

УДК 811.161.2'367.3'42'06-116:004]:82-3Р.

DOI <https://doi.org/10.32447/2663-340X-2018-4-70-74>

СТРУКТУРНЕ РОЗЗНАЧЕННЯ РЕЧЕНЬ З ЧУЖИМ МОВЛЕННЯМ У КОРПУСІ ТЕКСТІВ Р. ІВАНИЧУКА

Лотоцька Н. Я.

викладач кафедри іноземних мов та перекладознавства
Львівського державного університету безпеки життєдіяльності
вул. Клепарівська, 35, Львів, Україна

У статті розглянуто розвиток корпусних досліджень у лінгвістиці та можливості комп'ютерних технологій у дослідженні та описі мови. Зазначено, що із розвитком корпусної лінгвістики значно зросла ефективність лінгвістичної обробки великих масивів текстів, зокрема авторських, оскільки корпус надає великі можливості для проведення різноаспектних лінгвістичних досліджень мовної системи.

Подано визначення корпусу текстів (КТ) та його значення для проведення мовознавчих студій. Зауважено, що вивчення мовних явищ із використанням корпусу КТ дозволяє не лише підтвердити чи спростувати окремі аспекти теорії, а й уможливує дослідження мови письменника, його стиль, а також – виявлення закономірностей функціонування мовних одиниць на різних рівнях.

Метою дослідження є простежити та визначити типи синтаксичних конструкцій, завдання полягає у розробленні системи позначок для роззначення чужого мовлення в письменницькому корпусі.

У доробку окреслено найтипівіші способи передачі чужого мовлення в авторському тексті, яке, своєю чергою, дає змогу відобразити особливості мови персонажа. Чуже мовлення у художньому тексті може бути передано такими конструкціями, як пряма і непряма мова, невласне пряма мова, монолог, діалог, полілог, цитата.

Особливу увагу приділено структурному роззначенню чужого мовлення в корпусі текстів (КТ) Р. Іванчука, для цього запропоновано принципи роззначення чужого мовлення. Зазначено, що роззначення надає особливу додаткову інформацію про властивості текстів, які входять до корпусу, дозволяє швидко та ефективно знайти ті слова, форми та конструкції, які необхідні досліднику.

В результаті проведеного дослідження зазначено, що творчий доробок письменника містить величезний арсенал різних синтаксичних конструкцій передачі чужого мовлення, а отримані результати відкриють перспективи для вивчення синтаксичних конструкцій з прямою, непрямою, невласне прямою мовою та цитатами, дослідження чисельності та частоти цих синтаксичних конструкцій у корпусі текстів письменника.

Ключові слова: корпусна лінгвістика, корпус текстів, роззначення, структурне роззначення, пряма мова, діалог, невласне пряма мова, цитата.

Постановка проблеми в загальному вигляді та обґрунтування її актуальності. Сучасному суспільству характерне широке впровадження інформаційних технологій в усі напрямки мовознавчих досліджень [6, 30]. Відповідно сучасне мовознавство зорієнтоване на розвиток комп'ютерно-лінгвістичної методології дослідження й опису мови. Практичні потреби й нерозв'язані проблеми дослідження мови зумовили появу корпусу текстів.

Аналіз останніх досліджень і публікацій. Нині спостерігаємо стрімкий розвиток корпусних студій у всьому світі, що підтверджує як величезна кількість досліджень у цій галузі, так і різноманіття корпусів, що постійно з'являються. До напрямку залучають вивчення загальної теорії корпусної лінгвістики (Д. Байбер, Дж. Синклер, В. Тойберт), кореляції корпусної лінгвістики та інших лінгвістичних дисциплін (М. Вюн, В. Товберт), типології корпусів і методики інтерпретації корпусних даних (С. Гріс), розроблення

загальних засад створення природних мов тощо (праці Б. Алтенберга, М. Баньки, В. Френсиса, Г. Кеннеді, Г. Ліча, А. Баранова, М. Михайлова, Р. Рикова, Л. Ричкової, С. Шарова та ін.).

