

МІЖРЕГІОНАЛЬНА  
АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ



**МЕТОДИЧНІ МАТЕРІАЛИ  
ЩОДО ЗАБЕЗПЕЧЕННЯ САМОСТІЙНОЇ  
РОБОТИ СТУДЕНТІВ  
з дисципліни  
“ОСНОВИ ІНФОРМАТИКИ  
ТА ПРИКЛАДНОЇ ЛІНГВІСТИКИ”  
(для бакалаврів)**

Київ  
ДП «Видавничий дім «Персонал»  
2013

Підготовлено доцентом кафедри іноземних мов *О. І. Гузь*

Затверджено на засіданні кафедри теорії та практики перекладу  
(протокол № 6 від 30.01.09)

*Схвалено Вченою радою Міжрегіональної Академії управління персоналом*

**Гузь О. І.** Методичні матеріали щодо забезпечення самостійної роботи студентів з дисципліни “Основи інформатики та прикладної лінгвістики” (для бакалаврів). — К.: ДП «Вид. дім «Персонал», 2013. — 67 с.

Методичні матеріали щодо забезпечення самостійної роботи студентів містять пояснювальну записку, зміст самостійної роботи з дисципліни “Основи інформатики та прикладної лінгвістики”, питання для самоконтролю, а також список літератури.

- © Міжрегіональна Академія управління персоналом (МАУП), 2013
- © ДП «Видавничий дім «Персонал», 2013

## ПОЯСНЮВАЛЬНА ЗАПИСКА

Сучасний прагматичний час зумовив активізацію досліджень у сфері практичного застосування накопичених лінгвістичних досягнень та активний розвиток прикладної лінгвістики (ПЛ). Провідні університети світу не лише проводять підготовку фахівців за цією спеціальністю, а й залучають студентів до дослідження проблем ПЛ з практичним їх впровадженням у життя суспільства. Запорукою успіху в підготовці цих спеціалістів має стати чітке усвідомлення того, які напрями має прикладне застосування лінгвістичної теорії.

Розвиток лінгвістичних досліджень останніх десятиліть характеризується міждисциплінарною взаємодією з іншими науками (кібернетика, інформатика, семіотика, соціологія, психологія тощо). Інтегрованість зусиль вчених зумовила розвиток таких комплексних наук, як соціолінгвістика, психолінгвістика, нейролінгвістика, когнітивна лінгвістика, математична лінгвістика, кібернетична лінгвістика, комп'ютерна лінгвістика тощо.

Посилення міжнародних зв'язків стимулювало розвиток таких напрямів ПЛ, як інтерлінгвістика, що вивчає міжнародні мови як засіб міжнародного спілкування, та перекладознавство. Їх теоретичним підґрунтям можна вважати досягнення у галузі порівняльного мовознавства (контрастивної та конфронтативної лінгвістики).

Мова як явище соціальне функціонує в тісній взаємодії з суспільством: стан і розвиток мови залежить від суспільних процесів і навпаки — мова впливає на стан суспільства. Тому вивчення ПЛ як науки є дуже нагальним і важливим для майбутніх фахівців з іноземної мови.

Мета дисципліни — допомогти студентам опанувати основні положення ПЛ, розвинути їхні навички та вміння самостійно працювати над навчальним матеріалом.

# **ЗМІСТ САМОСТІЙНОЇ РОБОТИ**

*з дисципліни*

## **“ОСНОВИ ІНФОРМАТИКИ ТА ПРИКЛАДНОЇ ЛІНГВІСТИКИ”**

### **Тема 1. Прикладна лінгвістика, її співвідношення з загальним мовознавством та іншими дисциплінами**

Термін “прикладна лінгвістика” багатозначний. В Україні і за кордоном — різні інтерпретації: у них — в 30–40-х роках насамперед викладання іноземних мов, методика викладання, особливості опису граматики для навчальних цілей; у нас — з 50-х років пов’язано з розробкою комп’ютерних технологій і появою систем автоматичної обробки інформації. Тому в Україні ПЛ є комп’ютерною лінгвістикою, обчислювальною лінгвістикою, автоматичною лінгвістикою, інженерною лінгвістикою. За Звеґінцевим: “під ПЛ розуміють всі види автоматичної обробки мовленнєвої інформації — машинне розпізнавання усного мовлення, машинний переклад, автоматичне кодування... Але ПЛ не закривається в межах даної проблематики”.

Широке розуміння прикладної лінгвістики — застосовування наукових знань про устрій і функції мови у нелінгвістичних наукових дисциплінах і різних сферах людської діяльності, а також теоретичне осмислення цієї діяльності.

**Об’єкт:** мова в її природному стані, аналіз типових характеристик мови до їх застосовування в комунікації. Відкриття специфіки природного спілкування.

**Предмет** дослідження: конструктивне. Пізнання системних принципів устрою мови, розробка і вдосконалення структур і форм методів аналізу і опис мови.

**ПЛ і загальне мовознавство (ЗМ).** Спочатку ПЛ реалізовувала ідеї та теорії ЗМ. Класичні лінгвістичні дослідження займаються описом і вивченням мови лише на рівні спостереження (прим.: відтворення і сприйняття мовлення). ПЛ займається деталізацією мови на рівні лінгвістичних об’єктів, які важко спостерігаються, інколи не спостерігаються (прим.: рівні синтаксису, морфеми).

1. Емпіричні розділи Лінгвістики: семантика, діалектологія... Це — накопичення і первинна обробка мовних і мовленевих фактів.

2. Теоретична Лінгвістика узагальнює й усвідомлює всі ці факти, створює теоретичної концепції мови і текстотворення.

3. Прикладна Лінгвістика. Перевірка на практиці (методом моделювання) вірогідності гіпотез ЗМ. Потрібна пояснювальна сторона — як функціонує мова? Тільки тоді можна створити модель мови.

### **Зв'язок ПЛ з іншими науками**

Основна частина ПЛ з теоретичного мовознавства. Методологія логіки і математики, філософія мови — для повного абстрагування від дійсності, щоб створити цілісну модель. Фізіологія, логопедія (проблема афазій), експериментальна психологія, акустика, інформатика, кібернетика, системи штучного інтелекту (ШІ).

З функціональної точки зору ПЛ може бути визначена як наукова дисципліна, в якій вивчаються й розробляються способи оптимізації функціонування мови. Функції мови (комунікативна, когнітивна, інформаційна, епістемічна — як спосіб зберігання й передачі знань, соціальна...) дають точки відліку для класифікації великої галузі застосування лінгвістичних знань.

**Компоненти ПЛ:** 1) теорія і практика перекладу (+ машинний переклад); 2) теорія і практика викладання (рідна/іноземна мови) — лінгводидактика; 3) теорія і практика ППС; 4) теорія кодування; 5) соціолінгвістика; 6) орфографія і орфоепія; 7) теорія впливу; теорія аргументації; 8) політична лінгвістика; 9) лексикографія (в тому числі автоматична); 10) термінологія, термінографія; 11) польова лінгвістика (дослідження мов малих народностей); 12) комп'ютерна лінгвістика (моделювання окремих аспектів мови); 13) лінгвістична кримінологія; 14) психолінгвістика; 15) квантитативна лінгвістика.

