

УДК 81'33(075)

*Рекомендовано до друку вченою радою філологічного факультету  
Донецького національного університету імені Василя Стуса  
(протокол №12 від 17 червня 2020 р.).*

*Рецензент:*

**Ганна Ситар**, доктор філологічних наук, доцент (Донецький національний університет імені Василя Стуса).

**Гарбера І.В.**

Основи прикладної лінгвістики: Практикум. Вінниця, 2020. 138 с.

Практикум містить цикл завдань з основ прикладної лінгвістики. Кожен структурний блок складають короткий теоретичний мінімум, контрольні питання, практичні завдання. Доповнюють видання завдання до модульної контрольної роботи та ілюстративні матеріали.

Практикум адресований студентам-філологам, які здобувають спеціальність «Прикладна лінгвістика», викладачам вузів, аспірантам, усім, хто цікавиться проблемами прикладного мовознавства.

© Гарбера І.В., 2020

## ЗМІСТ

<b>Тема 1. РОЛЬ ПРИКЛАДНОЇ ЛІНГВІСТИКИ В ІНФОРМАЦІЙНОМУ СУСПІЛЬСТВІ</b>	<b>4</b>
<b>Тема 2. ПРИКЛАДНА ФОНЕТИКА. РОБОТА АУДІОПРОГРАМОЮ AUDACITY</b>	<b>5</b>
<b>Тема 3. ТРАДИЦІЙНІ СИСТЕМИ ПИСЬМА. РОБОТА ТЕКСТОВИМИ ПРОЦЕСОРАМИ</b>	<b>9</b>
<b>Тема 4. СПЕЦІАЛЬНІ СИСТЕМИ ПИСЬМА: ТРАНСКРИБУВАННЯ, ТРАНСЛІТЕРУВАННЯ, СТЕНОГРАФУВАННЯ, МОВА ЖЕСТІВ, КРИПТОГРАФУВАННЯ, АЗБУКА МОРЗЕ</b>	<b>13</b>
<b>Тема 5. ЦИФРОВІ (КОМП'ЮТЕРНІ) СИСТЕМИ ПИСЬМА. РОЛЬ СПЕЦІАЛЬНИХ СИСТЕМ ПИСЬМА</b>	<b>17</b>
<b>Тема 6. ПРИКЛАДНА МОРФОЛОГІЯ</b>	<b>19</b>
<b>Тема 7. ПРИКЛАДНА ЛЕКСИКОГРАФІЯ</b>	<b>23</b>
<b>Тема 8. ПРИКЛАДНИЙ СИНТАКСИС</b>	<b>26</b>
<b>Тема 9. АНАЛІТИКО-СИНТЕТИЧНЕ ОПРАЦЮВАННЯ ДОКУМЕНТІВ: БІБЛІОГРАФУВАННЯ, ІНДЕКСУВАННЯ, ІНФОРМАЦІЙНИЙ ПОШУК, АНОТУВАННЯ, РЕФЕРУВАННЯ, КОРЕКТУРА, РЕДАГУВАННЯ, ПЕРЕКЛАД, ГОТУВАННЯ ОГЛЯДІВ</b>	<b>30</b>
<b><i>МОДУЛЬНА КОНТРОЛЬНА РОБОТА</i></b>	<b>34</b>
<b><i>ЛІТЕРАТУРА</i></b>	<b>39</b>
<b><i>ІЛЮСТРАТИВНІ МАТЕРІАЛИ</i></b>	<b>40</b>

# ТЕМА 1. РОЛЬ ПРИКЛАДНОЇ ЛІНГВІСТИКИ В ІНФОРМАЦІЙНОМУ СУСПІЛЬСТВІ

## I. Теоретичний мінімум

Мова – основний носій інформації в суспільстві. Людина (адресант), генеруючи повідомлення, наділяє символ певним значенням (певною кількістю інформації). У комунікативному каналі (повітрі, наприклад) символ кодується певним фізичним сигналом (звуком). Інша людина (адресат) сприймає сигнал, співставляє його зі своїми символами й, відповідно, значеннями, властивими їм (загалом, вони мають бути тотожними тим значенням, які відомі адресантові). Так виглядає механізм передачі інформації мовою.

Прикладна лінгвістика (ПЛ) і комп'ютерна лінгвістика (КЛ) конструюють лінгвістичні системи, прискорюючи опрацювання інформації. Вони використовують комп'ютери – універсальні кібернетичні пристрої опрацювання інформації (їх досліджує кібернетика).

В Україні сьогодні в програмному забезпеченні комп'ютерів вживаються переважно іноземні мови (англійська та російська). Звідси існує потреба «українізувати» операційні системи, пакети канцелярських програм, комп'ютерні лінгвістичні системи, програмне забезпечення інтернету тощо, а також зробити українську мову доступною для носіїв інших мов, тобто помістити лінгвістичні бази даних української мови в інтернеті (наприклад, БД «Словники України»).

За об'єктом вивчення (мова) ПЛ є мовознавчою дисципліною та спеціальністю, а за методами вивчення – кібернетичною. ПЛ сформувалася повноцінно в 60-70-х роках ХХ століття на основі кількох мовознавчих напрямів – експериментального, структурного, математичного, прикладного, комп'ютерного (інженерного). У 60-80-х роках ПЛ називали «Структурна, прикладна і математична лінгвістика».

## **II. Контрольні питання**

1. Дайте визначення поняття «прикладна лінгвістика». Чи можна вважати її міждисциплінарною? Чому саме?
2. Чи можна розмежувати поняття «прикладна лінгвістика» та «комп'ютерна лінгвістика»?
3. Дайте визначення поняттям «мова», «мовлення», «інформація».
4. Яким чином між собою пов'язані поняття «символ», «сигнал», «значення» та «комунікативний канал»?
5. Чи вважаєте ви комп'ютер корисним винаходом? Дайте визначення поняттю «кібернетика».
6. Чому, на вашу думку, в Україні сьогодні в програмному забезпеченні комп'ютерів вживаються переважно іноземні мови (англійська та російська)?
7. Які, на вашу думку, перспективи розвитку прикладної та комп'ютерної лінгвістики?

## **III. Практичне завдання**

Перегляньте відео «Чи може комп'ютер мислити як людина» <https://www.youtube.com/watch?v=tbS6pr13gHc>. На вашу думку, в далекому майбутньому комп'ютер навчиться спілкуватися повноцінно, як справжня людина?

## **ТЕМА 2. ПРИКЛАДНА ФОНЕТИКА. РОБОТА З АУДІОПРОГРАМОЮ AUDACITY**

### **I. Теоретичний мінімум**

Мовний звук – це звук лише з певними акустичними характеристиками, які утворюють його значення серед інших лінгвістичних звуків. Мовні звуки не позначають фрагментів світу.

Усне мовлення щодо писемного є первинним й набагато швидшим (див. Схему 1).

<b>Різновид мовленнєвої діяльності</b>	<b>Швидкість, слово / секунда</b>
усне мовлення (без пауз)	2 – 4
письмо	0,5
друкування	0,2-0,5 (непрофесійне) 1,5-2,5 (професійне)

Схема 1. Порівняльна характеристика швидкості різних видів мовленнєвої діяльності

Теоретична фонетика досліджує опозиції, у які вступають мовні звуки: голосні – приголосні, глухі – дзвінкі, губні – негубні. Прикладна фонетика досліджує мовні звуки в таких аспектах: артикуляторний (артикуляція анатомічних органів гортані людини під час утворення звуків); фізіологічний (психофізіологічні й нейролінгвістичні механізми сприйняття і розпізнавання звуків); акустичний (фізичні характеристики звуків та можливості сприйняття органів слуху).