Дослідження основних теоретичних питань типології структур із чужим мовленням можна знайти в працях із граматики, зокрема в розділах із дослідження синтаксису [9]. У монографії Г. Чумакова [11] крім висвітлення питання типології всебічно проаналізовані компоненти структур із чужим мовленням, їх типи, різновиди й модифікації, розглянуто стилістичні функції чужого мовлення. Чимала кількість наукових студій присвячена вивченню діалогічного мовлення як одного зі способів передачі чужого мовлення. Так, Н. Шульжук з'ясовує питання функціонування різноманітних структур діалогічного мовлення [13, 10].

Формулювання мети і завдань статті. Мета статті – дослідити способи передачі чужого мовлення у корпусі текстів Р. Іванчука за допомогою структурного роззначення. Завдання статті полягає в розробленні системи позначок для виокремлення чужого мовлення у КТ письменника.

Виклад основного матеріалу дослідження.

Корпусна лінгвістика як галузь прикладного мовознавства покликана не моделювати функціонування мови в різних умовах, ситуаціях, проблемних галузях та послуговуватися цими моделями, а лише фіксувати всі аспекти функціонування мови, зберігаючи як інтра-, так і екстралінгвістичну специфіку, забезпечуючи оптимальну адекватність лінгвальних даних [2, 14].

Центральним для корпусної лінгвістики є поняття *корпус текстів*. Корпус текстів (КТ) – це збірка текстів, подана в електронному форматі, споряджена лінгвістично компетентною інформацією у придатній для опрацювання комп'ютером формі та за необхідності програмним зняттям, яке спрощує доступ до цієї інформації [5].

У світовій лінгвістиці поряд із національними корпусами все частіше з'являються корпуси текстів конкретних письменників. Вивчення ідіолекту автора стало одним із провідних напрямків сучасного мовознавства. КТ конкретного автора дає змогу вивчати мову письменника, його стиль на орфографічному, морфологічному, лексичному, логіко-семантичному, фразеологічному, синтаксичному, прагматичному, дискурсивному та інших рівнях.

Створення корпусу творів Р. Іванчука, на нашу думку, стане важливим поступом у розробці корпусу української мови, а також цікавим для лінгвістів та інших дослідників творчості цього письменника. Щоб створити КТ письменника, необхідно перетворити текст в електронну форму та формалізувати його – зробити роззначення тексту [7].

Роззначення – головна характеристика корпусу, що відрізняє корпус від простих колекцій текстів, оскільки надає особливу додаткову інформацію про властивості текстів, які входять до корпусу. Добре розмічений текст – безцінна знахідка для спеціаліста, адже у своїй дослідницькій роботі лінгвісти залежать перш за все від кількості і якості зібраного матеріалу.

Роззначення полягає у приписуванні текстам і їх компонентам спеціальних позначок: *зовнішніх*, або *екстралінгвістичних* (відомості про автора й відомості про текст); *структурних* (розділ, абзац, речення, словоформа); *власне лінгвістичних*, що описують лексичні, граматичні та інші характеристики елементів тексту [3, 72].

Міжнародний стандарт кодування текстової інформації ТЕІ (англ. Text Encoding Initiative – проект кодування текстів), який «репрезентує ті ознаки тексту, які потребують експліцитної ідентифікації з метою сприяти текстовому опрацюванню за допомогою комп'ютерних програм» [14].

Відомо, що структурне роззначення тексту відображає інформацію про структуру тексту, дає змогу відокремити одне слово від іншого, виділити розділові знаки, окремі слова, межі словосполучення, речення, абзацу, глави, розділу, заголовка та ін.

У цьому дослідженні ми зупинимося детальніше на способах передачі чужого мовлення та його роззначення у художньому тексті. Для цього ми розробили наступні позначки, які подані за ієрархією:

<s>...</s> – початок і кінець речення;

<q>...</q> – початок і кінець мови персонажів;

<poetry>...</poetry> – початок і кінець поезія.