### **Основні дослідницькі задачі ПЛ**

1. *Створення лінгвопроцесорів* — системи автоматичного аналізу та синтезу мовлення і тексту на природних мовах. Його компоненти: морфологічні, семантичні, синтаксичні аналізатори.

2. *Машинний переклад* — передача плану висловлення шляхом планів змісту різними мовами, в ідеалі — без участі людини.

Моделі: “текст — текст”, “текст — смисл — текст”, “текст — дійсність — смисл — текст”.

3. *Прикладна лексикологія і машинна лексикографія.* Автоматичне складання словників для розробки автоматичної ППС, машинного перекладу. Частотні словники, словопоказчик, конкорданси, термін. Бази даних, машинні словники, енциклопедії.

4. *Системи автоматичного аналізу й синтезу мовлення.* Спрямовані на розробку природно-мовного інтерфейсу. Розпізнавання мовлення потребує залучення всіх рівней організації мови.

5. *Автоматизована обчислювальна система.* Підвищення якості та ефективності навчання, спрямовані на самостійне вивчення.

Загалом прикладні аспекти лінгвістики забезпечення різних сфер людської діяльності сходяться до однієї спільної проблеми — проблеми обробки інформації, яка функціонує у суспільстві. Це інформація текстова і в письмовому вигляді, та усне мовлення як найбільш звичний спосіб комунікації. Особлива роль мовознавства у вирішенні практичних проблем і потреб суспільства визначається самою сутністю природної людської мови, яка є унікальним засобом зберігання й передачі інформації.

Актуальні практичні проблеми і потреби суспільства — це впровадження у виробництво станків і приладів, які керуються голосом оператора, систем автоматичного пошуку й обробки інформації різними мовами, вдосконалення організації зв'язку (телефонного, радіозв'язку та ін.), застосування фонетичних даних під час лікування порушень мови и слуху.

Зазначені проблеми включають в себе ряд нових напрямів ПЛ, пов'язаних з автоматичною обробкою тексту (мовлення), з широким використанням ЕОМ, таких як: лінгвістичне забезпечення інформативних систем різних типів; машинний переклад; комп'ютеризація навчання; розробка систем, які розуміють природну мову (лінгвістичні задачі в системах штучного інтелекту); розробка систем використання інформації, яку має звуковий мовленнєвий сигнал.

Прикладна лінгвістика виникла в результаті гострої потреби знати, що таке мова у дії, в комунікації, оскільки теоретична лінгвістика вивчала переважно мову в її стані, в системі. При цьому вона довго сторонилася своєї спорідненості з науково-технічною і діловою прозою. Не було опису, моделей, орієнтова-

них на власно прикладні задачі. Нарешті, в 70-х роках ХХ ст. вчені дійшли висновку, що рішення багатьох прикладних проблем не може бути суто лінгвістичним, а лежить на шляхах моделювання поведінки і мислення людини, семантики, синтезу формальних і семантичних засобів мови. Так з'явилась одна з найважливіших міжгалузевих фундаментальних проблем прикладної спрямованості — *проблема моделювання знань*.

Моделювання знань останнім часом виявилось додатком до різних наук: логіки, лінгвістики, математики, психології, кібернетики. Причина такого загального інтересу до цієї проблеми одна — створення автоматичних систем ШІ. Знання даються нам у мові і через мову, існують як у текстах спонтанного мовлення (монолог, діалог, репліки в різних ситуаціях), так і в письмових текстах (пам'ятники стародавньої письменності, художня література, наукова і технічна література). Через мову ми передаємо свої знання від покоління до покоління. Мова, таким чином, виступає і як форма зберігання знань, і як засіб та спосіб їх передачі. В науці і техніці немає знань без текстів, а моделювання семантики науково-технічних текстів — це моделювання системи знань цієї галузі. Так поступово проблема моделювання знань зійшлася з проблемою моделювання смислу тексту. І тут центральним є питання про створення семантичних презентацій тексту.

Семантика проходить крізь всі рівні мови і тим самим не є окремим рівнем. В семантиці постійно виявляється вплив, з одного боку, методів власне лінгвістичних, а з іншого — методів логічної семантики, теорії класифікацій.

Під час моделювання лексичної семантики у структурну лінгвістику виникли теорії семантичного поля, компонентного аналізу, тезаурусного опису лексики. В синтаксичній семантиці було розроблено велику кількість концепцій семантичної структури речення.

Ядром сучасної ПЛ є структурна і математична лінгвістика в усьому різноманітті їх методів і досягнень.

Основним предметом структурної і математичної лінгвістики є розробка і вдосконалення структури, форми методів аналізу й опису мови. В цілому, за своєю сукупністю цілей, напрямів, методів і підходів, ПЛ — це насамперед комплексна дисципліна, яка постійно торкається філософії, психології, фізіології, математики, логіки, соціології, інформатики.

### **Питання для самоконтролю**

1. Що таке прикладна лінгвістика? Об'єкт і предмет її дослідження.
2. Як прикладна лінгвістика пов'язана з загальним мовознавством?
3. Компоненти та основні дослідницькі задачі ПЛ?
4. Де застосовується моделювання знань?
5. Що є ядром прикладної лінгвістики?

*Література* [1–3; 5; 16; 29; 57; 58; 67]

### **Тема 2. Різні найменування області прикладної лінгвістики і їх смислові відмінності. Універсальні прикладні проблеми**

**Прикладна лінгвістика** — це комплексна наукова дисципліна, яка вивчає мову в різних ситуаціях і розробляє методи вдосконалення мовних систем і мовних процесів.

Термін “прикладна лінгвістика” (ПЛ) з'явився в кінці 20-х років ХХ ст., коли була усвідомлена необхідність у науковому рішенні прикладних задач з використанням методів формального лінгвістичного аналізу письмового і акустико-лінгвістичного аналізу усних повідомлень.

За кордоном під ПЛ розуміють вдосконалення методів навчання мови (дидактична лінгвістика). У нашій країні ПЛ розуміють як комп'ютерну лінгвістику, яка стає майже синонімом ПЛ.

Лінгвістика належить до комплексу когнітивних наук, які об'єднуються за їх інтересом до проблем організації, уявлення, обробки і використання знань.

#### **Синоніми ПЛ:**

*Комп'ютерна лінгвістика* (машинна лінгвістика) — дисципліна, яка розробляє лінгвістичні аспекти комп'ютеризації.

#### *Обчислювальна лінгвістика.*

Термін “комп'ютерна лінгвістика” ширший за термін “обчислювальна лінгвістика”, оскільки створює загальну орієнтацію на використання комп'ютерів для вирішення різних наукових і практичних задач, не обмежуючи способи вирішення цих задач.

Англійський еквівалент *computational linguistics* може перекладатися і як комп'ютерний, і як обчислювальний.



Термін “обчислювальна лінгвістика” може розумітися вужче, оскільки навіть під час широкого трактування поняття обчислювання за його межами залишається таке вирішення лінгвістичних задач, як представлення знань, організація банків мовних даних, психолінгвістичні аспекти взаємодії людини і комп’ютера та ін.

*Структурна лінгвістика* (СЛ) — сукупність поглядів на мову і методи її дослідження, в основі яких лежить розуміння мови як знакової системи з чітко виділеними структурними елементами (одинацями мови, їх класами) і прагнення до строгого (як у точних науках) формального опису мови.