Прикладна фонетика використовує такі методи для дослідження звуків і органів мовлення людини:

- палатографія (встановлює місця торкання язиком піднебіння під час утворення звуку);
- тензопалатографія (вимірює силу артикуляції);
- рентгенографія та кінорентгенографія (спостерігає за позицією органів мовлення та їх переміщеннями);
- осцилографія (визначає тривалість, частоту й інтенсивність звука);
- спектрографія (фіксує загальну акустичну картину звука).

Прикладна фонетика розрізняє такі звукові сигнали: первинні – безпосередньо породжені мовцем; вторинні – ті ж первинні, але пропущені через канали електроакустичних пристроїв (мікрофони, гучномовці, підсилювачі).

Прикладна фонетика у першу чергу аналізує фізичні характеристики мовних звуків. Серед них:

I. **Частота звука**, з якою згущені та розріджені ділянки повітря діють на органи слуху. Частоту звука вимірюють у кількості коливальних періодів за 1 с. Одиниця вимірювання – герц (Гц / Hz). Основний тон – найнижча частота генерування звука голосовими зв'язками. Резонанси (відбиті коливання) – інші тони, вищі за основний, що виникають у надгортанній порожнині. Мовні звуки складаються з одного основного й кількох резонансних тонів, серед яких виділяють тони з максимумом сконцентрованої в них енергії (максимальною інтенсивністю) – форманти. Людина не сприймає частоти, вищі за 20 КГц та нижчі за 20 Гц.

II. **Інтенсивність звука**, що вимірюють силою тиску, який діє на одиницю площі. Одиниця вимірювання – децибел (дБ / dB). Голосні більш інтенсивні, ніж приголосні. Для людини больовим порогом відчуття є звук, рівний чи більший 130 дБ.

III. **Тривалість звука**, вимірювана в мілісекундах, тобто одній тисячній секунди (мс / ms). Тривалість пауз у мовному потоці більше ніж 100 мс (0,1 с) дає змогу людині відокремлювати звуки, а менша – сприймати їх як один.

IV. **Спектральний склад звука**, тобто структура одного звука з кількох різних тонів – формант, від яких залежить те, з яким еталонним звуковим образом людина співвідносить отриманий органами слуху звук, тобто яким є значення отриманого звуку (чи це звук [a], чи [i]). Основний тон дорівнює 100-350 Гц. Крім нього, мовні звуки можуть містити, наприклад: [a] – одночасно приблизно 650, 1250 та 2200 Гц; [i] – одночасно приблизно 250, 2750 та 3400 Гц.

Звуковий редактор – спеціальна комп'ютерна програма, що візуально зображує звук (у дво- та тривимірному просторі). Звукова плата – пристрій, який вводить в комп'ютер звук та виводить його. Рентгенокінографування – фіксування на плівці рухів органів мовлення людини.

Серед звукових редакторів виділяють: звукові редактори загального призначення (для опрацювання музики); звукові редактори лінгвістичного призначення (для опрацювання усного мовлення); звукові редактори з відображенням і без відображення спектра звука.

Прикладна фонетика виконує такі завдання: створення таблиць звуків з їх фізичними характеристиками; розроблення методів сегментації мовного потоку; створення артикуляційних моделей усного мовлення; дослідження фізіологічних механізмів сприйняття мовних звуків органами слуху.

Комп'ютерна лінгвістика здійснює: створення комп'ютерних баз даних еталонних звуків мови (для розпізнавання усних текстів); створення спеціалізованих лінгвістичних звукових редакторів; створення систем автоматичного аналізу усного мовлення для дистанційного керування робототехнічними пристроями; створення систем автоматичного аналізу усного мовлення для його перетворення в писемний текст (створення електронних секретарів-«писарів»); створення систем автоматичного синтезування усного мовлення (створення електронних секретарів-«декламаторів»); створення систем автоматичного розпізнавання мовця (атрибуція усного тексту).

## **II. Контрольні питання**

1. Що таке кімограф? Яким чином він фіксує звук?
2. Що таке звуковий редактор? Чи здатні сучасні звукові редактори візуально зобразити звук?
3. Що таке звукова плата? Де вона знаходиться та як вона працює?
4. Які існують види звукових редакторів?
5. Що таке рентгенокінографування?

## **III. Практичні завдання**

1. Ознайомтеся зі змістом статті:  
<https://ru.wikipedia.org/wiki/%D0%90%D1%83%D0%B4%D0%B8%D0%>

BE%D1%80%D0%B5%D0%B4%D0%B0%D0%BA%D1%82%D0%BE%D1%80. Складіть план «Функції аудіоредактора».

2. Перейдіть за посиланням <https://audacity-pro.site/>. Ознайомтеся з інструкцією до використання та функціоналом аудіопрограми Audacity. Відредагуйте у скачаній програмі обраний аудіофайл (готову роботу надішліть викладачеві у двох форматах: оригінальне аудіо та відредаговане вами з детальним описом, які функції було використано).

### ТЕМА 3. ТРАДИЦІЙНІ СИСТЕМИ ПИСЬМА. РОБОТА З ТЕКСТОВИМИ ПРОЦЕСОРАМИ

#### **I. Теоретичний мінімум**

Писемна мова є похідною (вторинною) щодо усної. Письмо – графічна знакова система, призначена для фіксації усного тексту на площинному носії інформації. Графічна знакова система – сукупність (множина) графем (наприклад, абетка, множина лігатур, множина ієрогліфів). Графема – мінімальний двомірний графічний символ (рисунок, піктограма), що позначає одиницю мови.

Графема позначають:

- \* літера – фонему;
- \* літера чи лігатура (кілька літер, об'єднаних в одну, – *ш* як *ш* та *ч*) – частину або цілий склад;
- \* ієрогліф – частину слова або ціле слово.

Графіка – розділ прикладного мовознавства, який досліджує письмо.

Виокремлюють 2 види письма:

- традиційне – письмо, призначене для використання всіма членами суспільства (сербська мова використовує кирилицю і латиницю; японська – ієрогліфи, літерне письмо для слів власного походження (хірагана), літерне письмо для слів іноземного походження (катакана));



- спеціальне – письмо, призначене для використання лише якоюсь частиною суспільства (незрячі, люди з вадами органів мовлення) чи з якоюсь певною метою: зафіксувати усний текст на носії інформації; передати писемний текст за допомогою іншого алфавіту; зашифрувати писемний текст тощо.

З розвитком комп'ютерів графеми набули зафіксованих міжнародними стандартами цифрових кодів, носії інформації для запису текстів стали тривимірними.

Система письма: графіка, орфографія, напрям письма.

Графіка – сукупність усіх символів, які використовують у певний час у писемному мовленні якоїсь мови. До складу графіки української мови входять: абетка, пробіл, розділові знаки, арабські й римські цифри, апостроф, знак наголосу та ін. Абетка – це упорядкована множина графем, які позначають звуки певної мови й призначені для графічного відображення звукових образів слів. Ієрогліфічні мови замість абеток мають упорядковані списки ієрогліфів (наприклад, графіка китайської мови базується на 214 ключах, що вказують на розташування ієрогліфів у словнику, який нараховує їх багато тисяч). Алографи – різні варіанти написання літер. Наприклад, в кирилиці є 4 алографи: великі, малі, друковані й рукописні літери.

Орфографія – історично усталена й загальноприйнята система правил фіксування усних текстів певної мови на письмі.

Оптимальність абетки ( $A$ ) – співвідношення кількості літер і кількості звуків мови:  $A = L / S$  ( $L$  – кількість літер в абетці мови,  $S$  – кількість звуків у ній). Чим ближче  $A$  до 1, тим оптимальнішою є абетка.

Надлишковість графіки ( $H$ ):  $H = G / S$  ( $G$  – кількість графем у мові,  $S$  – кількість звуків у ній). Чим більша за 1 величина  $H$ , тим більша надлишковість графіки.

Для набирання текстів у комп'ютерах використовують спеціальне програмне забезпечення (текстові процесори), які дають змогу набирати й виправляти тексти. Для автоматичного набирання текстів (з готових

друкованих чи стилізовано написаних рукою оригіналів) використовують системи оптичного читання.

