

*ДИБІНА А. В.*

*Харківський гуманітарний університет “Народна українська академія”*

## ОГЛЯД ПІДХОДІВ ДО ДОСЛІДЖЕННЯ СТРУКТУРИ ЗВ’ЯЗНОГО ТЕКСТУ

У статті проведено огляд головних понять лінгвістики тексту – “текст” та “зв’язний текст”. Звертається особлива увага на підходи до вивчення структури зв’язного тексту.

**Ключові слова:** лінгвістика тексту, зв’язний текст, автоматичне реферування.

В статье проведено исследование главных понятий лингвистики текста – “текст” и “связный текст”. Обращается особое внимание на подходы к изучению структуры связного текста.

**Ключевые слова:** лингвистика текста, связный текст, автоматическое реферирование.

The article deals with the study of main concepts of the text’s linguistics – “text” and “coherent text”, focuses on approaches to studying the structure of coherent text.

**Key words:** text’s linguistics, coherent text, automatic summarization.

**Актуальність** проблеми дослідження тексту, який би сприймався як смислове та структурне ціле, обумовлена тим, що комплексне дослідження зв’язного тексту має першочергове значення для розробки систем автоматичної обробки текстів природною мовою.

**Метою** даної статті є огляд досліджень в області лінгвістики тексту.

**Об’єктом** дослідження є існуючі підходи до вивчення структури зв’язного тексту, **предметом** – поняття “тексту” та “зв’язного тексту”.

Дослідженням зв’язного тексту займається лінгвістика тексту, що сформувалася в 60–70-х роках ХХ століття. Виникнення даного напрямку філологічних досліджень є результатом всього періоду розвитку лінгвістичної науки, оскільки достатня вивченість речення поза контекстом показала, що самі по собі окремі речення не є закінченими та незалежними одиницями мовленнєвого твору, та не представляють можливості теоретичного опису мовних явищ. З’явилися багаточисленні публікації, монографії, навчальні посібники, почали проводитися наукові конференції та симпозіуми, присвячені лінгвістичним дослідженням тексту. Передумовою появи нової області лінгвістичного знання, яку Л. Гобін назвав “від речення до тексту” [1:53], є те, що вчені в центр уваги поставили текст як вищу комунікативну одиницю, структура якої відрізняється від структури окремого речення [2, 3], а також правила побудови зв’язного тексту та його смислові категорії, що виражаються за цими правилами [4:63], аналіз семантики різних видів тексту, показники текстуальної організації [5:64]. Атомічна генеративно-трансформаційна теорія, в центрі

вивчення якої знаходились фонологічні, морфологічні, синтаксичні та семантичні аспекти незалежних від контексту речень, була відкинута [6].

На необхідності вивчення зв'язного тексту, а не окремого речення, наголошують багато науковців. Проте саме поняття “текст” на сьогодні є найбільш невизначеним та спірним, оскільки різні вчені по-різному визначають його та акцентують увагу на вивченні тих чи інших аспектів тексту. Текстом називають і слово, і окреме речення, і абзац. Автори західних підходів представляють текст як структурне ціле, в якому поєднані зовнішні та внутрішні текстові ознаки. І тільки в ХІХ–ХХ ст. текст почали розуміти як структуру, зв'язну та цілісну (або ціліснооформлену) єдність, де структура представляє собою складну ієрархічну систему, в якій між реченням та текстом існує ціла ланка одиниць структури тексту [7]. У витоків такого розуміння тексту стояв російський фольклорист та літературознавець Володимир Якович Пропп, який у 1928 р. в роботі “Морфологія казки” розглянув російські казки як систему з повторюваною структурою. Ця книга була перекладена англійською та французькою, вплинувши на становлення та розвиток багатьох наукових напрямків, у тому числі семіотику та лінгвістику тексту. Інший російський вчений, Олександр Матвійович Пешковський, запропонував теорію про систематичне дослідження конститутивних ознак тексту, звернувши увагу на логічну побудову мовних одиниць, більших за речення, і тим самим заклав основи вивчення поняття складного цілого. З проблемою виявлення структурних особливостей тексту без врахування семантичної складової елементів пов'язане дослідження американського вченого Зелліга Харріса (Zellig Harris), засноване на побудові граматики тексту без врахування граматики речень, що входять до його складу.

Разом з тим в науковій літературі проблема вивчення закономірностей побудови структурованого тексту неодноразово привертала увагу дослідників. Так, ще 2000 років тому в класичній поезії та риториці були розроблені структурні моделі поезії, драми, права та політики, що були неперевершеними до появи та розвитку ідей структуралізму в лінгвістиці, наприклад, в глосематиці, де текст розглядався як єдина даність. Пізніше структурні методи дослідження з'явилися в психології, антропології, етнографії, дослідженнях з теорії літератури [6].