Свою назву структурна лінгвістика отримала завдяки особливій увазі до структури мови, яка є мережею відношень (протиставлень) між елементами мовної системи, які впорядковані і знаходяться в ієрархічній залежності в межах певних рівней. Структурний опис мови передбачає такий аналіз реального тексту, який дає можливість виділити узагальнені інваріантні одиниці (схеми речень, морфеми, фонеми) і співвіднести їх з конкретними мовленнєвими сегментами на основі строгих правил реалізації. Ці правила визначають межі можливого варіювання мовних одиниць в мовленні. Залежно від рівня аналізу правила реалізації формуються як правила позиційного розподілу конкретних, наприклад, принцип додаткової дистрибуції в фонології і морфології (дистрибутивний аналіз), або як трансформаційні правила в синтаксисі (під час трансформаційного аналізу), які регулюють перехід від інваріантної глибинної структури речення до значної кількості її реалізації. На базі СЛ розвинулась породжуюча граматики (генеративна лінгвістика); ідеї структурного аналізу багато в чому визначили постановку і вирішення задач, пов’язаних з машинним перекладом; СЛ дала можливість ширше проникнути в лінгвістику математичних методів (математична лінгвістика).

На СЛ вплинули праці багатьох учених, таких як: Сепір, Блумфілд, Ф. де Соссюр, один із провідних теоретиків – Якобсон; у росіян — Реформатський (знакова теорія мови), Ревзін, Холодович (загальна теорія моделювання); Апресян, Арутюнова, Гак, Залізник, Звєгінцев, Мельчук, Успенський та ін. (практичне використання методів СЛ).

*Математична лінгвістика* (МЛ) — математична дисципліна, предметом якої є розробка формального апарату для опису будови природних і деяких штучних мов. Виникла в 50-ті роки ХХ ст.; одним із головних стимулів появи математичної лінгвістики була нагальна потреба в уточненні основних лінгвістичних понять. Методи МЛ мають багато спільного з методами математичної логіки — математичної дисципліни, яка займається вивченням будови математичних роздумів, — і особливо таких її розділів, як теорія алгоритмів і теорія автоматів.

*Контрастивна лінгвістика* (КЛ) (порівняльна лінгвістика) — порівняльне вивчення двох, рідко кількох мов для виявлення їх подібностей і відмінностей на всіх рівнях мовної структури з метою типологічної класифікації мов. Як правило, контрастивна лінгвістика оперує матеріалами на синхронному зрізі мови. КЛ з'явилась й інтенсивно розвивалась в 50-ті роки ХХ ст., однак її появу підготували роботи Е. Д. Поліванова, Бодуена де Куртене, Л. В. Щерби, в яких були викладені теоретичні основи порівняння рідної та інших мов. У 70-ті роки контрастивні дослідження в окремих країнах (головним чином в США) використовували породжуючу модель Хомського, з приведенням явищ двох мов, які порівнюються, до спільної глибинної структури; нині спостерігається відхід від цієї методики на користь структурно-функціонального підходу.

### **Штучний інтелект (ШІ)**

Автоматичний переклад — дія, яка виконується на комп'ютері (перетворення тексту з однієї мови в еквівалентний за змістом текст іншою мовою).

Універсальні прикладні проблеми:

- створення і вдосконалення алфавітів і писемності (вирішення повністю і успішно) на трьох стадіях: поява писемності; книгодрукування; комп'ютеризація;
- створення систем транскрипції усної мови, систем транслітерації іншомовних слів;
- створення словників (лексикографія) (перші словники — глосарії — коментарі до церковних текстів), створення автоматичних словників, тезаурусів;
- уніфікація і стандартизація науково-технічної термінології;
- вивчення процесів і створення правил утворення нових назв виробів, товарів і т. д.;

- усний і письмовий переклад, розробка систем машинного перекладу;
- АРМи;
- навчання рідній та іноземній мовам, розробка відповідних методик (навчання дітей і дорослих, навчання емігрантів та ін.);
- автоматичне розпізнавання і синтез мовлення;
- створення і вдосконалення інформаційного методу для запису інформації;
- автоматичні методи переробки текстової інформації;
- створення автоматизованих систем інформаційного пошуку;
- створення автоматичних словників і систем машинного перекладу;
- розробка методів автоматичного анотування, реферування і перекладу;
- розробка експертних систем;
- лінгвістичне забезпечення АСУ;
- проблеми мови и роду (*politically correct non-sexist language*);
- створення систем стенографії, систем письма для сліпих;
- лікування мовленнєвих порушень;
- аналіз дискурсу.

### **Питання для самоконтролю**

1. Чим відрізняється комп'ютерна лінгвістика від обчислювальної?
2. Що зумовлює назву структурної лінгвістики?
3. Що досліджує математична лінгвістика?
4. Чим займається контрастивна лінгвістика?
5. Що досліджує штучний інтелект?

*Література* [4; 5; 7; 8; 10; 12; 13; 23]

### **Тема 3. Поняття рівня в теоретичній і прикладній лінгвістиці**

Ідея рівневої організації мови отримала широке розповсюдження в середині 20-х років ХХ ст. спочатку в американській дескриптивістській лінгвістиці, а пізніше і в інших напрямках, зокрема у вітчизняній лінгвістиці.

Рівні мови — деякі частини мови; підсистеми спільної мовної системи, кожна з яких характеризується сукупністю відносно однорідних одиниць і набором правил, які регулюють їх використання і групування в різні класи і підкласи. Поділ за рівнями в межах теоретичної лінгвістики:

- фонемний;
- морфемний;
- лексичний (рівень слова);
- синтаксичний (рівень речення).

Рівнеутворюючими якими володіють лише ті одиниці мови, які підкорюються правилам рівневої відповідності, тобто володіють здатністю вступати в парадигматичні й синтагматичні відношення лише з одиницями того самого рівня. З одиницями другого рівня одиниці будь-якого рівня вступають лише в ієрархічні відношення типу “складається з...”, “входить у...”. Так фонемні можуть утворювати класи і поєднуватися в мовленнєвому ланцюзі лише з фонемами, морфемами — с морфемами, слова — лише зі словами. В той же час фонемні входять у звукові оболонки морфем, морфемні — в слова, слова — в речення. Групування одиниць мови всередині рівней, наприклад, фонем (голосні й приголосні), морфем (кореневі, афіксальні), слів (знаменні, службові і т. д.) не є рівнеутворюючими.

Рівень мови слід відрізнити від рівня аналізу мови — фаз чи етапів розгляду мови. В лінгвістичній практиці онтологічний рівень мови і процедурний рівень аналізу (операційний) нерідко змішуються, хоча між ними немає прямої відповідності. Рівні аналізу залежать від цілей і завдань дослідження, тобто багато в чому визначаються точкою зору дослідника на об’єкт, що досліджується.

Б. Городецький визначає рівень мови як сукупність подібно до функціонуючих одиниць разом з відношеннями, які їх пов’язують.

Структура мови ділиться на два відносно замкнені (самостійні і незалежні) плани — план вираження і план змісту (ПЗ), всередині кожного з яких розрізняють індивідуальний набір рівней, тобто рівні плану змісту не ізоморфні рівням плану вираження (ПВ) (полісемія, омонімія, синонімія; українським глосним не відповідають ніякі одиниці ПЗ).