## **II. Контрольні питання**

1. Яке письмо називають предметним? Знайдіть його ілюстрації.
2. Опишіть різновиди графічно-символьного письма: піктографічне, ідеографічне, силабічне (складове), літерно-звукове. Наведіть приклади кожного з різновидів.
3. У яких випадках: 1) один звук може позначатися різними літерами; 2) одна літера може позначати різні звуки; 3) кілька звуків можуть позначатися однією літерою; 4) кілька літер можуть позначати один звук?
4. Які алографи має арабська абетка?
5. Які принципи ієрогліфічної системи китайського, японського, корейського письма?
6. У чому полягають фонетичний, морфологічний, традиційний, диференційний та транслітераційний орфографічні принципи? Наведіть приклади.
7. Які напрями написання мають арабське, давньогрецьке, латинське і кириличне, китайське письмо?
8. Які пристрої, інструменти й матеріали здавна використовували люди для письма?
9. Як працюють текстові процесори, призначені для опрацювання ієрогліфічних мов?

## **III. Практичні завдання**

1. Ознайомтеся зі змістом статей  
[https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2%D0%B8%D0%B9\\_%D0%BF%D1%80%D0%BE%D1%86%D0%B5%D1%81%D0%BE%D1%80](https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2%D0%B8%D0%B9_%D0%BF%D1%80%D0%BE%D1%86%D0%B5%D1%81%D0%BE%D1%80) та

[https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2%D0%B8%D0%B9\\_%D1%80%D0%B5%D0%B4%D0%B0%D0%BA%D1%82%D0%BE%D1%80](https://uk.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2%D0%B8%D0%B9_%D1%80%D0%B5%D0%B4%D0%B0%D0%BA%D1%82%D0%BE%D1%80). Чим текстові процесори відрізняються від текстових редакторів?

2. У обраному текстовому процесорі відредагуйте поданий текст, зробіть його візуально красивим, використовуючи щонайменше 5 шрифтів:

У слові відбитоморфемну структуру і морфологічні категорії конкретної мови. Граматичне значення слова, передаване службовими елементами, є абстрактним, незалежним безпосередньо від його конкретного лексичного значення. Наприклад, у слові біг («видруху») наявні такі граматичні значення, як належність до розряду іменника, дочоловічого роду однини, називного відмінка. Класифікацію типів морфологічних значень слів проводять на основі їх відношення до відображення позамовної дійсності, синтаксичної системи, словозміни, словотвору тощо. Значення, що відображають дійсність, кваліфікують як номінативні, а спрямовані у внутрішню синтаксичну структуру мови як синтаксичні (релятивні). Внутрішньоморфологічна класифікація значень ґрунтується на широкому розумінні морфології як такої системи, що складається з власне морфології і словотвірної морфології, і відповідно до характеру службових морфем (афіксів) охоплює словозмінні (власне граматичні, грамеми, флективні) та словотвірні (дериваційні, дериватеми, класифікаційні) значення.

## ТЕМА 4. СПЕЦІАЛЬНІ СИСТЕМИ ПИСЬМА: ТРАНСКРИБУВАННЯ, ТРАНСЛІТЕРУВАННЯ, СТЕНОГРАФУВАННЯ, МОВА ЖЕСТІВ, КРИПТОГРАФУВАННЯ, АЗБУКА МОРЗЕ

### I. Теоретичний мінімум

**Транскрибування** – це вид спеціального письма, за допомогою якого фіксують вимову усних текстів. Для опису звуків використовують звукове, а для опису синтагм і фраз – інтонаційне транскрибування.

У звуковому транскрибуванні виділяють:

1) *фонетичне*, за допомогою якого фіксують усі основні особливості мовлення ([коужу'х]);

2) *фонематичне*, коли вказують лише зафіксовані в мові фонемі, а не реально вимовлені в текстах (<кожу'х>);

3) *практичне*, яке застосовують для відтворення особливостей вимови чужомовних слів літерами своєї абетки (mother як *мазе*).

При інтонаційному транскрибуванні синтагми розділяють однією, фрази – двома, а кінець тексту – трьома косими лініями. Фразовий (логічний) наголос позначають знаком " перед відповідним словом.

Фонетичне транскрибування використовують в орфоепічних словниках (орфоепія – це сукупність норм вимови одиниць літературної мови, їх наголошення й інтонації у фразах усного тексту). Практичне транскрибування використовують для найпростіших типів перекладних словників-розмовників, путівників, перекладів рідною мовою іноземних власних назв, термінів, номенклатур, топонімів. Інтонаційне транскрибування широко застосовують у ЗМІ для розмічування текстів дикторам радіо й телебачення, а також для акторів театру й кіно.

**Транслітерування** – це вид спеціального письма, за допомогою якого тексти, написані однією графікою, політерно відтворюють за допомогою іншої графіки. Найчастіше графікою виступає абетка. Транслітерування за

допомогою латинської абетки називають романізацією (а, наприклад, запис японських текстів літерами латинської абетки – ромаджі).

**Стенографування** – це вид спеціального письма, за допомогою якого усний текст можна записувати в кілька разів швидше, ніж за допомогою звичайного письма. Зростання швидкості забезпечується за рахунок спрощення написання літер. Українська стенографія виникла лише в 20-30-х роках ХХ ст. Зараз її використовують у дипломатії, під час синхронного перекладу тощо.

**Мова жестів** – спеціальне письмо, яке дає змогу позначати літери, а також цілі слова жестами. Розрізняють: 1) мову жестів для людей без вад органів мовлення (наприклад, в австралійських племенах вдови після смерті чоловіків рік розмовляють лише мовою жестів); 2) мову жестів для людей з вадами органів мовлення – глухих чи глухонімих (таких людей до 1,5%). Структурно мова жестів складається з двох компонентів: тисячі-півтори слів, для позначення яких використовують окремі жести (найчастіше вживані слова), а також решти слів, для яких використовують пальцеву абетку (пальцями позначають окремі літери). Розрізняють одно- та дворучну абетки. Мови жестів не залежать від конкретних звукових мов (наприклад, мова амслен є спільною для США, Канади й Франції, але не Англії). Одиницею мови жестів є жест, або херема (від грец. *heiros* – рука). Херема може вказувати на місце виконання жесту, на конфігурацію руки, на характер руху. Кількість херем приблизно збігається з кількістю фонем. У мові жестів існує свій поділ за частинами мови (наприклад, прикметники належать до одного з різновидів предикатів, для дієслів вказується кількість об'єктів). Крім жестів, використовується міміка (наприклад, для розрізнення омонімів).

**Криптографування** (від грец. *krupto* – таємний і *grafo* – пишу, таємне писання) – це вид спеціального письма, який дає змогу записати зафіксований природною писемною мовою текст в такий спосіб, який максимально ускладнює його політерне розпізнавання. Криптографування складається з шифрування й дешифрування. Шифрування – це зміна в тексті зображення

символів, їх кодів чи порядку за наперед заданими правилами (ключем). Дешифрування – це відновлення в тексті за допомогою ключа первинного зображення символів, їх кодів чи порядку. Розрізняють дешифрування текстів, які були спеціально зашифровані (у дипломатії, розвідці), та текстів, які не були спеціально зашифровані (наприклад, написані мертвими мовами). Прикладна лінгвістика займається саме останнім різновидом дешифрування – лінгвістичним. Розпізнати текст допомагає наявність паралельних перекладів, знання власних імен та історичних подій. У результаті створюється ключ, який має містити перекладний словник і граматику мови.

*Азбука Морзе* – спосіб кодування традиційного письма, який використовували в електричних каналах зв'язку. Знаки кодують короткими («крапками») й довгими («тире») електричними сигналами. Крапка – одиниця тривалості, тривалість тире дорівнює тривалості трьох крапок. Пауза між сигналами в літері – одна крапка, пауза між літерами – три крапки.