При роззначенні треба пам'ятати, що жодні позначки не повинні перехрещуватись, а бути лише вкладені одна в одну – <tag1> ... <tag2> ... <tag3> ... </tag3> ... </tag3> ... </tag2> ... </tag2> ... </tag1>.

Мову автора не позначаємо, а мову персонажів обрамлюємо позначками <q>...</q>. При цьому потрібно враховувати, що мови автора чи персонажа можуть охоплювати (кожна зокрема) і декілька речень чи абзаців, а можуть перемежовуватися в одному реченні. У будь-якому випадку потрібно дотримуватися правила вкладення позначок.

Тепер необхідно визначити, які є шляхи передачі чужого мовлення та структурно їх роззначити. Відомо, що в художньому тексті чуже мовлення може бути передано прямою і непрямою мовами, невласне прямою мовою, монологом, діалогом, полілогом, цитатою.

Твори Р. Іванчука презентують різноманітні способи передачі чужого мовлення. Найтипівішим засобом передачі чужого мовлення у художніх текстах є пряма мова, оскільки вона дає змогу авторові використовувати мову як засіб характеристики персонажів, передати їхній внутрішній світ, світогляд [9, 95].

<s>Бувало, що хтось із простих карликів підіймався і вигукував у бік вибраних: <q>«Ось той щойно загадково посміхався, я не знаю, що він мав на думці, але посміхався не випадково, мені здається, що він прихований ворог карлицького ладу». </q></s> [Р. Іванчук «Орда»].

<s><q> – Не треба, добре? </q> – прошепотіла вона, коли моя рука почала супроти волі оповивати її шию. </s><s><q> – Говоріть ... </q></s><s><q> Розкажіть якийсь парадокс, хоч би про Франка. </q></s> <s><q> Був же він при всій своїй величі звичайною грішною людиною – навіщо його весь час ставити на п'єдестал? </q></s> [Р. Іванчук «Шрами на скалі»].

Часто в художніх творах чуже мовлення передається у формі діалогу. Діалогічне мовлення – найяскравіша форма передачі розмовного мовлення, оскільки саме тут повідомлення оформлюється в безперервне взаємне спілкування людей. Характерною ознакою діалогу є речення – репліки, що змінюють одна одну [12, 366]. Звісно, що творчість Р. Іванчука не є винятком:

<s>Шенгер підвівся і спитав Яновського:

<q>– А де працюватиме Умкашівлі?</q></s>

<s><q>– О, не турбуйтеся, Роберте Карловичу, </q>– відказав надміру ласкаво Яновський, <q> – ми не дамо йому пропасти!</q></s>.

<s>І ось весь інцидент вичерпався: Шенгер звів національне лихо до клопоту про одну особу і так цим догодив кураторові, що той, розчулившись, промовив:

<q>– Я завжди вірив у благонадійність професури вашої гімназії.</q></s> <s><q>Спасибі!</q></s> [Р. Іваничук «Четвертий вимір»].

Окрім діалогічного мовлення у романах письменника можна простежити вживання монологів та полілогів. В монологі мова ведеться лише від однієї особи, слів автора немає, тому позначки <q>...</q> ми не вживаємо, на відміну від полілога:

<s>Що ж повинен вчинити я – маю визнати?</s><s> А де моє поле: там, в Оріяні, де посіяв зерна Христової науки, чи у спаленій рідній Юдеї?</s><s> З ким мені бути: з оріяньським чужим народом, осяяним новою вірою, чи із своїм, юдейським, який з прадавньою ідеєю вигнання?</s><s> Який рішенець має набрати форму слова й лягти в перших рядках мого Євангелія?</s> [Р. Іваничук «Євангеліє від Томи»].