Як і мовна структура, в цілому рівні керують обома сторонами мовленнєвої діяльності — аналізом і синтезом, оскільки є їх структурною основою, однак не слід прив’язувати мовний рівень

до конкретної процедури аналізу/синтезу. Мовний рівень — лінгвістична універсалія.

**Склад рівней ПВ:**

- 1) рівень фонему;
- 2) рівень морфеми;
- 3) рівень слова;
- 4) рівень словосполучення;
- 5) рівень речення.

**Склад рівней ПЗ:**

- 1) морфо-семантичний рівень (одиниці — значення морфем);
- 2) лексико-семантичний рівень (одиниці — лексеми (значення слів)).

**Питання для самоконтролю**

1. Які рівні мови слід розрізняти?
2. Які одиниці мови володіють рівнеутворюючими якостями?
3. Які плани має структура мови?

*Література* [2; 3; 6–8; 16; 20; 51]

**Тема 4. Типи лінгвістичних моделей; основні вимоги до них і критерії оцінки**

**Модель** у лінгвістиці — це штучно створене лінгвістом реальний чи мисленнєвий прилад, який відтворює, імітує своєю поведінкою (звичайно у спрощеному вигляді) поведінку оригіналу в лінгвістичних цілях.

Типи лінгвістичних моделей:

- 1) за охопленням структури мови: загальні (глобальні) намагаються охопити всю мову (*vocabulary, grammar*) і специфічні: фонетична модель української мови, модель системи голосних тощо;
- 2) за типологічним статусом: універсальні намагаються охопити всі мови світу; специфічні характерні для певної мови чи групи мов: м'якість — твердість приголосних української мови (не діє в англійській, французькій);
- 3) за гносеологічним статусом: моделі мови; моделі лінгвістичних знань різних фонетичних шкіл; моделі діяльності лінгвіста;
- 4) за зображеним аспектом мови і мовленнєвої діяльності.

Моделі розрізняються не лише за спрямованістю на певний об'єкт, а й за використаними засобами моделювання (алгоритму чи обчисленню). Алгоритм — строга незмінна послідовність правил. Обчислення — множина дозволених правил (порядок виконання не важливий).

Аналізуючі моделі моделюють процес розуміння, використовують логічний засіб — алгоритм. Синтезуючі моделі моделюють процес вербалізації, смисл мовленнєвого уривку.

У породжуючих моделях (автор Н. Хомський) об'єкт моделювання — множина правильних мовленнєвих уривків. Складаються правила розрізнення прийняттого і неприйняттого; логічний засіб — обчислення, не є вираженням смислу; на виході — ланцюг елементів (граматично правильних речень).

Власно структурні моделі є основою всіх інших об'єктів моделювання — структура мови як така; логічний апарат — логіка.

5) за кінцевою метою дослідження: теоретичні, описові, прикладні;

6) за використаними методами: математичні, психологічні, соціологічні моделі;

7) за функціональним статусом: абстрактно узагальнюючі моделі, діючі;

8) за використаними матеріальними засобами: графічні, символні, комп'ютерні.

Певна модель звичайно входить в набір певних моделей, які описують певний рівень мови:

- фонологічний;
- морфологічний;
- синтаксичний;
- лексико-семантичний.

#### **Основні теоретичні вимоги до моделі:**

1) повнота моделі — здатність відображати всі факти, на які вона розрахована, на охоплення яких вона претендує;

2) простота — використання якомога меншої кількості засобів (символів, правил) для досягнення поставленої наукової мети;

3) пояснювальна сила — здатність моделі відкривати причини фактів, які спостерігаються і передбачати нові факти (наприклад, моделі історичної зміни слова; системи машинного перекладу в дуже незначному ступені пояснення);

4) адекватність — якість максимальної схожості об'єкта, що моделюється, на оригінал можна звести до пояснювальної сили чи теоретико-множинної відповідності;

5) економічність — економічне використання енергетичних і тимчасових ресурсів під час використання моделі;

6) точність — можливість виконання операцій, які надаються моделлю, формальним апаратом;

7) естетичні якості — краса моделі.

Прикладні критерії: головне — зручність моделі. Для моделювання мови дуже важливі логічні засоби реалізації моделі (комп'ютерне втілення моделі).

### **Питання для самоконтролю**

1. Що являє собою модель у лінгвістиці?
2. Критерії класифікації лінгвістичних моделей?
3. Чим відрізняється алгоритм від обчислення?
4. Що описує набір певних моделей?
5. Основні теоретичні вимоги до моделі.

*Література [28; 34; 39; 51]*

### ***Тема 5. Метамови формального опису семантичних структур***

Семантичні метамови (СМ) розрізняються:

- за об'єктом, який вони описують (морфема, лексема, словосполучення, речення, текст у цілому);
- за аспектом мовної структури, який вони відображають (парадигматичний аспект).

Сходження СМ — можливість перекладати з однієї СМ на іншу.

За об'єктом, що описується:

1. Значення морфем СМ:

- СМ компонентного аналізу (Найда, Апресян, Катс);
- СМ пресуппозиційного аналізу (прототипічний аналіз): якщо висловлювання піддати запереченню, то заперечується лише імпліцитно висловлена частина; те, що не заперечується — асерція.

2. Значення лексем досліджується в синтагматичному аспекті — сполучення лексем: теорія семантичної валентності, поняття моделі управління (Апресян); в парадигматичному аспекті: глибинні відмінки (агенс, інструменталіс), відмінкова рамка Філмора.

### 3. Значення словосполучень досліджується:

- у парадигматичному аспекті за допомогою тих самих СМ опису, що й лексеми;
- у синтагматичному плані: мова лексичних параметрів і функцій (Апресян), поняття лексичного параметра пов'язано з поняттям стійкості сполучення слів у словосполученні.

### 4. Значення речення в парадигматичному аспекті:

- семантична сітка (Скрегг: “Семантическая сеть как модель памяти”), мова обчислення предикатів (Дж. Лакофф “Постулаты речевого общения”), імплікація, пресупозиція, пропозиція, умови щирості, умови умотивованості).

5. Значення тексту в цілому: в парадигматичному аспекті використовуються такі макроструктури, як сценарії, фрейми, плани, сцени.

Фрейм — базова структура уявлення знань.

Сценарій динамічний — набір фреймів, які розгортаються в часі.

Семантична сітка (Чейф) — теорія організації пам'яті.

Існує три види пам'яті:

- поверхнева (surface);
- короткочасна (shallow);
- довготривала (deep);
- ієрархія подій -> ієрархія подій різної особистісної значущості;
- особиста залежність;
- поняття існуючої свідомості (consciousness);
- мова концептуальних залежностей (Шенк);
- мова шаблонів (templates, Уілкс:).

### 6. Процеси:

- розуміння (Шенк);
- вербалізації (Макк'юін).

Роджер Шенк: знання не можуть бути пояснені лінгвістичними термінами: -> мова концептуальних залежностей: P-Trans (фізичне переміщення), M- Trans (інтелектуальне переміщення).

