Загнітко. — Вінниця: ДонНУ ім. Василя Стуса, 2016. — Вип. 32. — С. 100—108.

22. Ситарь Анна. Модели синтаксических фразеологизмов с повторами в украинском языке: попытка классификации / Анна Ситарь // Язык, сознание, коммуникация: Сб. статей / Редкол. М.Л. Ремнёва, Е.Л. Бархударова, А.И. Изотов, В.В. Красных, Ф.И. Панков. — М.: МАКС Пресс, 2013. — Вып. 47. — С. 485—493.

23. Шабат-Савка С.Т. Категорія комунікативної інтенції в українській мові: [монографія] / Світлана Тарасівна Шабат-Савка. — Чернівці: «Букрек», 2014. — 412 с.

Purpose: To create an algorithm of recognition of syntactic idioms in the corpus of texts (for four words in the sequence in the text).

Results: Based on the results of previous theoretical studies we highlighted formal and statistical criteria that enable to distinguish syntactic idioms from other language units in the corpus of Ukrainian-language texts.

The author developed a block diagram of syntactic idioms recognition, incorporating two branches constructed accordingly. The first branch is based on the presence of word repeats (full words concurrence or presence of other word forms of the word) and the list of core components determined on previous stages of the study (*ε ... i/ü/ta;з/із/зі ucix/vcix, ne*).

The second branch was created for another type of syntactic idioms — with a three-term core component. It takes into account the following properties of the analyzed units: the presence of combinations of service parts of speech, service parts of speech with pronoun or adverb, pronoun and adverb; compliance of words combinations with the register of the syntactic idioms core components currently comprising 108 structures; association measure of mutual information ≥ 19.8 .

Discussion: Offered algorithm enables automatic identification of syntactic idioms and removal of contexts of their use, it can be used to improve the procedure of automatic text processing and creation of automated translation systems.

Стаття надійшла до редакції 06.04.2017

УДК:371.134:004

ХАРИЦЬКА Світлана

РОБОТА ЗІ СЛОВНИКОВОЮ СТАТТЕЮ ЯК ПОЧАТКОВИЙ ЕТАП ФОРМУВАННЯ НАВИЧОК АНОТУВАННЯ ТА РЕФЕРУВАННЯ ТЕХНІЧНИХ ТЕКСТІВ

У статті автор вказує на необхідність формування навичок реферування та анотування для отримання необхідної інформації з іншомовних науково-технічних джерел через роботу зі словниковими статтями. Вказується на необхідність вміти працювати зі словниково-довідниковою літературою. Перелічуються причини, що спонукають майбутніх спеціалістів технічної галузі вміти оперативно вилучати необхідну інформацію, роблячи анотації та реферати науково-технічної літератури.

Автор зазначає, що головним завданням сучасної освіти є підготовка людини до морального та професійного самовизначення в ринковій системі відносин, до самостійної взаємодії з динамічним світом.

Ключові слова: словник; словникова стаття; анотація; анотування; реферат; реферування; наукова-технічна інформація.

Постановка проблеми. На сучасному етапі розвитку суспільства неперервного навчання вміння працювати з інформацією стає передумовою успіху в усіх галузях життєдіяльності людини. В умовах зміни завдань сучасної освіти, які визначаються варіативністю навчально-виховних закладів, модернізацією змісту, розширенням інноваційних процесів, виникає необхідність по-новому розглянути специфіку педагогічної діяльності і, зокрема, фахову підготовку сучасного спеціаліста.

Слід зазначити, що суспільство інформаційних технологій значно більшою мірою зацікавлене в особистостях, здатних самостійно й активно діяти, приймати рішення, гнучко пристосовуватися до змінних умов життя. Тому, на наш погляд, головним завданням сучасної освіти є підготовка людини до морального та

професійного самовизначення в ринковій системі відносин, до самостійної взаємодії з динамічним світом. На освіту нині суттєво впливають зміни у сучасному суспільстві. На даний час не тільки педагогічна спільнота, але й суспільство загалом розуміє, що володіння інформаційними технологіями є одним із базових елементів сучасної освіти. Зокрема, протягом останніх десятиліть у багатьох країнах світу та й в Україні змінилась суспільна парадигма — від технократичної до індустріальної, від індустріальної до інформаційної [6, с. 5]. На сучасному етапі розвитку суспільства інформація набуває ознак стратегічного продукту [7].

Протягом останнього десятиліття розвинулі країни Європи та світу, серед яких Австрія, Велика Британія, Канада, Нова Зеландія,