Теоретичну основу дослідження складають класичні роботи вітчизняних та іноземних вчених, в основному французьких структуралістів та семіологів Клода Леві-Стросса (Claude Lévi-Strauss), Ролана Барта

(Roland Barthes), Цветана Тодорова (Tzvetan Todorov), Альгірдаса-Жюльєна Греймаса (Algirdas Julien Greimas) та інших дослідників.

З формуванням лінгвістичної теорії тексту пов'язані роботи західних вчених Роланда Харвега (Roland Harveg), Адріануса ван Дейка (Adrianus Van Dijk), Вольфганга Дресслера (Wolfgang Dressler), Карела Гаузенблаза (Karel Gauzenblaz), Петра Сгалла (Petr Sgall), Ірени Беллерт (Irena Bellert) та інших.

Зазначимо, що в становленні нової наукової дисципліни існувало два етапи: перший – розглядав текст з точки зору його таксономічних одиниць; другий етап розвитку переосмислив текст як комунікативний процес, що і призвело до виникнення дискурсивного аналізу. Так, один із засновників лінгвістики тексту та її найбільш активний організатор та популяризатор голандський лінгвіст Т. А. ван Дейк у другій половині 1970-х років перейшов з текстової на дискурсивну термінологію, що знаменувало перехід від статичного до динамічного, процедурного підходу. Важливу роль зіграли роботи Вольфганга Дресслера та Роберта де Богранда, а також роботи Роджера Шенка, присвячені моделюванню процесу розуміння та синтезу зв'язного тексту в рамках штучного інтелекту.

У вітчизняній лінгвістичній науці методологічні основи аналізу тексту, його структурних одиниць, проблеми розуміння в межах цілого тексту, а не окремого речення, категорії смислової та структурної зв'язності тексту розглядали І. Р. Гальперін, І. П. Севбо, Г. Г. Москальчук, С. І. Гіндін, Н. М. Леонтєва, Ю. М. Марчук, Л. В. Орлова та інші. Проблемою когнітивних та прагмастилістичних аспектів розподілу інформації в тексті в межах побудови зв'язного тексту займалася Л. Г. Лузіна. Закономірності організації науково-реферативного тексту як структури описані в роботах української вченої В. І. Перебийніс. Дослідження зв'язності тексту Е. Ф. Скороходько, І. І. Ратцевої, Г. В. Бондаренко, В. Є. Берзона та деяких інших спрямовані на розробку автоматичних методів його згортання.

Особливий внесок в дослідження питань глибинної структури, актуального членування речення був зроблений чеським лінгвістом В. Матезіусом. Ще один представник празької лінгвістичної школи Ф. Данеш вивчав динамізм тем та рем в тексті. Проблеми семантичної пам'яті були досліджені в роботах англійського вченого Р. Шенка та голандського лінгвіста Т. А. ван Дейка.

В штучному інтелекті також відбулася зміна парадигм, викликана необхідністю моделювання розуміння мовлення та розробкою програм

автоматичної обробки текстів: моделювання знань про навколишній світ, автоматичний лінгвістичний аналіз та синтез смислу текстів.

Автор першої монографії з лінгвістики тексту у вітчизняній науці Зинаїда Яківна Тураєва [8] визначає ті завдання, вирішення яких є найбільш актуальним у наш час:

1. Вивчення тексту як складного структурно-семантичного цілого, основними характеристиками якого є цілісність та зв'язність. Саме ці характеристики дозволяють організувати текст в семантичне та структурне ціле, що дає можливість говорити про відмінність тексту та “не-тексту”. Смисл зв'язного тексту не є сумою смислів кожного окремого речення, що входить до його складу. Зв'язний текст – це система, в якій всі елементи взаємопов'язані та взаємообумовлені. Ця взаємопов'язаність проявляється у повторі деяких смислових одиниць та компресії речень, які в цілому тексті можуть бути доповнені іншими реченнями. Тому, аналізуючи структуру тексту, слід враховувати міжфразові зв'язки (лексичні та граматичні), міжфразовий контекст та прагматичну спрямованість тексту.

Дія інтегративних, або ієрархічних відношень, проявляється в поєднанні одиниць різних рівнів на семантичному та структурному рівнях [8:23]. Структура тексту забезпечується завдяки двом основним ознакам – цілісності та зв'язності. Ознака цілісності запозичена з природничих наук та проявляється як властивість об'єкта в цілому. Цілісність складається з відношень між окремими елементами об'єкту, веде до смислової організації тексту. Ознака зв'язності властива математиці, проявляється через зовнішні структурні показники та веде до форми, структурної організації тексту [3].

