

УДК 81`33.811.92

Леута О.І.

СУЧАСНІ КОМУНІКАТИВНІ ТЕХНОЛОГІЇ ТА МОВОЗНАВСТВО

Статтю присвячено огляду проблем, що постали перед мовознавством на сучасному етапі розвитку суспільства та упровадження нових комунікативних технологій.

Ключові слова: комунікативні технології, інформація, інтернет, комп'ютерна лінгвістика, корпус мови, тезаурус.

Сучасне суспільство визначають як інформаційне. Цей термін почали вживати у 60-х роках ХХ ст. для відзначення тієї ролі, яку поступово набувала інформація в житті людства. На думку Е.Тоффлера, “Для цивілізації Третьої хвилі одним із найважливіших видів сировини, причому невичерпним, буде інформація” [Тоффлер 1984: 33]. Якщо основу індустріального суспільства складали вироблення та розподіл енергії, то в постіндустріальному провідними галузями економіки стали сфера послуг, наука, освіта. Економічною основою інформаційного суспільства стало виробництво й поширення інформації. Процеси комунікації вийшли на такий рівень, що окремі країни почали їх використовувати у військовій галузі, зокрема в інформаційних операціях чи навіть війнах. Антологію воєн можна представити у такому вигляді: війни аграрного періоду вели за території, війни індустріального періоду – за засоби виробництва, постіндустріального – за сировину, інформаційного – за засоби оброблення, зберігання й використання інформації.

Щоб побачити, якими темпами розвиваються комунікативні технології, американський дослідник Фредерік Уільямс стиснув хронологію створення засобів спілкування (36 тис.років) у межах однієї доби [Williams 1983].

00-00	36 тис.років тому	Поява homo sapiens, звукове спілкування
08 год-00 хв	14 тис.років тому	Наскельний живопис
20 год -00 хв	6 тис. років тому	Піктографія та ієрогліфія
22 год.-00 хв	5 тис. років тому	Буквено-звукове письмо
23 год. 38 хв.	1453 р.	Книгодрукування
23 год. 54 хв. 15 сек.	1832 р.	Телеграф
23 год. 55 хв. 43 сек.	1876 р.	Телефон
23 год. 55 хв. 47 сек.	1895 р.	Кіно, радіо
23 год. 57 хв. 40 сек.	1942 р.	Комп'ютер
23 год. 57 хв. 52 сек.	1947 р.	Транзистор
23 год. 58 хв. 02 сек.	1951 р.	Кольоровий телевізор
23 год. 59 хв. 01 сек.	1978 р.	DVD

Варто відзначити цікавий парадокс сучасного етапу комунікаційного розвитку людства – ”демасовизацію” масової комунікації. Якщо з 60-х років ХХ ст. завдяки телебаченню, радіо, інформаційним агенціям світ перетворився на глобальну спільноту, яка на різних континентах різними мовами обговорювала одні події, новини, переглядала одні серіали, то з 80-х років ситуація змінюється. Нестримний розвиток комунікаційних мереж, розширення можливостей вироблення різних інформаційних продуктів спричинили диверсифікаційні процеси, спрямовані на створення конкретного продукту для окремих сегментів споживача. ХХ ст. було епохою телебачення, а телевізорів за даними ЮНЕСКО у 1992 р. було більше мільярда, тобто один на чотири чоловіки. Саме в надрах телебачення й зародилися процеси децентралізації. З 1980 р. до ефірного додалося кабельне й супутникове телебачення, різко збільшилася кількість телекомпаній, каналів. Монополія державного телебачення закінчилася, приватні телекомпанії почали спеціалізуватися на виробленні інформаційного продукту для певних категорій, який важко контролювати.

ХХІ ст. є епохою Інтернету, який став інформаційним середовищем існування кожної людини. Інформатика та Інтернет нині займають таке важливе місце в житті молодшої частини суспільства, що вчені говорять про формування інтернет-комп'ютерної субкультури. У комп'ютерщиків є своя міфологія і фольклор, свій пантеон – розробники найвідоміших ігор, легендарні хакери, творці програм і апаратного продукту (С. Джобс, Б. Гейтс); своя мова, соціальна ієрархія.¹

Інтернет на очах перетворюється на комунікативно-інформаційне середовище перебування людини. У мовному аспекті інтернет-комунікація надзвичайно різноманітна й багатогранна, у ній є все, що й у загальнонародній мові. З іншого боку, Інтернет має величезний вплив на немережне спілкування в різних його формах. Найвиразніше мовна своєрідність інтернет-комунікації виявляється в таких особливостях:

1. *Посилення метамовної рефлексії у користувачів Інтернетом.* Це зумовлено тим, що у мережному спілкуванні використовують полісеміотичні (гібридні, креолізовані) знаки, а також паралельним застосуванням двох чи більше алфавітів. Полісеміотичні знаки – смайлики, що утворені комбінаціями розділових знаків, математичних і логічних символів тощо. Цікавим є фонетичне і нечислове використання назв цифр і літер (for you – 4 U; see you latter – CUL8R). Універсальним символом в адресах електронної пошти є @. Цей знак походить з англійського прикметника *at*, його називають *собака*, *абез'яна*, *жаба*, *плюшка*, *масямба*, *краказямбла*.

2. *Графіко-морфемні та графіко-лексичні гібриди.* Характерним явищем цього різновиду комунікації є змішування кирилиці та латиниці, що зумовлено двомовністю користувачів, відсутністю прямих відповідників англійським

¹ Соціальна ієрархія залежить від ступеня професіоналізму: від початківців – *чайників*, непрофесійних користувачів – *юзерів*, *лазерів* до професіоналів – *програмерів*, *сисадмінів*, *геймерів*, *хакерів*, *кракерів*.

номінаціям в українській: *створити піскпате, почали з «helpів», e-mail-адреса, IT-бізнес, on-line-конференція.*

3. *Ігрова фонетизація письма.* Тут перестають діяти правила орфографії і написання базоване на фонетичному принципі або асоціативному обігруванні: *аффтар, падонки, пазитиф, фтему, пруща.*

У результаті взаємодії спеціальної підсистеми інформатики з розмовним мовленням і молодіжним арго сформувався субстратний варіант – інтернет-комп'ютерний сленг.

Можна вважати, що третя комунікативна революція – створення персональних комп'ютерів (поряд з винаходом письма та початком книгодрукування) сформувала новий тип комунікації поряд з класичними – усним і писемним мовленням.

Постійне збільшення обсягів комунікації зумовлює “ущільнення” й водночас глибшу диференціацію семіотичного континууму: кількість класів і видів знакових систем зростає, хоча кількість природних мов невинно зменшується. Також утрачають свою роль паралінгвістичні засоби – міміка, жести, інтонація. Нові семіотики, тобто нові мови комунікації взаємодіють з традиційними мовами культур і так видозмінюються усі засоби спілкування.

З другої половини ХХ ст. відбувається зближення мовознавства з точними науками, зокрема з математикою і формується особлива галузь досліджень – математична лінгвістика, яка об'єднує математичну формальну теорію граматик і статистичну теорію мови, що використовує методи математичної статистики, теорію інформації. Мовознавство виявилось тією гуманітарною наукою, яка не лише стала використовувати інструментальні методи спостереження (експериментальна фонетика) та інструментальні прийоми (психолінгвістика), але й освоїла математичні способи моделювання мови.

У ХХІ ст. особливої гостроти набула проблема аналізу мови у її реальному функціонуванні із застосуванням досягнень інших наук: логіки, математики, психології, кібернетики. Найважливішим завданням такого семіозису наук, на думку М. Алефіренка, є “створення автоматичних систем штучного інтелекту, які моделюють знання” [Алефіренко 2005: 231]. Оскільки різноманітні знання зберігають, використовують і передають засобами мови, то моделювання науково-технічних і художніх текстів – це моделювання системи знань у відповідних галузях науки й техніки.

Наприкінці ХХ ст. сформувалася комп'ютерна лінгвістика – один із напрямів прикладного мовознавства, головним завданням якого є максимальна формалізація мовознавчих досліджень.

Завданням комп'ютерної лінгвістики спочатку було створення складних систем обслуговування комп'ютерів за допомогою мови, які роблять можливим безпосереднє спілкування людини з комп'ютером, автоматичну переробку, запам'ятовування, пошук і подачу інформації засобами мови. При цьому багато традиційних галузей мовознавства суттєво змінюють методіку дослідження завдяки можливості побудови програм, що реконструюють альтернативні варіанти фонологічних і граматичних систем прамов; машинному визначенню

часу розпаду споріднених мов за допомогою лексикостатистики; укладання машинних словників для величезних обсягів давніх писемних текстів; здійсненню допоміжних операцій для дешифрування давніх систем письма тощо.