## **II. Контрольні питання**

1. Коротко передайте зміст статті:

[https://uk.wikipedia.org/wiki/%D0%9C%D1%96%D0%B6%D0%BD%D0%B0%D1%80%D0%BE%D0%B4%D0%BD%D0%B8%D0%B9\\_%D1%84%D0%BE%D0%BD%D0%B5%D1%82%D0%B8%D1%87%D0%BD%D0%B8%D0%B9\\_%D0%B0%D0%BB%D1%84%D0%B0%D0%B2%D1%96%D1%82](https://uk.wikipedia.org/wiki/%D0%9C%D1%96%D0%B6%D0%BD%D0%B0%D1%80%D0%BE%D0%B4%D0%BD%D0%B8%D0%B9_%D1%84%D0%BE%D0%BD%D0%B5%D1%82%D0%B8%D1%87%D0%BD%D0%B8%D0%B9_%D0%B0%D0%BB%D1%84%D0%B0%D0%B2%D1%96%D1%82).

2. Назвіть основні орфоепічні норми сучасної української літературної мови.

3. Охарактеризуйте стенографічні алфавіти кількох різних мов (використовуючи знайдені в інтернеті їх ілюстрації).

4. Використовуючи мову жестів, передайте слова: *батько, мати, дідусь, бабуся, жінка, чоловік, дівчинка, хлопчик, місто, село, дім, вулиця, магазин, їдальня, школа, аптека.*

5. Знайдіть таблиці для пальцевої латинської та кириличної абетки. Дивлячись на них, передайте своє ПІБ двома мовами: англійською та українською.
6. Чи можна перетворювати мову жестів в усний текст за допомогою комп'ютерних систем?
7. Наведіть приклади лінгвістичного дешифрування.
8. Ознайомтеся з азбукою Морзе для латинської та кириличної абеток. Хто нею користується в наш час?

### **III. Практичні завдання**

1. Запишіть різними видами звукового транскрибування (фонетичним, фонематичним, практичним) речення: *Студенти спілкуються сучасною українською літературною мовою.*
2. Передайте будь-які 20 українських власних назв графемами англійської мови.
3. Використовуючи ресурс <https://translit.net/>, транслітеруйте фрагмент тексту публіцистичного стилю довжиною в 1000 знаків без пробілів українською мовою латиницею, фрагмент тексту наукового стилю довжиною в 1000 знаків без пробілів англійською мовою російськими літерами.
4. Застенографуйте текст: *Стенографія – це скорочене швидкісне письмо. Для стенографії характерні особливі короткі знаки, скорочення й пропуски елементів слів, зливе написання слів і словосполучень. Стенографія дає змогу синхронно записувати усну мову й раціоналізувати техніку писання. Стенографічне письмо перевищує за швидкістю звичайне в 4-7 разів. Стенографія відома з глибокої давнини. У сучасній стенографії розрізняють системи геометральні й курсивні. Створили їх в Англії: геометральну – Дж. Уїлліс (1602), курсивну – С. Бордлі (1789). Геометральні системи застосовуються для мов з відносно короткими словами (наприклад, для англійської, французької,*

іспанської), а курсивні – для мов з довгими словами (наприклад, для слов'янських, скандинавських, німецької). За способом скорочень розрізняють системи стенографії морфологічні, або орфографічні (дотримуються орфографії звичайного письма, скорочуючи слова здебільшого їх початковими буквами) й фонетичні (для скорочення викидають звуки, які погано чути). Російська стенографія вперше застосована в 1860 році в Петербурзі на диспуті про походження Русі між істориками М. Погодіним та М. Костомаровим. У 1933 році в СРСР прийнято Державну єдину систему стенографії (ДЄСС), в основу якої покладено курсивну морфологічну систему М.М. Соколова. Українська стенографія створена в пожовтневий час. У Києві в 1920 році вперше почали працювати державні курси стенографії. Потім такі курси було відкрито в Харкові, Одесі, Вінниці та інших містах. На основі ДЄСС розроблено український стенографічний алфавіт.

5. Придумайте й зобразьте 5 знаків, які б зашифрували 5 ключових для освіти понять: «навчання», «студент», «викладач», «іспит», «університет».
6. За допомогою азбуки Морзе закодуйте повідомлення: *Привіт! Я – студентка університету.*

## **ТЕМА 5. ЦИФРОВІ (КОМП'ЮТЕРНІ) СИСТЕМИ ПИСЬМА. РОЛЬ СПЕЦІАЛЬНИХ СИСТЕМ ПИСЬМА**

### **I. Теоретичний мінімум**

З метою забезпечення взаємодії комп'ютерів з людьми постало завдання створити спеціальні шрифти (шрифтокомплекти), у яких літери позначалися б не графічними (як у традиційному письмі), а цифровими образами (цифровими кодами). До кожного цифрового коду прикріплювали відповідне графічне зображення знака.



До складу шрифтокомплектів, які поставляють з операційною системою Windows, як обов'язкові входять Ariel, Courier New, Times New Roman. У кожен із цих шрифтокомплектів входять 224 знаки, серед яких є латинська абетка, а також можуть бути деякі інші.

Шрифтокомплекти UNICODE містять понад 65 тис. знаків. У кожен із цих комплектів входять абетки та ієрогліфи більшості мов світу. У цих шрифтокомплектах коди 0-8191 (шістнадцяткові коди 0-1FFF) відведено під абетки всіх європейських мов, гебрейську, абетку арабських та індійських мов, а також фонетичні знаки; коди 8192-12287 (шістнадцяткові 2000-2FFF) – під пунктуаційні, математичні, технічні символи й орнаменти; коди 12288-16383 (шістнадцяткові 3000-3FFF) – під фонетичне китайське, японське та корейське письмо; коди 16384-59391 (шістнадцяткові 4000-E7FF) під китайські, японські та корейські ієрогліфи, а також каліграфію хань; коди 59392-64023 (шістнадцяткові E800-FDFF) – під приватне використання; коди 64024-65535 (шістнадцяткові FE00-FFFF) – під забезпечення сумісності з іншими шрифтами.

Завдання ПЛ у спеціальних системах письма:

- створення таблиць транскрибування як загальних (універсальних), так і конкретних (тобто для мов узагалі й для конкретних мов);
- створення таблиць транслітерування для конкретних мов на основі різних абеток;
- розроблення й стандартизація графіки стенографування й створення підручників зі стенографування;
- створення словників, граматик і підручників мови жестів та її вдосконалення;
- розроблення методів лінгвістичного дешифрування і здійснення самого дешифрування невідомих писемностей;
- відбір літер, лігатур, ієрогліфів тощо для UNICODE-шрифтокомплектів (тобто відбір множин знаків – літер різних абеток, лігатур, ієрогліфів,

цифр, пунктуаційних знаків, графічних, математичних, біологічних, хімічних тощо);

- кодування знаків шрифтокомплектів.

Завдання комп'ютерного забезпечення у спеціальних системах письма:

- створення систем транскрибування;
- створення систем транслітерування;
- створення електронних «секретарів» для незрячих, що дають їм змогу працювати на комп'ютері, а також читати та вводити тексти;
- створення систем перетворення мови жестів у писемну чи усну мову й навпаки (для людей з вадами мовного апарату);
- створення систем шифрування / дешифрування текстів;
- укладання комп'ютерних шрифтокомплектів.

## **II. Контрольні питання**

1. Опишіть етапи створення й еволюції шрифтокомплектів.
2. Що таке UNICODE? Опишіть його функціонал.

## **III. Практичне завдання**

Напишіть невелике висловлення на тему *«Який із вивчених видів спеціального письма я вважаю найважливішим у нашу еру і чому саме»*.