<s><q>– Водохреще, кажете?</q> – перепитав офіцер.</s> <s><q> – А як називається ця справа?</q></s>

<s><q>– Кутя,</q> – майже приязно мовила господиня.</s>

<s><q>– Кутя...</q></s><s><q> А-а, кутя!</q>– вигукнув лейтенант і враз насторожився: щось шеберхнуло за вікном.</s>

<s><q>– Ні, то не Василько,</q>– полегшено зітхнув Іван.</s><s><q> – То, мабуть, обірвався сопел з даху.</q></s>

<s><q>– І де він забарився?</q> – видихнула Марія.</s>

<s><q>– Ще в банду піде з моїм револьвером,</q> – занетерпелився офіцер.</s> [Р. Іваничук «Вогненні стовпи»].

Ще одним шляхом передачі чужої мови є непряма мова, за допомогою якої мовець передає висловлювання третьої особи від свого імені, причому це він робить не буквально, а тільки переказує його загальний зміст. Непряма мова відрізняється від прямої граматично, вона не може бути самостійним реченням, – це завжди підрядне речення, що стоїть після головного (слів автора) [9, 97]. У реченнях з непрямою мовою ми не використовуємо <q>...</q>:

<s>Антін хотів було ще поцікавитися, як почуває себе батько, коли тут таке недоречне й дивне його запитання...</s> [Р. Іваничук «Край битого шляху»].

Що стосується невластивості прямої мови, то вона є вільною, не зв'язаною з такими структурно-синтаксичними формами, як непряма мова, невластива пряма мова повніше й точніше передає емоції героїв не тільки за допомогою лексичних і фразеологічних засобів, а й за допомогою синтак-

сично-окличних речень, риторичних звертань [4, 113]. Синтаксичне вираження невластивості прямої мови ледве помітне, оскільки вона пропущена крізь призму свідомості героя. Невластива пряма мова поєднує в собі властивості прямої і непрямої мови, тому позначки <q>...</q> не вживаємо:

<s>Він куняв, немов підхмелений при неділі, та я помітив, як чоловічок враз сторожко витягнув шию й клепаню зсунув на потилицю, щоб краще мене чути.</s> <s>О, маю ще одного слухача, подумав я й відчув, що цей слухач мені заважає, ніби я соромлюся читати своє писання в його присутності...</s> [Р. Іваничук «Хресна проща»].

Треба також згадати цитату як різновид чужого мовлення в авторському тексті. Цитата – це особливий різновид чужого мовлення, дослівний уривок із якогось твору, що наводиться іншим автором для підтвердження чи пояснення своєї думки [9, 98]. У КТ ми її позначатимемо <q>...</q>:

<s>...Мені важко було повірити, коли в одному номері місячника я прочитав повідомлення: <q>«У збірнику київських футуристів «Семафор у майбутнє» опублікував свої новели «На півдорозі» і «Бій» молодий прозаїк і драматург, виходець із Західної України Мирослав Ірчан»</q></s> [Р. Іваничук «Через перевал»].

Для виокремлення поетичного тексту ми пропонуємо позначку <poetry>...</poetry>:

<s>А тоді закурилися вогнища, захмеліли голови козацькі і покотилася над чужою тісною землею широка, як дикі степи, могутня, як води на Дніпровських порогах, – пісня:

<poetry>Ой що ж бо то та за чорний ворон,

Що над морем крякає,

Ой що ж бо то та й за бурлака,

Що всіх бурлак скликає!</poetry></s> [Р. Іваничук «Мальви»].

Висновки та перспективи подальших досліджень. Досліджуючи тексти Р. Іваничука ми можемо простежити всі відомі нам способи передачі чужого мовлення. В результаті проведеного дослідження, можна впевнено стверджувати, що творчий доробок письменника містить величезний арсенал різних синтаксичних конструкцій передачі чужого мовлення – від еліптичних речень до складних синтаксичних структур.

Запропонована модель роззначення речень з чужим мовленням у корпусі текстів Р. Іваничука дасть змогу відобразити синтаксичні конструкції з прямою мовою, діалогами, полілогами, цитатами, визначити їхню частотність та чисельність у творах письменника. Отримані результати допоможуть визначити особливості ідіолекту Р. Іваничука, а описане роззначення суттєво полегшить мовознавчі та літературознавчі студії.