Однією з найважливіших проблем комп'ютерної лінгвістики є моделювання природної мови. Прикладні моделі вирізняються певним спрощенням мовної реальності, хоча це не завжди означає ігнорування реальної складності модельованого об'єкта. Дослідники намагаються не обходити увагою багатоаспектність, багатолінійність, відкритість мовного механізму, але у новій дослідницькій парадигмі це не завжди можна реально враховувати й реалізовувати. Багато принципових моментів організації природних мов взагалі не вписуються в систему конкретних однозначних досліджень, вибудованих на засадах точних наук: це нечіткість, невиразність і дифузність значень мовних одиниць, динамічність мовної системи, образність номінації (насамперед метафоричність), неспівмірність обсягів словника й позначуваної ним інформації, величезні потенції в освоєнні нових знань, різноманітність функцій мови у суспільстві.

Комп'ютерне моделювання мови й мовленнєвої діяльності з кінця ХХ ст. ставить своєю метою розроблення спеціальних теорій у царині моделювання мовної здатності людини. До найважливіших напрямів цієї діяльності належать типологія штучних і природних мов, теорія розуміння тексту, теорія вербалізації комунікативних намірів, теорія комунікативних невдач, теорія семантичної еквівалентності повідомлень та ін. На думку В. Базилєва, головними проблемами, які останнім часом постали перед представниками комп'ютерної лінгвістики, є три: проблема розуміння, проблема знань і проблема комунікативних невдач [Базылев 2007: 108].

Сучасна комп'ютерна лінгвістика характерна суттєвим розширенням номенклатури лінгвістичних процесів, серед яких автоматизовані інформаційно-пошукові, інформаційно-експертні; системи машинного перекладу; автоматичні словники різних типів, зокрема термінологічні бази даних; системи автоматичної компресії тексту, логіко-семантичні й гіпертекстові системи автоматичного реферування; навчальні лінгвістичні автомати багатofункціонального та вузькоспеціалізованого типів; редакційно-видавничі системи для корекції орфографії, граматики, стилю; системи автоматичної атрибуції та адаптації текстів; текстові мовленнєві експертні системи з природним мовним інтерфейсом і розвинутими компонентами пояснення й набуття знань; автоматизовані системи розуміння й синтезу зв'язного тексту; комп'ютерні лінгвістичні ігри; системи переведення мовленнєвих повідомлень у текст і навпаки; системи автоматичної транскрипції; системи розпізнавання голосу та ін.

Головні труднощі моделювання різноманітних видів мовленнєвої діяльності пов'язані з тим, що мовленнєві утворення – тексти – моделюють реальну навколишню дійсність. Закономірно, що практично неможливо створити універсальну модель, яка буде відтворювати усі різновиди

мовленнєвої поведінки людини й аналізувати чи синтезувати різні тексти, що передають усе різноманіття навколишнього світу.

Звичайний лінгвістичний опис структури мови можна кваліфікувати як інтелектуальні структурні моделі, що можуть доволі ефективно бути застосовані у лінгводидактиці, редагуванні.

Останнім часом дослідники особливу увагу приділяють моделюванню систем мовленнєвої поведінки, орієнтованих на вузьку предметну галузь. Однак, навіть при орієнтації комп'ютерної системи на вузьку предметну ділянку ефективність автоматичної обробки великих за обсягом текстів суттєво залежить від універсальності моделі мовного механізму, тому виникає потреба побудови загальної моделі, що є основою комп'ютерної моделі мовленнєвої поведінки людини.

Крім створення *математичних моделей* мови, що є компонентом математичного забезпечення, для комп'ютерних моделей мовленнєвої поведінки, необхідно розробити *структурні лінгвістичні моделі*, які відтворюють процеси функціонування мовного механізму, а також *структурні моделі* мовного механізму.

У моделюванні окремих рівнів мовної системи є певні закономірності, зумовлені обсягом, специфікою функціонування, функціями одиниць.

Звукова форма мови є вихідною для усіх природних мов і тому основними напрямками вивчення звукової форми є такі:

- 1) аналіз акустичних характеристик мовленнєвого сигналу для створення систем передавання мовлення каналами зв'язку, для вироблення процедур автоматичного розпізнавання й синтезу мовлення, для покращення методик відновлення мовлення та слуху хворих;
- 2) аналіз артикуляційних характеристик процесів мовотворення для виявлення механізмів керування органами мовлення відповідними центрами кори головного мозку, що полегшує розуміння процесів вищої нервової діяльності людини;
- 3) дослідження процесів сприймання мовленнєвих повідомлень органами слуху для оцінки значущості акустичних ознак цих повідомлень і для побудови моделі слухової обробки мовленнєвого сигналу. Такі розробки надзвичайно важливі для спеціалістів в галузі запису й передачі мовленнєвих повідомлень, а також автоматичного розпізнавання й синтезу мовлення.