# **ТЕМА 6. ПРИКЛАДНА МОРФОЛОГІЯ**

## **I. Теоретичний мінімум**

Автоматичний морфологічний аналіз (АМА) є частиною процесу автоматичного опрацювання текстів (АОТ), основним завданням якого є розпізнавання їх змісту. Виокремлюють 2 види опрацювання: формальне – перетворення фрагментів тексту без аналізу його змісту; змістове (смісловне) – перетворення фрагментів тексту з аналізом його змісту,

встановленням логіко-семантичних відношень між його компонентами (використовується додаткова, семантична інформація, виражена в тексті імпліцитно). Завдання АМА: визначення частиномовної приналежності текстових одиниць; ідентифікація словоформ однієї лексеми.

У результаті АМА кожна словоформа тексту набуває свого коду частини мови і значення граматичних категорій (рід, число, відмінок, вид, час, особа тощо). АМА – процес, який забезпечує проведення морфемного, синтаксичного й семантичного аналізів, неможливих без попереднього визначення частин мови. Формальні граматичні ознаки дозволяють провести первинну семантичну характеристику, адже іменник узагальнено – предмет, дієслово – дія, прикметник – ознака, прислівник – ознака дії тощо. Подальший АМА (визначення відмінку й числа для іменників, часу й особи для дієслова тощо) дозволяє визначити синтаксичну будову речення: підмет, присудок, додаток, обставину (суб'єкт і об'єкт дії, детермінанти). Тобто морфологічна інформація – основа для визначення смислу речення.

Морфологія – це частина граматичної будови мови, що охоплює частини мови (граматичні класи слів), морфологічні (граматичні) категорії цих частин мови та їх форми. Серед основних завдань морфології як науки: 1) з'ясування принципів класифікаційного виділення частин мови; 2) виокремлення т.зв. морфологічної семантики слова; 3) опис формальних засобів, закріплених за певними частинами мови.

Автоматичне кодування слів тексту (приписування їм кодів граматичних класів) пов'язане з усвідомленням принципів традиційної граматичної класифікації. Необхідно виділити формальні ознаки слів, за допомогою яких можна було б автоматично визначати приналежність лексеми до певної частини мови.

Основною одиницею морфології є слово. Предметом вивчення морфології є граматичне значення – показник тих відношень, у які вступає слово в межах словосполучення, речення, тексту. Об'єктом вивчення

морфології є структура слова, форми словозміни, способи вираження граматичних значень.

Словоформи – це граматичні форми одного слова, тотожні лексично (спільне лексичне значення), але протиставлені граматичним значенням. Граматичне значення виявляється в певній граматичній формі (одне й те ж граматичне значення може мати різні матеріальні засоби вираження, або форми: *писатиму* й *буду писати* – для майб. часу; *студенту*, *студентові* – для Д.в.).

Морфологічне слово – це сукупність усіх словоформ (граматичних форм слів). Якщо така сукупність упорядкована, вона утворює парадигму.

Способи вираження граматичного значення:

- 1) синтетичний (граматичні значення в межах морфологічного слова виражаються за допомогою афіксів: закінчення, суфікс, префікс, інтерфікс тощо): *дядькові*, *кошеняти*, *відбігти*;
- 2) аналітичний (показник граматичних значень – службове слово): *буду мріяти*, *мріяв би*, *хай мріє*;
- 3) аналітико-синтетичний (поєднання двох попередніх – афіксальне граматичне оформлення слова + аналітичні елементи): *на вікні*, *в університеті* (грамема М.в.);
- 4) суплетивний (творення граматичних форм від різних коренів): *ти – тебе – тобі*, *поганий – гірший*.

Граматична форма – це мовний знак, за допомогою якого виражається граматичне значення. Наприклад: флексія: *поле – поля*; суфікс: *ягня – ягняти*; префікс: *читати – прочитати*; постфікс: *будувати – будуватися*; чергування: *сон – сну*; наголос: *вода – вóди*; службові слова: *хай думає*; суплетивізм: *я – мене*; порядок слів: *День змінює ніч. / Ніч змінює день*.

Для позначення класів однотипних граматичних значень вживається термін «граматична категорія». До морфологічних граматичних категорій належать категорії роду, числа, відмінка, особи, часу, способу, стану, виду.

Комп'ютерний морфологічний аналіз і синтез є основою функціонування практично всіх комп'ютерних лінгвістичних систем (КЛС). Аналіз передбачає здійснення: 1) розщеплення словоформи на морфеми; 2) лематизацію словоформи (перетворення її у вихідну форму); 3) визначення її граматичних характеристик. Синтез передбачає зворотній процес: якщо попередньо задано вимоги до словоформи (відомо, якою повинна бути її словотвірна / словозмінна основа, а також граматичні характеристики), то КЛС на основі баз даних морфем повинна вміти сама утворити потрібну словоформу.

Основними завданнями прикладної морфології є:

- створення таблиць словозміни;
- укладання списків морфем;
- визначення продуктивності й частотності морфем;
- визначення частот реалізації в текстах різних граматичних категорій;
- автоматичне визначення морфологічної будови слів у текстах;
- автоматична лематизація словоформ із текстів;
- автоматичне визначення морфологічних характеристик словоформ із текстів;
- створення систем морфологічного синтезу;
- автоматичне ймовірнісне визначення для нових невідомих слів їх морфологічної будови, способу їх відмінювання й приписування їм граматичних характеристик.

## **II. Контрольні питання**

1. Наведіть приклади формального та змістового АОТ.
2. Які формальні засоби можуть вказувати на ту чи ту частину мови?
3. Основною одиницею вивчення якого ще розділу науки про мову є слово?

### **III. Практичні завдання**

1. Як ви розумієте твердження Марчука Юрія Миколайовича:  
*«У комп'ютерній лінгвістиці поняття морфологічного аналізу є операційним. Якщо у традиційній лінгвістиці до морфологічного аналізу належить те, що характеризує форму і відповідає на питання «що» класифікують, то у прикладній лінгвістиці важливо не «що», а «як» отримують ту чи ту інформацію».*
2. Визначте частини мови в поданому реченні: *Студенти старанно виконують складне завдання.* Визначте, яким членом речення виступає кожна словоформа. Зробіть висновок, яким чином пов'язана частиномовна приналежність словоформ та їх синтаксична роль.
3. Чим лексичне значення відрізняється від граматичного? Наведіть приклади.
4. Охарактеризуйте формальні засоби вираження граматичних значень (флексія, суфікс, префікс, постфікс, чергування, наголос, службове слово, суплетивізм, порядок слів). З конкретними прикладами.
5. Що є вихідною формою для кожної з частин мови? Наведіть приклади.

## **ТЕМА 7. ПРИКЛАДНА ЛЕКСИКОГРАФІЯ**

### **I. Теоретичний мінімум**

Лексикографія (словникарство) (грецьк. *lexikos* – *словесний, словниковий*; *grapho* – *пишу*) – складова лексикології, яка розробляє теоретичні принципи й практичні засади укладання словників. Розмежовують теорію лексикографії (розроблення моделей будови словників, методик їх укладання, структур статей тощо) й практику лексикографії (сам процес укладання словників).

У суспільстві лексикографія виконує такі функції:

- описує, нормалізує та зберігає лексику мови;

- навчає мові (рідній – за допомогою тлумачних, іноземній – за допомогою перекладних словників);
- забезпечує міжмовні контакти (за допомогою дво- та багатомовних перекладних словників);
- фіксує результати дослідження мови (в етимологічних, синонімічних, антонімічних, омонімічних, частотних, обернених, ідеографічних тощо словниках; ці ж словники можуть бути використані, відповідно, й для вивчення мови).

Об'єктами лексикографії є морфеми, лексеми (а також їх парадигми), словосполучення (словники сполучуваності, словники фразеологізмів) і речення (словники прислів'їв і приказок).