СПИСОК ЛІТЕРАТУРИ

1. Бук С. Структурне анотування у корпусі текстів (на прикладі прози Івана Франка). Українська мова. 2009. № 3. С. 59–71.
2. Демська-Кульчицька О. Основи національного корпусу української мови: монографія. Київ: Інститут української мови Національної академії наук України, 2005. 218 с.
3. Жуковська В. В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. 142 с.
4. Загнітко А. Український синтаксис: теоретико-прикладний аспект. Донецьк: ДонНУ, 2009. 137 с.
5. Кульчицький Ігор. Корпуси текстів як лінгвотехнологічне підґрунтя виявлення змін в українській мові. ХХ–ХХІ століття: жанрово-стильові й лінгвістичні метаморфози в українській мові та літературі: Монографія / А. Архангельська (гол. ред.), О. Левченко, О. Тищенко, та ін. Оломоуць: Університет ім. Ф. Палацького, 2016. С. 269–298.
6. Кульчицький І. М. Технологічні аспекти укладання корпусів текстів. Дані текстових корпусів у лінгвістичних дослідженнях: монографія / В. А. Широков, І. В. Шевченко, А. П. Загнітко та ін.; за ред. О. П. Левченко. Львів: Видавництво Львівської політехніки, 2015. С. 29–45.
7. Лотоцька Н. Я. Структурне роззначення речень у корпусі текстів Р. Іваничука. Актуальні проблеми філології та перекладознавства. Збірник наукових праць Хмельницького національного університету, 2017. Вип. № 12. С. 147–151
8. Сучасна українська літературна мова: Синтаксис / За заг. ред. І. К. Білодіда. К., 1972. 515 с.
9. Синтаксис української мови: навчальний комплекс / уклад. О. В. Дуденко. Умань: ВПЦ «Візаві», 2015. 208 с.
10. Широков В. А. Корпусна лінгвістика / В. А. Широков, О. В. Бугаков, Т. О. Грязнухіна та ін. К.: Довіра, 2005. 471 с.
11. Чумаков Г.М. Синтаксис конструкций с чужой речью. К., 1965. 219 с.
12. Шульжук К. Ф. Синтаксис української мови: підруч. / К. Ф. Шульжук. К.: Видавничий центр «Академія», 2004. 408 с.
13. Шульжук Н.В. Репліка як компонент діалогічної єдності. Актуальні проблеми сучасної філології. Мовознавчі студії. Зб. наук. праць. – Вип.11. Рівне., 2003. С. 36.
14. McEnery T., Xiao R., Tono Y. Corpus-based Language Studies: an Advanced Resource Book. London: Routledge, 2006.
15. Neveu F. Dictionnaire des sciences du langage [Електронний ресурс]. Paris: A.Colin, 2004. Режим доступу до ресурсу: <http://www.franckneveu.fr/mediapool/76/768102/data/IDIOLECTE.pdf>