Шенком була постульована незалежність мовного уявлення від електронного мовознавства, тим не менш мова комп'ютерного засобу виявилася прив'язаною до поверхневого синтаксису англійської мови. Кожне висловлювання запускає ланцюг концептуальних ви-



снівків (inferences), які дають можливість правильно зрозуміти ситуацію (Міла була голодна, вона взяла путівник Мітчеліна).

Модель SAM (Script Applying Mechanism) є комп'ютерною програмою, яка дає можливість зрозуміти зв'язність тексту за рахунок використання сценаріїв:

POLITICS (веде діалог, моделює політичну ідеологію);

RAM -> TALE-SPIN — породження казок;

FRUMP — машинне реферування повідомлень кількома мовами, читання, яке спирається на поняття інтересу (Integral Partial Parser).

Категорії, які зустрічаються у Шенка:

- інтерес;
- пояснення;
- пам'ять (організація концептуальної пам'яті);
- очікування;
- розуміння того, хто помиляється.

Уілкс “Аналіз речень англійської мови” вводить єдину форму уявлення — шаблон (template), інтуїтивною відповідністю якого можна вважати базисну форму уявлення “агенс — дія — об'єкт”. Шаблони складаються з більш дрібних блоків — формул, які відповідають тлумаченню окремих лексем. Для того щоб побудувати повне семантичне уявлення тексту (семантичний блок), шаблони об'єднуються за допомогою структур більш високого рівня — надшаблонів (paraplates) і правил доходження висновків; кожна з готових структур побудована на базі 80 атомарних семантичних елементів, а також функцій і предикатів, які задаються цією множиною.

Система на LISPe перекладає тексти з англійської на французьку.

Для рішення неоднозначності використовується ланцюг умовиводів. Немає синтаксичного аналізу в звичайному розумінні, одразу переходить до семантичних уявлень; немає в явному вигляді ніякої синтаксичної інформації, лише формула (Ф), яка має вигляд ДЗ; головним вважається сам правий елемент Ф, який задає фундаментальну категорію для всієї Ф; має глибинні відмінкові елементи; більша гнучкість під час описування смислу, ніж у методі Фодора і Катса, в яких немає глибинних відмінкових елементів. Шаблони мають вигляд сітки, яка складається з фреймів.

Шаблон складається з 3-х вузлів: агенс, дія, об'єкт.

## 2. Гіпертекстові системи.

Термін “гіпертекст” (ГТ) було введено Тедом Нельсоном в 1965 р. для описування документів, які виражають нелінійну структуру ідей. Однак, сама ідея ГТ пов'язується з Ванневаром Бушем (1945 р.), який теоретично обґрунтував проект технічної системи “Метех”, яка дозволяла зв'язувати тексти і їх фрагменти за асоціативними відношеннями. Але через відсутність комп'ютерів проект виявився складним для технічного впровадження. Його було впроваджено в системі Т. Нельсона “Ксанаду”, яка вже передбачала використання комп'ютерів.

З одного боку, ГТ — це особлива форма пред'явлення текстового матеріалу, його нелінійна організація. Гіпертекст — це текст, смислові елементи якого можуть читатися в різній послідовності.

З іншого боку, ГТ — це новий спосіб, інструмент, інформаційна технологія, яка робить можливим пред'явлення і засвоєння великих обсягів текстового матеріалу в нелінійній формі.

### Технологічні якості ГТ:

- різноманітність (гіпермедія, тобто не тільки текст, але й звук, і графіка, і відео);
- нелінійність (немає стандартної послідовності читання).

Послідовність читання — вільна. Між текстовими фрагментами зазначено дозволені переходи. Як правило, від одного фрагмента можна перейти до кількох інших. Читати можна з будь-якого елемента в різних напрямках. Комп'ютери дають можливість миттєво переходити від одного фрагмента тексту до другого, що дає змогу читати нелінійні тексти так само легко, як і лінійні.

Гіпертекст — комп'ютеризований нелінійний текст. Особливості ГТ порівняно з традиційними нелінійними текстами (текст з коментарями, виносками, енциклопедія):

1. Вираження сітьової структури: вузли (текстові фрагменти (ТФ) і зв'язки (можливість переходу від одного формульного тексту (ФТ) до другого -> навігація). Зв'язок може бути односпрямованим від фрагмента до коментаря чи двоспрямованим між двома фрагментами. Зв'язок може носити ієрархічний характер — від спільного до цілого.

## **Компоненти ГТ:**

Структурно ГТ являє собою граф, у вузлах якого знаходяться традиційні тексти/їх фрагменти, зображення, таблиці, відеоролики. Вузли, пов'язані різними відношеннями, які можуть бути одно- або двонаправленими. Ланцюг вузлів, через які проходить читач при перегляді компонентів тексту, отримує шлях (маршрут). Сукупність суміжних вузлів отримує оточення цього вузла. Оточення вузла створюють ті вузли, в яких утримується інформація, близька за семантикою до змісту цього вузла. Вузли сітки, в які входить і виходить багато стрілок-відношень, отримують центральну частину ГТ, а ті, які майже ізольовані від інших вузлів — його периферію.

## **Види ГТ:**

I. Структурно ГТ може бути:

- **ієрархічний** — деревоподібним — суттєво обмежує можливість переходу між його компонентами. Відношення нагадують структуру тезауруса, який спирається на родовидових зв'язках (Gohper);
- **сітьовим** — дає можливість використовувати різні типи відношень між компонентами, не обмежуючись відношеннями “рід — вид”.

II. За можливістю програмного забезпечення:

- **простий** (напр., електронне найменування документа, що дає можливість перейти до будь-якої частини документа, минувши перегляд всього документа);
- **складний** — багатосистемність переходів між компонентами ГТ, в них немає уявлення про базовий текст, з яким пов'язані другорядні за значенням тексти.

III. За способом існування:

- **статистичний** (не змінюється у процесі експлуатації);
- **динамічний** (зміни є нормальною формою існування). Звичайно функціонують там, де необхідно постійно аналізувати потік інформації, тобто в інформаційних службах різного роду.

IV. За зв'язком між елементами:

- **ГТ жорсткої структури** — відношення між елементами фіксуються творцями;

- **ГТ м'якої структури** — відношення виникають кожного разу, коли користувач звертається до ГТ. Технологія спирається на семантичний аналіз близькості документів та інших джерел інформації один до одного. Нині розповсюджено використання технологій м'якої структури на ключових словах. Перехід від першого вузла до другого здійснюється в результаті пошуку ключових слів. Так як набір ключових слів кожного разу може бути різним, то змінюється і структура ГТ.

Чим більше вузлів, тим переважно м'якша структура, оскільки кодування жорстких зв'язків віднімає надто багато часу. Структура Інтернет часто функціонує як ГТ м'якої структури.

Інтерфейс користувача: перехід — не більше двох клавіш.

Користувач повинен мати засоби орієнтації: до кожного вузла дається заголовок — локальна орієнтація. Глобальна орієнтація — наочне зображення структури гіпертекстів мережі, в якій відмічається шлях, який пройшов користувач. Багатовіконна система дає можливість бачити одночасно кілька формульних текстів на екрані.

2. Відкритість гіпертексту (для включення нових формульних текстів, читач може виступати співавтором).