Словник – множина певним чином упорядкованих одиниць мови, до яких можуть бути подані коментарі різного типу. Одиниці мови (морфеми, слова, словосполучення, речення), включені до словника як заголовкові, називають реєстровими. У словнику кожна реєстрова одиниця мови утворює статтю (мовою баз даних – запис). Кожен запис складається з окремих полів. Поля разом з їх розділювачами й способами виділення утворюють формат запису словника (комп'ютерних баз даних за аналогією).

Словники поділяють на енциклопедичні та лінгвістичні. Серед лінгвістичних є:

- за кількістю поданих мов (одномовні й багатомовні);
- за особливостями інформації (тлумачні, орфографічні, орфоепічні, синонімічні, антонімічні, омонімічні, фразеологічні, етимологічні, асоціативні, частотні тощо);
- за способом сортування інформації (алфавітні, гніздові, ідеографічні, обернені);
- за повнотою (мінімуми, середньої повноти й повні);
- за ступенем дотримання норм у реєстрових словниках (нормативні й ненормативні);
- за метою (навчальні й довідкові).

Термінографія – складова лексикографії, що досліджує теорію та практику укладання термінологічних словників.

Завдання прикладної лексикографії:

- укладання словників різних типів (зокрема, частотних);
- дослідження проблем теорії лексикографії;
- створення комп'ютерних словників – копій паперових та таких, що призначені для функціонування в складі КЛС;
- створення інтегрованих лексикографічних баз даних.

## **II. Контрольні питання**

1. Дайте визначення лексикології, лексикографії, термінології, термінографії.
2. Які періоди розвитку можна виділити в історії лексикографії? Опишіть їх.
3. Наведіть приклади лексикографічних комп'ютерних лінгвістичних систем.
4. Які наразі проблеми є ключовими у лексикографії загалом, та комп'ютерній зокрема?
5. Чим займається Державний комітет стандартизації в Україні?
6. Для чого потрібні термінологічні словники? Дайте визначення номенклатурам і тезаурусам.

## **III. Практичні завдання**

1. Ознайомтеся зі змістом статті:  
[https://uk.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA\\_%D1%81%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%D1%96%D0%B2\\_%D1%83%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D1%81%D1%8C%D0%BA%D0%BE%D1%97\\_%D0%BC%D0%BE%D0%B2%D0%B8](https://uk.wikipedia.org/wiki/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA_%D1%81%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%D1%96%D0%B2_%D1%83%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D1%81%D1%8C%D0%BA%D0%BE%D1%97_%D0%BC%D0%BE%D0%B2%D0%B8). Оберіть будь-які три словники й



- опишіть їх за такими критеріями: кількість поданих мов; спосіб сортування інформації; повнота; нормативність; мета; вид словника.
2. Укладіть свій власний термінологічний словничок (реєстр – 10 слів-термінів).

## ТЕМА 8. ПРИКЛАДНИЙ СИНТАКСИС

### I. Теоретичний мінімум

Виокремлюють три компоненти синтаксису: синтаксис непередикативних словосполучень, синтаксис передикативних словосполучень, синтаксис надфразних єдностей і блоків (останній називають ще граматиною тексту).

Для моделювання породження й сприймання синтаксичних конструкцій (словосполучень і речень) використовують граматику породження й граматику сприймання. Наприклад, для опису граматики породження текстів використовують формальну граматику Н. Хомського (G):

$$G = \{S, V, F, R\}, \text{ де}$$

**S** – початковий службовий символ, з якого починають породження словосполучення чи речення;

**V** – службові символи, за допомогою яких породжують речення (одним із цих службових символів є S);

**F** – слова, які на останньому етапі породження підставляють замість службових символів;

**R** – правила, які визначають послідовність породження й перетворення одних синтаксичних конструкцій на інші.

Приклад застосування такої граматики подано нижче.

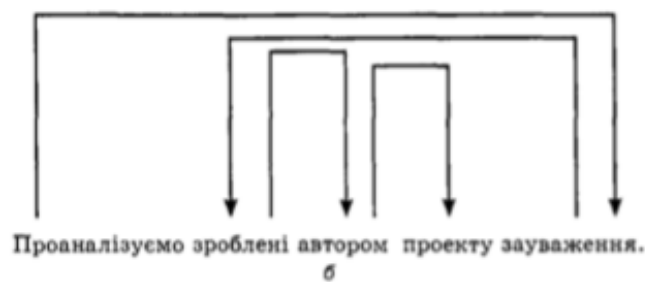
*Letan* – надання переліченим змінним таких значень (знак « := » позначає операцію заміни символа, який стоїть зліва від цього знака, символами, що стоять справа від нього):

$B = \{S, C, V, O\};$	(службові символи)
$F = \{\text{читає, Марічка, книжку}\};$	(слова для підставлення)
$R = \{$	(множина правил)
$S := CV;$	(правило 1)
$V := VO;$	(правило 2)
$C := \text{Марічка};$	(правило 3)
$V := \text{читає};$	(правило 4)
$O := \text{книжку}$	(правило 5)
$\}$	

2 етап – підставлення потрібних символів:

- |   |  |
|---|--|
| 1) $S;$                                 | (починаємо з початкового<br>символа)                           |
| 2) $S := CV;$                           | (застосовуємо перше правило<br>підстановки)                    |
| 3) $CV := CVO;$                         | (застосовуємо друге правило<br>підстановки)                    |
| 4) $CVO := \text{Марічка читає книжку}$ | (застосовуємо третє, четверте<br>й п'яте правила підстановки). |

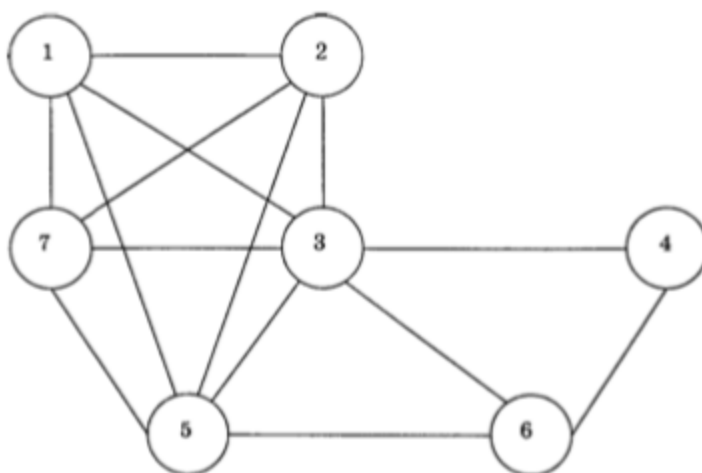
Для опису будови словосполучень, речень, надфразних єдностей і блоків використовують граfi – спеціальний математичний інструмент. Наприклад:



У правильно збудованому реченні дуги, що з'єднують слова, не можуть перехрещуватися.

За допомогою графів, але вже не ієрархічних, а сіткових, моделюють будову надфразних єдностей (кілька речень, об'єднаних засобами граматики тексту; на письмі часто збігається з абзацом) і блоків (кілька надфразних єдностей, також об'єднаних засобами граматики тексту; на письмі часто збігається з розділом, підрозділом чи параграфом). У надфразних єдностях граматика тексту реалізується за допомогою повторів, конекторів, паралельних синтаксичних конструкцій, аналогічних морфологічних характеристик словоформ, речень тощо. Нижче подано приклад сіткоподібного графа індійської казки, що утворює надфразну єдність (номер речення відповідає номеру вузла (в кружечках), а кожна дуга (лінія) між вузлами (реченнями) позначає граматикотекстовий зв'язок між цими реченнями).

(1) Був колись в одного багатого купця син і звали його Мугдхабуддхі.  
(2) Поїхав якимось Мугдхабуддхі торгувати на острів Катаха і навантажив корабель, попри все інше, порядним вантажем сандалового дерева. (3) От усі товари він уже продав, а на сандал нема покупця та й нема! (4) Ніхто в тій країні знати не знав та й ніколи не чув, для чого сандал використовують. (5) Дізнався він, що тамтешні люди у дроворубів вугілля купують. (6) Тоді перетворив дурень увесь сандал на вугілля, продав його за ціною вугілля й урочисто повернувся додому. (7) Всюди розтеревеним він, як кмітливо зі справами упорався, чим приніс людям утіху чималу.