REFERENCES

1. Buk S. (1970). Strukturne anotuvannia u korpusi tekstiv (na prykladi prozy Ivana Franka): № 3. [Structure annotation in text corpus (on the material of Ivan Franko's prose)] Ukrainska mova [in Ukrainian].
2. Demska-Kulchytska O. (2005). Osnovy natsionalnoho korpusu ukrainskoi movy: monograph. [Fundamentals of the National Corps of the Ukrainian Language]. Kyiv: Instytut ukrainskoi movy Natsionalnoi akademii nauk Ukrainy [in Ukrainian].
3. Zhukovska V. V. (2013). Vstup do korpusnoi linhvistyky: navchalnyi posibnyk [Introduction to corpus linguistics: manual]. Zhytomyr: Vyd-vo ZhDU im. I. Franka [in Ukrainian].
4. Zahnitko A. (2009). Ukrainyski syntaksys: teoretyko-prykladnyi aspekt [Ukrainian syntax: theoretical and applied aspect]. Donetsk: DonNU [in Ukrainian].
5. Kulchytskyi Ihor. (2016). Korpusy tekstiv yak linhvotekhnolohichne pidgruntia vyavlennia zmin v ukrainskii movi: monograph. [Texts corpus as a linguistic basis for the identification of changes in Ukrainian language]. Olomouts: Universytet im. F. Palatskoho [in Ukrainian].
6. Kulchytskyi I. M. (2015). Tekhnolohichni aspekty ukladannia korpusiv tekstiv [Technological aspects of creating of text corpus]. Lviv: Vydavnytstvo Lvivskoi politekhniki [in Ukrainian].
7. Lototska N. Ya. (2017). Strukturne rozznachennia rechen u korpusi tekstiv R. Ivanychuka: v. № 12 [Structural Annotation of Sentences in the R. Ivanychuk's Text Corpus]. Khmelnytskyi natsionalnyi universytet [in Ukrainian].
8. Bilodid I.K. (1972). Suchasna ukrainska literaturna mova: Syntaksys [The Modern Ukrainian language: Syntax]. K. [in Ukrainian].
9. Dudenko O.V. (2015). Syntaksys ukrainskoi movy: navchalnyi kompleks [The Ukrainian Syntax: manual]. Uman: VPTs «Vizavi» [in Ukrainian].
10. Shyrokov V. A. (2005). Korpusna linhvistyka [Corpus linguistics]. K.: Dovira [in Ukrainian].
11. Chumakov H.M. (1965). Syntaksys konstruktsyi s chuzhoi rechiu [The syntax of sentences with direct speech]. K. [in Russian].
12. Shulzhuk K. F. (2004) Syntaksys ukrainskoi movy: pidruch. [The Ukrainian Syntax: manual]. K.: Vydavnychiy tsentr «Akademiia» [in Ukrainian].
13. Shulzhuk N.V. (2003). Replika yak komponent dialohichnoi yednosti: v.11 [Replica as a component of dialogic unity]. Rivne [in Ukrainian].
14. McEnery T., Xiao R., Tono Y. Corpus-based Language Studies: an Advanced Resource Book / T. McEnery, R. Xiao, Y. Tono.– London: Routledge, 2006.
15. Neveu F. Dictionnaire des sciences du langage [Elektronnyi resurs] / F. Neveu // Paris: A.Colin, 2004. Rezhym dostupu do resursu: <http://www.franckneveu.fr/mediapool/76/768102/data/IDIOLECTE.pdf>

STRUCTURAL ANNOTATION OF SENTENCES WITH DIRECT AND INDIRECT SPEECH IN THE R. IVANYCHUK'S TEXT CORPUS

Lototska N. Ya.

*Lviv State University of Life Safety
Kleparivska Str., 35, Lviv, Ukraine*

The object of the research is R. Ivanychuk's text corpus. The study of writer's idiolect has become one of the leading areas of modern linguistics, it might help to find out quantitative and qualitative particularities of author's idiolect.

The research deals with modern techniques and their value for studying writer's language in linguistics. The methods of corpus linguistics concern such lingual phenomena in corpus texts that are difficult to explore without computer technologies.

The main aim of the article lies in determining the types of sentences via structural annotation that gives possibility to study the author's idiolect.

The objective of the research is to work out the principles of structural annotation of sentences in the writer's text corpus.

The paper presents the role of corpus linguistics in modern linguistics, the importance of computer technologies opportunities in research and description of language. The definition of corpus text and its importance in linguistic research are given.

The article deals with the development of linguistics studies, the appearance of writer's corpus text beside the national corpus text. The notion of text annotation and its types are defined, the principles of structural annotation are described.

The concept of sentence as a multifaceted unit in modern linguistics and syntactic sentence status with direct speech are analysed. Special attention is devoted to the status of sentence with direct speech.

Particular attention is concerned with marking-up of sentences in the text, such annotation helps to reveal different types of sentences in R. Ivanychuk's corpus text.

The paper presents perspective directions of further research to study quantity, frequency, length of syntactic constructions by applying structure annotation.

Key words: *corpus linguistics, text corpus, annotation, structure annotation, idiolect, direct speech, indirect speech, dialogue.*