Приклад ГТ: зображувальна і звукова система Гіпермедіа (комплексне використання інформації різного походження, синтез різних мистецтв). Динамічний ГТ постійно доповнюється новими текстовими фрагментами (необхідно знаходити зв'язки для ФТ, що надходять ззовні).

Гіпертекст — сама форма організації матеріалу і технологія, без якої неможлива його організація. ГТ — накопичення інформації в базі даних (БД), доступ до даних через запити. Зв'язок важливий для користувача, тому в БД-системі немає раніше встановлених зв'язків. Видача інформації — відтворення фрагментів сіток, сформованих до цього моменту в БД. Видаються окремі вузли і частки сітці у графічній формі разом з маршрутами руху. У ГТ-сітки (ГТС) можна легко відображати ідеї учасників проекту, для подальшого вивчення їх в інтерактивному режимі одразу кількома учасниками.

Існує чотири види ГТС (огляд Коніліна):

1. Бібліотечні макросистеми (ширше, ніж пошук літератури) XANADU, Нельсон (система зберігання й актуалізації інформа-

ції), TEXTNET (принцип динамічного упорядкування, перелік вузлів).

2. Засоби дослідження проблем: IBIS, Риштель (аспекти, позиції, аргументи); JOG (вивчення довідкової енциклопедичної літератури, введення сторінок).

3. Системи для перегляду БД (подібні бібліотечним, але менші, і служать для отримання довідкових даних) Browsing Systems, Web-середовище потрібне для опису системи перегляду, легкість доступу, додавати нову інформацію не дозволено.

4. Системи широкого призначення (експериментування в різних напрямках): INTERMEDIA фірми XEROX.

В архітектурі гіпертекстової системи розрізняється БД і систему управління. Система управління ГТ включає два інструментальних комплекси. Один використовується як інструмент інтерактивного управління процесами браузеру (перегляд крок за кроком у напрямі до інформації, що шукається), а другий — у процесах створення і ведення ГТ. Першим браузером, який отримав широку славу, стала програма Mosaic, розроблена Марком Андресеном у 1993 р. У 1994 р. на основі цієї програми був створений один із самих популярних браузерів — Netscape Navigator. Пізніше — Internet Explorer.

Найбільш популярний програмний пакет для створення і ведення ГТ нині — програмні пакети HyperCard компанії Apple. ГТ у формі HyperCard подається у вигляді каталожних карток. Користувач за допомогою досить простого інтерфейсу організує структуру картки — зв'язки між ними. Картки можуть мати різні види інформації.

### **Сфери використання ГТ**

Бізнес- і професійні комунікації:

- інтерактивні гіпермедіа-презентації для маркетингу і збуту;
- бізнес-звіти;
- розробка планів і пропозицій;
- керівництво з корпоративної політики і процедур;
- керівництво з експлуатації;
- юридична документація;
- медична інформація.

Навчання і науково-дослідницька діяльність:

- навчальні посібники;
- навчальні завдання;
- довідники, енциклопедії, словники, бібліотеки, підручники тощо;
- складання рефератів, звітів.

Інше:

- імітація і моделювання;
- всевітня павутина (World Wide Web);
- проектування інтерфейсів;
- організаційне управління;
- мистецтво і культура.

### **Питання для самоконтролю**

1. Значення морфем СМ.
2. Значення словосполучень і речень.
3. Види пам'яті.
4. Що таке шаблони?
5. З чого складаються шаблони?
6. Критерії поділу на типи ГТ.
7. З чого складаються гіпертекстові системи?
8. Види ГТС.
9. Де використовуються ГТ?

*Література* [7; 8; 17; 19; 26; 31; 34; 40; 43; 65]

### **Тема 6. Теорія фреймів**

Теорія фреймів (ТФ) — це парадигма для представлення знань з метою використання цих знань комп'ютером. Вперше була представлена Мінським як спроба побудувати фреймову сітку, чи парадигму з метою досягнення більшого ефекту розуміння. З одного боку Мінський намагався сконструювати базу даних (БД), яка містить енциклопедичні знання, але з іншого боку, він хотів створити найбільш ілюструючу базу, яка має інформацію у структурованій і впорядкованій формі. Ця структура дала змогу комп'ютеру вводити інформацію в гнучкішій формі, маючи доступ до того розділу, який потрібен у цей момент. Мінський розробив таку схему, в якій інформація утримується в спеціальних осередках, які зветься “фреймами”.

Вони об'єднані у сітку, яка зветься системою фреймів. Новий фрейм активізується з появою нової ситуації. Відмінною

рисою є те, що він одночасно має великий обсяг знань і водночас є достатньо гнучким для того, щоб бути використаним як окремий елемент БД.

Термін “фрейм” був найпопулярнішим у середині 70-х років, коли існувало багато його тлумачень, відмінних від інтерпретації Мінського. Щоб краще зрозуміти цю теорію, розглянемо один з прикладів Мінського, який ґрунтується на зв’язку між очікуванням, самопочуттям і відчуттям людини, коли вона відкриває двері і заходить до кімнати. Уявімо, що ви збираєтесь відкрити двері й зайти в кімнату незнайомого вам дому. Знаходячись у домі, перед тим, як відкрити двері, у вас є певні уявлення про те, що ви побачите, коли увійдете в кімнату. Наприклад, якщо ви побачите пейзаж чи морський берег, спочатку вам буде важко розпізнати їх. Потім ви будете здивовані, і кінець кінцем дезорієнтовані, оскільки ви не зможете пояснити інформацію, що надійшла, і пов’язати її з тими уявленнями, які були у вас до того. У вас виникнуть ускладнення з тим, щоб передбачити подальший хід подій. З аналітичної точки зору це можна пояснити як активізацію фрейму “кімнати” у момент відкривання дверей і його провідну роль в інтерпретації інформації, що надходить. Якби ви побачили за дверима ліжку, то фрейм “кімнати” отримав би більш вузьку форму і перетворився б у фрейм “ліжка”.

Іншими словами, ви б мали доступ до найбільш специфічного фрейму з усіх доступних. Можливо, що ви використаєте інформацію, яка є у вашому фреймі кімнати для того, щоб розпізнати меблі, що зветься процесом зверху-вниз, чи в контексті теорії фреймів фрейморухоючим розпізнаванням. Якби ви побачили пожежний гідрант, то ваші почуття були б аналогічні першому випадку. Психологи підмітили, що розпізнавання об’єктів легше проходить у звичайному контексті, ніж у нестандартній ситуації. З цього прикладу ми бачимо, що фрейм — це модель знань, яка активізується в певній ситуації і служить для її пояснення і передбачення. У Мінського є досить непевні ідеї щодо самої структури такої БД, яка могла б виконувати подібні речі. Він запропонував систему, яка складається з пов’язаних між собою фреймів, багато з яких складається з однакових підкомпонентів, об’єднаних у сітку. Таким чином, у випадку, коли хтось заходить в дім, його очікування контролюються операціями, які входять у сітку системи фреймів.

У розглянутому випадку ми маємо справу з фреймовою системою для дому, і з підсистемами для дверей і кімнати. Активізовані фрейми з додатковою інформацією в БД про те, що ви відкриваєте двері, будуть служити переходом від активізованого фрейму дверей до фрейму кімнати. При цьому фрейми дверей і кімнати матимуть однакову підструктуру. Мінський назвав це явище розділом терміналів і вважав його важливою частиною теорії фреймів.