Завдання прикладного синтаксису:

- типологізація непередикативних і передикативних синтаксичних конструкцій і визначення їх частотності;

- типологізація зв'язків граматики тексту в надфразних єдностях і блоках і визначення їх частотності;
- створення підручників практичного синтаксису;
- автоматичне виділення в текстах непередикативних і передикативних словосполучень і визначення їх будови;
- автоматичне визначення членів речення;
- автоматичний поділ речення на тему й рему;
- автоматичне встановлення граматичних зв'язків між реченнями в надфразних єдностях і блоках;
- створення систем синтаксичного синтезування словосполучень, речень, зв'язування речень у надфразних єдностях і блоках.

## **II. Контрольні питання**

1. Наведіть приклади непередикативного словосполучення, передикативного словосполучення, надфразної єдності.
2. Дайте визначення синтаксичного примітива. Для чого його використовують?
3. Дайте визначення конекторів у надфразних єдностях. Наведіть приклади.
4. Які речення в межах надфразної єдності вважаються зв'язаними?

## **III. Практичне завдання**

Знайдіть приклад надфразної єдності (приблизно 10 речень). Позначте кольорами засоби граматичного зв'язку в її межах. Побудуйте сіткоподібний граф дібраного вами прикладу.

## **ТЕМА 9. АНАЛІТИКО-СИНТЕТИЧНЕ ОПРАЦЮВАННЯ ДОКУМЕНТІВ: БІБЛІОГРАФУВАННЯ, ІНДЕКСУВАННЯ, ІНФОРМАЦІЙНИЙ ПОШУК, АНОТУВАННЯ, РЕФЕРУВАННЯ, КОРЕКТУРА, РЕДАГУВАННЯ, ПЕРЕКЛАД, ГОТУВАННЯ ОГЛЯДІВ**

### **I. Теоретичний мінімум**

*Аналітико-синтетичне опрацювання документів* – це аналіз документів і наявної в них інформації (текстової, графічної, аудіальної, відео тощо), який дає змогу синтезувати за певними правилами інший текст. Синтезування – або генерування на основі раніше створеного тексту зовсім нового, або трансформування за певними правилами раніше створеного тексту в модифікований. Аналітико-синтетичне опрацювання здійснюють тільки на основі зовнішніх текстів. Виконуючи його, використовують загальноприйняті алгоритми.

*Бібліографування* – це вид аналітико-синтетичного опрацювання документів, який полягає в укладанні опису документа. Такі описи можуть бути бібліографічними (описують опубліковані паперові видання), архівними (описують архівні документи), кінематографічними (описують кінофільми), інтернетними (описують веб-сторінки) тощо. Види й структуру бібліографічного опису визначають стандарти чи спеціально укладені норми. Періодичні збірники бібліографічних описів, які містять дані про нові публікації, називають сигнальною інформацією. В Україні такі збірники публікує Книжкова палата України. Існує низка систем бібліографування, призначених для бібліотек, які дають змогу створювати бібліографічні бази даних, вести їх і використовувати для пошуку потрібної літератури (наприклад, система «ІРБІС»). Існують також системи бібліографування, призначені для користувачів (авторів), що дають змогу створити індивідуальну базу даних бібліографічних описів (наприклад, система EndNote).

**Індексування** – це вид аналітико-синтетичного опрацювання текстів документів, який полягає у визначенні його основної теми і її записі інформаційно-пошуковою мовою. Інформаційно-пошукова мова – це штучна або обмежена природна мова, призначена для опису тем документів і проведення за її допомогою інформаційного пошуку. Зразком обмеженої природної мови може служити мова ключових слів, а прикладом спеціально укладеної штучної мови – Універсальна десяткова класифікація (УДК). Органи науково-технічної інформації часто використовують також рубрикатори.

**Інформаційний пошук** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого отримувач інформації на підставі сформульованого ним на комп'ютері запиту отримує з масиву документів ті, які за певними критеріями відповідають цьому запиту.

**Анотування** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого генерують короткий текст (до 600 знаків), що характеризує зміст, форму, цільове та реципієнтське (читацьке, слухацьке, глядацьке) призначення документа. Як правило, анотацію супроводжує бібліографічний опис документа.

**Реферування** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого генерують короткий текст (до 1000 знаків), що передає основний зміст документа. Для великих документів (книг) обсяг реферату може сягати 2500 знаків. Він повинен містити: тему документа; об'єкт і предмет дослідження; мету й характер роботи; використані методи й методики; отримані результати; висновки; сферу застосування результатів; бібліографію.

**Коректура** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого перевіряють відповідність копії документа його оригіналові й на основі виявлених у копії спотворень вносять у неї необхідні виправлення.

**Редагування** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого текст документа приводять у відповідність до чинних норм, а також здійснюють його творчу оптимізацію, тобто удосконалення. Серед алгоритмів редагування виокремлюють методи контролю (параметричні, спискові, шаблонні, структурні, аналітичні, когнітивні, положеннєві, компаративні, спеціальні) та методи виправлення (переставлення, видалення, заміна, вставлення, скорочення, опрацювання, перероблення, спеціальні методи).

**Переклад** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого одиниці мови тексту оригіналу заміняють на еквівалентні за змістом одиниці іншої мови за умови максимального збереження інформації, яку містить текст оригіналу.

**Готування оглядів** – це вид аналітико-синтетичного опрацювання текстів документів, у процесі якого за певною темою добирають низку документів, проводять критичну оцінку, систематизацію й узагальнення інформації, яку вони містять, а потім на цій основі генерують (створюють) нове повідомлення (огляд), що завершується висновками.

## **II. Контрольні питання**

1. Чим аналітико-синтетичне опрацювання відрізняється від створення текстів документів?
2. Ознайомтеся з роботою Книжкової палати України, розділ «Національна бібліографія». Які «літописи» вона укладає та за якими принципами?
3. Що таке рубрикатор? Наведіть приклади.
4. Наведіть приклади кількох інформаційно-пошукових систем, що є популярними наразі в Україні і світі.
5. Наведіть приклади анотації до книги, статті, кінострічки.
6. Користуючись ресурсом <http://www.nbuv.gov.ua/node/512>, знайдіть 3 реферати за темою своєї курсової роботи.
7. Де може знадобитися коректура?

8. Наведіть приклади норм редагування.
9. У чому суть перекладознавчих денотативної, трансформаційної, семантичної теорій, теорії рівнів еквівалентності?
10. Хто такі аналітики? Для чого потрібні огляди?

### **III. Практичні завдання**

1. Користуючись одним із ресурсів (<http://aphd.ua/pryklady-oformlennia-bibliorafichnoho-opysu-vidpovidno-do-dstu-83022015/> або <https://psycholing-journal.com/downloads/apa-style.pdf>), укладіть бібліографію до своєї курсової роботи (приблизно 20 найменувань).
2. Здійсніть індексування своєї курсової роботи: 1) вкажіть ключові слова (10 слів); 2) користуючись ресурсом <http://www.udcsummary.info/php/index.php?lang=uk&pr=Y>, визначте УДК.
3. Які ключові слова (словосполучення) ви формулюєте, щоб відшукувати інформацію в інтернеті для своєї курсової роботи? Яку інформаційно-пошукову систему використовуєте?
4. Сформулюйте анотацію до своєї курсової роботи (до 600 знаків). Перекладіть її англійською.
5. Напишіть реферат (1000-2500 знаків) за темою своєї курсової роботи.
6. Наведіть приклади коректурних знаків.
7. Знайдіть приклад тексту, що потребує редагування (до 1000 знаків). Користуючись ресурсом <https://www.languagetool.org/>, відредагуйте його. Виділіть кольором відредаговані місця.
8. Здійсніть переклад вашої улюбленої українськомовної поезії іншою мовою (російською / англійською на вибір) в будь-якому онлайн-перекладачі. Подайте цей переклад. Виділіть кольором місця, які, на вашу думку, потребують редагування.
9. Чи можливо автоматизувати процес створення оглядів? Відповідь аргументуйте.