У моделюванні *синтаксису* найважливішим завданням є вироблення алгоритму розпізнавання синтаксичної структури чи окремих її елементів. Нині широко використовують три основних способи представлення синтаксичної структури речення: структури складників (дерева безпосередніх складників); структури залежностей (дерева залежностей); комплексної структури, що об'єднує елементи перших двох.

З кінця ХХ ст. кількість електронних текстів і документів, які існують лише у комп'ютерному варіанті, надзвичайно виросла і це зумовило необхідність створення технологій, що пов'язані з обробкою даних. Обробка

документів, створених на матеріалі природних мов (NLP – “*natural language processing*”) – це нові технології, які допомагають людині використовувати електронні тексти і легко знаходити та опрацювати документи.

Наприклад, така процедура, як резюмування тексту дозволяє зменшити його обсяг і зробити можливим показ навіть на невеликому екрані (планшети, мобільні телефони). Штучне спрощення і узагальнення тексту може допомогти дітям, людям похилого віку, іноземцям легше сприйняти його зміст. Для того, щоб здійснити „компресію” тексту, треба застосувати такі процедури перефразування, які б враховували довжину і стиль тексту, його “читабельність”, не порушуючи при цьому його зміст. Автоматичний перифраз – це свого роду машинний переклад інформації в рамках однієї і тієї ж мови. До того ж здатність до перефразування тісно пов’язана зі здатністю розуміти мовлення.

Психолінгвісти вважають, що якщо може бути створена автоматизована система, що здатна передавати текст іншими словами, то вона (хоча б на рівні формальної інтерпретації) уже здатна “розуміти” текст, що, у свою чергу, надзвичайно важливо для створення штучного інтелекту.

Одним з найперспективніших напрямів застосування комп’ютера в лінгвістичних дослідженнях є автоматичний пошук бібліографічної інформації. Сучасному науковцю все складніше орієнтуватися в значному обсязі наукової інформації, тому вчені активно працюють над створенням штучних мов, призначених для запису семантичної інформації і подальшого її використання в *інформаційно-пошукових системах*. Інформаційно-пошукова система – це “комплекс пов’язаних між собою частин текстів, призначених для пошуку й вияву елементів інформації, які є відповіддю на інформаційний запит, пред’явлений системі” [Селіванова 658]. Мови, що їх обслуговують, називають *інформаційно-пошуковими мовами*, а словник, розроблений для зіставлення природної та інформаційно-пошукової мови називають *інформаційно-пошуковим тезаурусом*. Прикладами інформаційно-пошукових мов є УДК (універсальна десяткова класифікація), Синтол (Ж.Гарден), Інфол (В.Рубашкін). Тезаурус містить перелік *дескрипторів*, тобто лексичних одиниць інформаційно-пошукової мови, а дескриптор відповідає групі ключових слів природної мови. Групі синонімів відповідає один дескриптор, а багатозначному слову – група дескрипторів. Процедура укладання інформаційно-пошукового тезауруса об’єднує такі операції:

- визначення тематичного обсягу тезауруса;
- збирання вихідного масиву ключових слів;
- лексикографічна обробка ключових слів;
- виявлення еквівалентності між ключовими словами;
- відбір на цій основі дескрипторів і встановлення парадигматичних відношень між ними;
- остаточне оформлення тезауруса.

Уся ця робота побудована на лінгвістичній основі: наприклад, виявлення еквівалентності проводять на основі текстів з ключовими словами. Якщо заміна

одного слова іншим в усіх контекстах не порушує змісту, то ці одиниці є еквівалентними.

Важливим магістральним напрямом сучасних лінгвістичних досліджень є *корпусна лінгвістика*, яка ґрунтується на використанні *корпуса*, тобто дуже значного за обсягом мовного матеріалу, вибраного з різних функціональних стилів і зведеного в комп'ютеризовану систему. Національний корпус – це інформаційно-довідкова система, що базована на зібранні великих обсягів текстів в електронній формі, наприклад Cobuild Bank of English нараховує близько 500 млн слововживань. Корпусна лінгвістика дає можливість залучати до аналізу величезні обсяги матеріалу і цим дозволяє уникнути неправомірних висновків та узагальнень, випадкових фактів, які можуть спотворити реальну картину функціонування мовних одиниць усіх рівнів.