Мінський також ввів термінологію, яка могла б використовуватися при вивченні цієї теорії (фрейми, слоти, термінали, тощо). Хоча приклади цієї теорії були розділені на мовні і перцептуальні, а Мінський розглядав їх як такі, що мають спільну природу, у мові є ширша сфера її використання. В основному більшість досліджень було зроблено в контексті загальноновживаної лексики і літературної мови. Як найбільш доступну ілюстрацію розпізнавання, інтерпретації і передбачення можна розглянути дві послідовності речень, які взяті у Шранка і Абельсона. На глобальному рівні послідовність *A* явно відрізняється від *B*:

*A.* John went to a restaurant.

He asked the waitress for a hamburger.

He paid the tip & left.

*B.* John went to a park.

He asked the midget for a mouse.

He picked up the box & left.

Хоча всі ці речення мають однакову синтаксичну структуру і тип семантичної інформації, їх розуміння кардинально відрізняється. Послідовність *A* має доступ до деякого виду структури знань вищого рівня, а *B* не має. Якби *A* не мала такий доступ, то її розуміння зводилося б до рівня *B* і характеризувалося б як дезорієнтування. Цей контраст є наочним прикладом миттєвої роботи вищого рівня структури знань. Була запропонована програма під назвою SAM, яка відповідає на запитання і видає зміст таких розповідей.

Наприклад, SAM може відповісти на такі запитання, на які у тексті немає відповіді, за допомогою доступу до запису подій, що передбачаються, які передували обіді в ресторані:

Did John sit down in the restaurant?

Did John eat the hamburger?



Таким чином, SAM може розпізнати описану ситуацію як обід у ресторані і потім передбачити оптимальний розвиток подій. У нашому випадку розпізнавання не викликало труднощів, але у більшості випадків воно досить непросте і є дуже важливою частиною теорії.

Розглянемо інший приклад:

He plunked down \$5 at the window.  
She tried to give him \$ 2.50, but he wouldn't take it.  
So when they got inside,  
she bought him a large bag of popcorn.

Він цікавий тим, що у більшості людей він викликає цикл повторюваних неправильних чи незакінчених розпізнавань і реінтерпретацій. У випадках з багатозначними словами багатозначність вирішується за допомогою активізованого раніше фрейму. Для цих цілей необхідно створити лексикон до кожного фрейму. Коли фрейм активізується, відповідному лексикону віддається перевага під час пошуку відповідного значення слова. В контексті ТФ — це розпізнавання процесів, які контролюються фреймами, які, в свою чергу, контролюють розпізнавання вхідної інформації. Інколи це називається процесом зверху-вниз фрейморухоючого розпізнавання.

Використання цих процесів знайшло своє відображення в програмі FRAMP, яка може підсумовувати газетні повідомлення і класифікувати їх відповідно до класу подій, наприклад, тероризм чи землетрус. Ця програма зберігає набір об'єктів, які мають бути описані в кожному різновиді текстів, і цей набір допомагає процесу розпізнавання подій, що описуються.

### **Маніпуляція фреймами**

Такі питання, як розмір фрейму чи доступ до нього, пов'язані з організацією пам'яті і не потребують спеціального розгляду. В літературі є багато думок щодо процесів, які стосуються розпізнавання фреймів і доступу до структури знань вищого рівня. Незважаючи на те, що люди можуть розпізнати фрейм без особливих зусиль, для комп'ютера у більшості випадків це досить складне завдання. Тому питання розпізнавання фреймів залишаються відкритими і складними для вирішення за допомогою штучного інтелекту.

## **Розмір фрейму**

Розмір фрейму тісніше пов'язаний з організацією пам'яті, ніж це здається на перший погляд. Це відбувається тому, що в розумінні людини розмір фрейму визначається не тільки семантичним контекстом, але й багатьма іншими факторами. Розглянемо фрейм візиту до лікаря, який складається з підфреймів, одним з яких є кімната очікування. Таким чином ми можемо сказати, що розмір фрейму не залежить від його семантичного змісту і який представлено (такого, як, наприклад, візит до лікаря), але залежить від того, які компоненти описаної інформації у фреймі (такому, як кімната очікування) використовуються в пам'яті. Це означає, що коли певний набір знань використовується пам'яттю більш ніж в одній ситуації, система пам'яті визначає це, потім модифікує цю інформацію у фрейм, і реструктуризує початковий фрейм так, щоб новий фрейм використовувався як його підкомпонент. Перелічені операції також залишаються відкритими питаннями в ТФ.

## **Ініціалізаційні категорії**

Рош запропонував три рівні категорій пред'явлення знань: базову, субординатну і суперординаційну. Наприклад, у сфері меблювання концепція “крісла” є прикладом категорії основного рівня, а концепція “меблі” — це приклад суперординаційної категорії. На мову пред'явлення знань впливає таксономія, яка включає їх як різні типи даних. У сфері людського спілкування категорії основного рівня є первинними категоріями, які впізнає людина, інші ж категорії випливають з них. Тобто, суперординаційна категорія — це узагальнення базової, а субординатна — це підрозділ базової категорії.

Кожний фрейм має свій певний слот. Так, для фрейму дія слот може бути заповнена лише виконавцем цієї дії, а сусідні фрейми можуть наслідувати цей слот. Деякі дослідники пропонували, що випадки граматичних відмінків співпадають зі слотами в ТФ, і ця теорія була названа теорією ідентичності слоту і відмінку. Було запропоновано число таких відмінків — від 8 до 20, але точне число не визначене. Але якщо агентив повністю співпадає зі своїм слотом, то інші відмінки викликали суперечки. До цього часу точно не встановлено, скільки всього існує відмінків.

Також викликав труднощі той факт, що слоти не завжди можуть бути перехідними. Наприклад, відповідно до ТФ можна

сказати, що фрейм — це живий предмет і може мати живий слот, фрейм “людина” може мати слот “чесний”, а фрейм “блоха” не може мати такий слот, і він до нього ніколи не перейде. Іншими словами, зв’язки між слотами в ТФ не досліджені до кінця. Слоти можуть передаватися, можуть бути багатofункціональними, але в той же час не розглядаються як функції.

### **Гібридні системи**

Система фреймів (СФ) інколи адаптується для побудови описів чи визначень. Була створена змішана мова — KRYPTON, яка складалася з фреймових компонентів і компонентів предикатних обчислень, які допомагають робити висновки за допомогою термінів і предикатів. Коли активізується фрейм, факти стають доступними користувачу. Також існує мова Loops, яка поєднує об’єкти, логічне програмування і процедури. Існують також фреймоподібні мови, які за початкову позицію приймають один тип даних в пам’яті, якусь концепцію, а не дві (наприклад, фрейм і слот), і уявлення цієї концепції в пам’яті повинно бути цілісним.

### **Об’єктно-орієнтовані мови**

Паралельно з мовами фреймів існують об’єктно-орієнтовані програмні мови, які використовуються для складання програм, але мають деякі якості мов фреймів, такі, як використання слотів для детальної, досконалої класифікації об’єктів. Відмінність їх від мов фреймів у тому, що фреймові мови спрямовані на більш узагальнене уявлення інформації про об’єкт.