# МОДУЛЬНА КОНТРОЛЬНА РОБОТА

## *ВАРІАНТ 1*

1. Мова, у трактуванні прикладних лінгвістів, є:

- а) основним носієм інформації в суспільстві;
- б) найважливішим засобом спілкування і пізнання;
- в) глибинним пластом духовного життя народу, його історичною пам'яттю, найціннішим надбанням століть.

2. У чому спільне й відмінне прикладної лінгвістики, комп'ютерної лінгвістики, кібернетики?

3. Мовний звук – це:

- а) коливання, які сприймаються сенсорною системою тварин і людини;
- б) звук лише з певними акустичними характеристиками, які утворюють його значення серед інших лінгвістичних звуків;
- в) найменша неподільна одиниця, яка служить будівельним матеріалом для одиниць вищого порядку: морфем і слів.

4. Які методи використовує прикладна лінгвістика для дослідження звуків і органів мовлення людини? Назвіть фізичні характеристики мовних звуків, що аналізує прикладна фонетика.

5. Встановіть відповідність між терміном та його значенням:

- |                             |   |
|-----------------------------|---|
| 1) письмо                   | а) мінімальний двомірний графічний символ (рисунок, піктограма), що позначає одиницю мови         |
| 2) графема                  | б) графічна знакова система, призначена для фіксації усного тексту на площинному носії інформації |
| 3) графічна знакова система | в) письмо, призначене для використання всіма членами суспільства                                  |

- |                      |   |
|----------------------|---|
| 4) традиційне письмо | г) сукупність (множина) графем (наприклад, абетка, множина лігатур, множина ієрогліфів)   |
| 5) графіка           | г) упорядкована множина графем, які позначають звуки певної мови й призначені для графічного відображення звукових образів слів |
| 6) спеціальне письмо | д) сукупність усіх символів, які використовують у певний час у писемному мовленні якоїсь мови                                   |
| 7) алографи          | е) письмо, призначене для використання лише якоюсь частиною суспільства   |
| 8) абетка            | є) різні варіанти написання літер   |
| 9) орфографія        | ж) історично усталена й загальноприйнята система правил фіксування усних текстів певної мови на письмі                          |

**6.** Ви, як філолог, якою зі спеціальних систем письма (транскрибування, транслітерування, стенографування, мова жестів, криптографування, азбука Морзе) хотіли б досконало оволодіти й чому саме?

**7.** Які з зазначених завдань щодо спеціальних систем письма виконує прикладна лінгвістика, які – комп'ютерне забезпечення?

- створення таблиць транскрибування як загальних (універсальних), так і конкретних (тобто для мов узагалі й для конкретних мов);
- створення систем транскрибування;
- створення таблиць транслітерування для конкретних мов на основі різних абеток;
- створення систем транслітерування;
- розроблення й стандартизація графіки стенографування й створення підручників зі стенографування;
- створення систем шифрування / дешифрування текстів;
- створення електронних «секретарів» для незрячих, що дають їм змогу працювати на комп'ютері, а також читати та вводити тексти;

- створення словників, граматики і підручників мови жестів та її вдосконалення;
- укладання комп'ютерних шрифтокомплектів;
- створення систем перетворення мови жестів у писемну чи усну мову й навпаки (для людей з вадами мовного апарату);
- розроблення методів лінгвістичного дешифрування і здійснення самого дешифрування невідомих писемностей;
- відбір літер, лігатур, ієрогліфів тощо для UNICODE-шрифтокомплектів (тобто відбір множин знаків – літер різних абеток, лігатур, ієрогліфів, цифр, пунктуаційних знаків, графічних, математичних, біологічних, хімічних тощо);
- кодування знаків шрифтокомплектів.

## 8. Що таке шрифтокомплект?

## 9. Встановіть відповідність між терміном та його значенням:

- |                                      |  |
|--------------------------------------|--|
| 1) автоматичний морфологічний аналіз | а) приписування словам тексту кодів граматичних класів   |
| 2) автоматичне кодування слів тексту | б) процес, який забезпечує проведення морфемного, синтаксичного й семантичного аналізів, неможливих без попереднього визначення частин мови          |
| 3) морфологія                        | в) граматичні форми одного слова, тотожні лексично (спільне лексичне значення), але протиставлені граматичним значенням                              |
| 4) слово                             | г) основна одиниця морфології  |
| 5) морфологічне слово                | г) частина граматичної будови мови, що охоплює частини мови (граматичні класи слів), морфологічні (граматичні) категорії цих частин мови та їх форми |

- |   |   |
|---|---|
| б) комп'ютерний морфологічний аналіз і синтез | д) основа функціонування практично всіх комп'ютерних лінгвістичних систем               |
| 7) граматичне значення                        | е) мовний знак, за допомогою якого виражається граматичне значення                      |
| 8) граматична форма                           | є) показник тих відношень, у які вступає слово в межах словосполучення, речення, тексту |
| 9) словоформи                                 | ж) сукупність усіх словоформ (граматичних форм слів)                                    |

10. Сформулюйте основні завдання прикладної морфології.

### **ВАРІАНТ 2**

1. Укладіть деревну модель системи приголосних фонем української мови за способом творення / за місцем творення / за твердістю і м'якістю (*оберіть один із параметрів*).
2. Перейдіть за посиланням <https://www.youtube.com/watch?v=LHeUY4fvunk&feature=youtu.be>. Які типи інформації (за класифікацією А. Білецького) наявні в ролику? Відповідь аргументуйте.
3. Знайдіть уривок тексту публіцистичного стилю українською мовою (*5 речень*). Використовуючи будь-яку систему машинного перекладу, перекладіть його російською та англійською мовами. Результат закріпте. Випишіть типові помилки. Запропонуйте свій варіант коригування перекладених уривків.
4. Запропонуйте фрагмент інфологічного етапу проектування лінгвістичної бази даних «Рівні української мови» у вигляді структурованої відповідним чином таблиці.
5. Чому штучний інтелект – модель природного?

6. Заскріньте приклади діалогової «мовленнєвої діяльності» вашого комп'ютера з вами. Такий діалог матиме жорстку чи м'яку структуру? Доведіть.
7. Представте структуру фрейма для терма «комп'ютер».
8. Представте алгоритм для лексичного / морфемного / морфологічного / синтаксичного розбору слова / речення (*оберіть один із параметрів*).
9. Який словник кращий – паперовий чи комп'ютерний? Відповідь аргументуйте.

## ЛІТЕРАТУРА

1. Баранов А.Н. Введение в прикладную лингвистику. М., 2001. 360 с.
2. Белоногов Г.Г. Компьютерная лингвистика и перспективные информационные технологии. М., 2004. 159 с.
3. Волошин В.Г. Комп'ютерна лінгвістика. Суми, 2004. 382 с.
4. Дарчук Н.П. Комп'ютерна лінгвістика. К., 2008. 351 с.
5. Карпіловська Є.А. Вступ до прикладної лінгвістики: Комп'ютерна лінгвістика. Донецьк, 2006. 188 с.
6. Марчук Ю.Н. Компьютерная лингвистика. М., 2007. 317 с.
7. Николаев И.С., Митренина О.В., Ландо Т.М. Прикладная и компьютерная лингвистика. М., 2016. 315 с.
8. Партико З.В. Прикладна і комп'ютерна лінгвістика: Вступ до спеціальності. Львів, 2008. 224 с.
9. Соснина Е.П. Введение в прикладную лингвистику. Ульяновск, 2012. 110 с.