Дія категорії зв'язності проявляється у взаємодії глибинного смислу тексту та його формального вияву [8:23]. Дія даної категорії визначає спосіб організації тексту в семантичне ціле [2:8].

2. Лінгвістичний аналіз одиниць, що утворюють текст. Щодо критеріїв виокремлення одиниць тексту, то вчені по-різному підходять до вирішення цього питання.

Так, у якості основної одиниці тексту виділяють надфразові єдності (НФЄ) (І. Р. Гальперін, А. І. Москальська, Є. А. Реферовська та інші) та складне синтаксичне ціле (ССЦ) (О. М. Пешковський, Л. А. Булаховський, М. С. Поспелов, І. А. Фігуровський, Г. Я. Солганик та інші вчені). НФЄ характеризуються однією мікротемою, єдиним смислом його речень, смисловою та комунікативною єдністю [8:115], представляють собою закінчену частину комунікації [9:69]. Між собою ці єдності поєднуються

повторами, синонімічними рядами, ключовими словами [8:11], сполучниками, видо-часовими формами, пов'язаними з розподілом інформації в тексті [8:118]. Деякі вчені (З. Я. Тураєва) виділяють більші одиниці тексту, поєднуючи кілька НФС в єдиний предикативний комплекс [9:72; 8:74]. Поняття ССЦ, що було введено в лінгвістику тексту Олександром Матвійовичем Пешковським, характеризується інтонацією та ритмом, завдяки чому представляється можливим відокремити різні відрізки мовленнєвого ланцюжку.

3. Вивчення міжфразових зв'язків та відношень, що дозволяє розробити синтаксис складних мовленнєвих структур в рамках зв'язного тексту, але потребує глибинного проникнення в семантику тексту, приховану від безпосереднього спостереження.

4. Виявлення категорій тексту, тобто спеціальних текстових понять, притаманних певному функціональному стилю.

5. Побудова класифікації текстів за комунікативними параметрами.

Огляд існуючих підходів до аналізу структури зв'язного текста показав, що кожен з них розглядає текст з різних позицій, але об'єднуються ці підходи загальними законами зв'язності та загальною спрямованістю на цілісність тексту. Все різноманіття підходів умовно можна розділити на два напрямки.

*Перший напрямок* досліджень тексту характеризується виявленням змістових та структурних компонентів тексту і пов'язується із забезпеченням правильної побудови тексту – структурно та семантично:

1. Прагмастилістичний підхід виявляє “метасмисл”, тобто глибинні смислові відношення в кожному тексті, всю ту інформацію, яка імпліцитно не виражається в тексті, але присутня, утворюючи деяке організоване ціле [10].

2. Композиційно-тематичний підхід враховує розподіл інформації в тексті, залежить від місця розташування певних одиниць тексту, що веде до розуміння тексту в цілому не як набору окремих речень, а як зв'язного цілого. В рамках цього підходу уникнути врахування функціонального стилю здається неможливим, оскільки саме стиль тексту визначає його комунікативне та смислове структурування. Зокрема, інформація в науковому тексті розподіляється таким чином, що в кінці тексту його інформативність значно знижується, тоді як в початковій частині тексту та в його середній частині можна спостерігати високий рівень інформативності [7; 11; 12]. Зазначені вище частини тексту поєднуються між собою за



допомогою повторів, без яких неможливо організувати текст в логічне ціле. Це свого роду “елементарні схеми нанизання” [12], які дозволяють впорядкувати інформацію таким чином, щоб реципієнт мав можливість співвідносити відому інформацію з новою [11].

3. Структурно-семантичний підхід вивчає формально-граматичні критерії організації тексту на гіперсинтаксичному рівні [5].

4. Комунікативно-орієнтований та комунікативний підходи з’явилися на початку 70-х років – враховують прагматичну функцію тексту, закономірності передачі та сприйняття інформації відповідно автором та читачем, що характеризує становлення лінгвістики як все більш прагматично орієнтованої [5]. В такому розумінні текст завжди розглядається як елемент комунікації, мовні структури як інструмент здійснення певних інтенцій автора, а лінгвістика тексту повинна перш за все вивчати функціональну спрямованість тексту, тобто переважно враховувати комунікативну функцію мови.

5. Граматичний підхід спрямований на вивчення граматичної структури тексту, виявленні закономірностей побудови тексту [5].

6. Структурний (структуральний) або системний підхід описує текст в системних відношеннях між одиницями тексту [5].