Окрім того, постійне оновлення текстів сучасними текстами представляє реальний стан мови на сучасному етапі розвитку, виявляє динаміку поведінки одиниць, семантичні зрушення, граматичні процеси, чого не можуть дати традиційні словники й граматики із застиглим на певному синхронному зрізі матеріалом. Це так звані *моніторні* корпуси, що відтворюють зміни у мовній системі, наприклад, Бірмінгемський корпус Дж. Синклера.

Електронні бази даних, що дають можливість швидко опрацьовувати величезні обсяги матеріалу в автоматичному режимі відкривають нові можливості для досліджень у різних галузях мовознавчої науки, зокрема в фразеології та фразеографії. Фахівці наводять приклад успішного застосування електронних баз і методик статистичного аналізу лексичної комбінаторики, розроблені в Інституті німецької мови в Мангеймі. Спираючись та цю базу, фразеології виявляють нові значення фразеологічних одиниць, варіативність ідіом, особливості їх уживання й сполучуваності, визначають їх еквіваленти.

Як відзначають А. Молдаван і В. Плуґян, національний корпус мови – це і *база*, й *інструмент лінгвістичного дослідження*, а також *третій обов'язковий формат лінгвістичного представлення мови* (разом із традиційними грамами й словниками) [Молдаван, Плуґян 2005].

Сучасні тенденції розвитку суспільства ставлять перед дослідниками безліч нових проблем, розв'язання яких потребує знань з різних галузей і вироблення нових методик.

ЛІТЕРАТУРА

1. Алефиренко Н.Ф. Современные проблемы науки о языке: Учебное пособие. – М.: Флинта, 2005. – 412 с.
2. Базылев В.Н. Общее языкознание. – М.: Гардарики, 2007. – 285 с.
3. Молдаван А.М., Плуґян В.А. Национальный корпус русского языка // Труды Отделения историко-филологических наук РАН. – М., 2005. – С. 378-386.
4. Селіванова О.О. Сучасна лінгвістика: напрями та проблеми. – Полтава: Довкілля-К, 2008. – 711 с.
5. Тоффлер На пороге будущего// "Американская модель": с будущим в конфликте. – М.: Наука, 1984. – С. 5-39.

6. Williams F. The Communications Revolution / Relised ed. N.Y. etc., 1983. – 312 p.

Стаття посвящена обзору проблем, которые стали актуальными для языкознания на современном этапе развития общества и внедрения новых коммуникативных технологий.

Ключевые слова: коммуникативные технологии, информация, интернет, компьютерная лингвистика, корпус языка, тезаурус.

Article is devoted to review problems facing linguistics in the present stage of development of the society and the introduction of new communication technologies.

Keywords: communication technologies, information, Internet, computer linguistics, languages, thesaurus.

УДК 811. 161. 1' 374/354

Лобода Т.М.

ЗНАЧЕННЯ ЛЕКСИКОГРАФІЧНОГО ДОРОБКУ Б.Д. ГРІНЧЕНКА ДЛЯ ФОРМУВАННЯ ПРАВОПИСНОЇ СИСТЕМИ УКРАЇНСЬКОЇ МОВИ

Стаття присвячена аналізу впливу «Словаря української мови» Б. Д. Грінченка на становлення правописної системи української мови.

Ключові слова: лексикографія, словник, жанр словника, принципи добору реєстру словника, правописне оформлення словника, грінчевичівка.

Лінгвістичні аспекти творчого феномену Б.Д. Грінченка привертали увагу не одного покоління дослідників, до нього зверталися на початку 20 століття Аг. Кримський, І. Огієнко, С. Єфремов, пізніше М.М. Пилинський, О.О. Тараненко, В.В. Німчук, М.В. Кравченко, Н.Б. Копиленко та інші. Ювілейні дати видатних людей минулого спонукають нас прискіпливо переглядати сторінки життя великого ювіляра, висвітлюючи в його творчому доробку саме те, що суголосне нашій добі. У цьому сенсі 150-річчя від дня народження визначного українського письменника, перекладача, педагога, лексикографа, публіциста й громадського діяча, етнографа та фольклориста, найвизначнішого знавця і пропагандиста української мови Бориса Грінченка – не виняток, оскільки тривалий час довкола його спадщини точилися гострі суперечки, висловлювалися діаметрально протилежні думки: від буржуазно-націоналістичного діяча до буржуазно-ліберального.