7. Стилiстичний підхід виявляє стилістичні засоби виразності в тексті [5] та інші підходи.

*Другий напрямок* засновано на пошуку універсальних рис текстів безвідносно до конкретної мови. Цей напрямок активно використовує індуктивний метод мислення та розвиває дослідження граматики тексту як основи розвитку лінгвістичних знань взагалі.

Дослідження зв’язного тексту провадяться також у зв’язку з прикладними завданнями автоматичної обробки тексту – у першу чергу, автоматичного реферування, тобто комп’ютерного стислого викладу первинного тексту.

У вітчизняній науці цими проблемами займалися вчені Е. Ф. Скороходько та І. П. Севбо, яка розробила метод стилістичної діагностики текстів, порівнюючи кількісні характеристики синтаксичної структури речень. Серед зарубіжних вчених маємо зазначити американського спеціаліста зі штучного інтелекту Девіда Румельхарта (David Everett Rumelhart), який створив оповідну граматику (story grammars), американських лінгвістів Уільяма Манна (William Mann)

та Сандру Томпсон (Sandra Tompson), які у 1980 р. розробили методику описання структури дискурсу у вигляді дискурсивних одиниць.

Побудові системи автоматичного реферування присвячені і наші дослідження структури тексту. Проте наш підхід відрізняється від зазначених вище тим, що першочергово була вивчена синтаксична та семантична структура реферативного тексту та побудована модель реферату. Наявність такої моделі дозволяє проводити аналіз первинного тексту з урахуванням смислових аспектів, що наявні в рефераті, що також істотно полегшує вибір в тексті необхідної для реферату інформації.

Таким чином, в роботі проаналізовано поняття “текст” та “зв’язний текст”, існуючі підходи до аналізу структури зв’язного тексту, обґрунтована необхідність проведення подібних досліджень для побудови системи автоматичного реферування.

**Перспективи дослідження.** Для розробки системи автоматичного реферування ми взяли за основу прагмастилістичний та композиційно-тематичний підходи, оскільки вони найбільш повно відповідають завданням аналізу структурно-семантичної організації тексту та реферату. Методи аналізу, що використовуються в рамках даних підходів, дозволяють виділити в первинному тексті смислові одиниці для заповнення розробленої нами моделі реферату.

## ЛІТЕРАТУРА

1. Радзиевская Т. В. Об одном из направлений современных лингвистических исследований / Татьяна Вадимовна Радзиевская // Языкознание. – 1991. – № 3. – С. 53–58.
2. Орлова Л. В. Структура сверхфразового единства в научных текстах / Лариса Васильевна Орлова. – К. : Наукова думка, 1988. – 153 с.
3. Валгина Н. С. Теория текста / Нина Сергеевна Валгина. – М. : Логос, 2003. – 279 с.
4. Левицкий Ю. А. Лингвистика текста / Юрий Анатольевич Левицкий. – М. : Высшая школа, 2006. – 208 с.
5. Стриженко А. А. Об особенностях организации текстов, относящихся к разным функциональным стилям / Адель Алексеевна Стриженко, Любовь Ивановна Кручинина. – Иркутск : Изд-во Иркут. ун-та, 1985. – 176 с.
6. Дейк ван Т. А. Стратегии понимания связного текста / Адрианус ван Дейк, Вальтер Кинч // Новое в зарубежной лингвистике. – М., 1988. – Вып. 23 : Когнитивные аспекты языка. – С. 153–211.
7. Бондаренко Г. В. Методы лингвистического анализа структуры текста на сверхфразовых уровне (на материале научно-технических текстов) : автореф. дисс. на соискание науч. степени канд. филол. наук / Г. В. Бондаренко. – М., 1976. – 20 с.
8. Тураева З. Я. Лингвистика текста / Зинаида Яковлевна Тураева. – М. : Просвещение, 1986. – 127 с.
9. Гальперин И. Р. Текст как объект лингвистического исследования / Илья Романович Гальперин. – [6-е изд.]. – М. : Либроком, 2009. – 144 с.
10. Волков В. В. Прагмастилистика и филологическая герменевтика / Валерий Вячеславович Волков // Материалы VII междунар. конф. по проблемам семантических исследований [“Текст и методика его анализа”], (Харьков, 1994). – Харьков, 1994. – Ч. I : Теоретические основы лингвистики текста. – С. 15–17.
11. Москальчук Г. Г. Структура текста как синергетический процесс / Галина Григорьевна Москальчук. – М. : Эдиториал УРССЭ, 2003. – 296 с.
12. Севбо И. П. Структура связного текста и автоматизация реферирования / Ирина Платоновна Севбо. – М. : Наука, 1969. – 135 с.