Однією з труднощів уявлення знань і мови фреймів є відсутність формальної семантики. Це ускладнює порівняння якостей пред’явлення знань різних мов фреймів, а також повне логічне пояснення мови фреймів.

### **Питання для самоконтролю**

1. Мета створення теорії фреймів?
2. На чому ґрунтується поняття “фрейм”?
3. Як вирішується проблема з багатозначними словами?
4. Від чого залежить розмір фрейму?
5. Як використовується слот?
6. Яка мета створення об’єктно-орієнтованих мов?

*Література [5; 7; 11; 22; 23; 25–27; 30]*

## *Тема 7. Загальнонауковий метод моделювання і специфіка його використання в лінгвістиці*

Метод моделювання — центральний дослідницький метод в науці. Моделювання в науці — це виявлення якостей будь-якого предмета за допомогою побудови його моделі.

Моделлю можна назвати образ будь-якого об'єкта, який використовується в певних умовах як його замісник (фотографія в паспорті — модель людини).

Якості моделей:

- умовність — образ може бути не лише матеріальним, а й мисленнєвим і передаватися за допомогою знакової системи;
- моделлю може бути не лише образ, а й праобраз оригіналу, модель частіше є гомоморфною до оригіналу (тобто багатьом елементам оригіналу відповідає менша кількість елементів моделі на відміну від ізоморфізму).

Модель у лінгвістиці — це пристрій, який штучно створюється лінгвістом (реальне чи мисленнєве), і відтворює, імітує своєю поведінкою (звичайно у спрощеному вигляді) поведінку оригіналу в лінгвістичних цілях.

Власне є такі лінгвістичні моделі:

- моделі мовленнєвої діяльності, процесуальні моделі (найбільш складні);
- моделі мовної системи, мовної структури (також дуже складні);
- модель пам'яті та ін.

Лінгвістичне моделювання передбачає необхідне використання абстракції та ідеалізації. Відображуючи релевантні суттєві (з точки зору дослідження) якості оригіналу і відволікаючись від несуттєвих, модель виступає як деякий абстрактний ідеалізований об'єкт. Будь-яка модель будується на основі гіпотези про можливий устрій оригіналу і є його функціональним аналогом, що дає можливість переносити знання з моделі на оригінал. Критерієм адекватності моделі є експеримент.

В ідеалі модель повинна бути формальною (тобто в ній повинні бути і однозначно задані вихідні об'єкти, зв'язки, які їх пов'язують, і правила поводження з ними) і володіти пояснювальною силою (тобто не лише пояснювати факти чи дані експериментів, які не пояснює вже існуюча теорія, але й передбачати

невідому раніше, хоча й принципово можливу поведінку оригіналу, яка пізніше має підтверджуватися даними спостережень чи експериментів).

Поняття лінгвістичної моделі виникло у структурній лінгвістиці, але увійшло в науковий вжиток у 60–70 роки ХХ ст. з виникненням математичної лінгвістики і проникненням в лінгвістику математичних методів.

Зміст терміна “модель” в сучасній лінгвістиці значною мірою охоплювалось раніше терміном “теорія” (особливо Ельмслевим).

Вважається, що найменування “модель” заслуговує лише та теорія, яка достатньо експліцитно викладена і в достатньому ступені формалізована (в ідеалі кожна модель повинна допускати реалізацію на ЕОМ).

Конструювання моделі — не лише один із засобів відображення мовних явищ, а й об’єктивний практичний критерій перевірки істинності знань про мову. Разом з іншими методами вивчення мови моделювання виступає засобом поглиблення пізнання прихованих механізмів мовленнєвої діяльності, його руху від відносно примітивних до більш змістовних моделей, які повніше розкривають сутність мови.

У середині мови як системи існує принцип моделювання: одні його підсистеми моделюють інші, наприклад, система писемного мовлення є моделлю усного мовлення; в середині писемного мовлення ми маємо справу з кількома моделями (друкованою, рукописною); план вираження є моделлю плану змісту.

Метод моделювання звичайно спирається на знакові системи, але мова — сама знакова система, тобто слова ми моделюємо за допомогою слів.

Головна ціль моделювання в лінгвістиці — це моделювання цілісної мовної здатності людини.

### **Питання для самоконтролю**

1. Що таке модель і моделювання в лінгвістиці?
2. Які моделі?
3. Чому лінгвістичне моделювання використовує абстракцію й ідеалізацію?
4. Що є дійсним критерієм перевірки знань?

*Література* [9; 10; 12; 14; 16; 18; 22; 27; 35; 37; 39; 56; 63]

## *Тема 8. Методи ПЛ*

Різноманіття методів ПЛ можна порівняти з різноманіттям конкретних галузей використання знань про мову. Примітка: квантитативна лінгвістика спирається на методичний інструментарій статистики, комп'ютерна лінгвістика — на методи програмування і пред'явлення знань, теорія впливу — на уявлення про значуще варіювання мовних структур. Різноманіття методів добре видно під час порівняння методів описової, теоретичної ПЛ.

1. Завдання описової лінгвістики — опис фактів мови різних рівнів. На першому плані — метод класифікації, тобто виявлення сітки параметрів, яка дає можливість охопити всі релевантні (в теорії) якості мовних структур.

2. Теоретична лінгвістика формує саме уявлення про релевантність в концептуальних моделях мови. Концептуальні моделі функціонування мови не просто описують, а пояснюють факти, які спостерігаються, і передбачають, наприклад, умови їх появи. Моделювання в теоретичній лінгвістиці у своїх суттєвих рисах відповідає ідеї моделювання у сфері природничих наук.

Метод моделювання використовується в тих випадках, коли безпосереднє вивчення об'єкта моделювання в тому чи іншому відношенні ускладнене — наприклад, об'єкт руйнується під час безпосереднього контакту або сам контакт ускладнений. У цьому випадку об'єкт моделювання є для дослідника своєрідним “чорним ящиком”. (Примітка: мислення людини і мова в усіх його проявах.) Так само для створення моделі чого-небудь необхідно знати, як веде себе об'єкт моделювання в тих чи інших умовах. (Примітка: створення моделі синтаксису природної мови спирається на факти можливості/неможливості здійснення тих чи інших синтаксичних трансформацій.)

Обмеження методу моделювання за Апресяном: “моделювати можна тільки ті якості об'єкта, які не визначаються його фізичною природою”. Іншими словами, успішно моделюються тільки ті характеристики, які пов'язані зі структурною організацією об'єкта.

У теоретичній лінгвістиці часто використовуються такі типи моделей:

- компонентні моделі чи моделі структури (з чого зроблено  $X$ );

- передбачаючі моделі (передбачити поведінку  $X$  у тих чи інших обставинах);
- імітуючі моделі (зовнішньо вести себе як  $X$ );
- діахронічні моделі (як і чому змінюється  $X$  з часом).

ПЛ також використовує і метод класифікації і метод моделювання. Однак, оскільки задачі ПЛ зосереджені в області оптимізації функцій мови, а оптимізація визначається конкретною задачею, то у прикладках мовознавства використовується пізнавальна установка, відома під трохи іронічною назвою “Фокус-покус” (більш шановано можна було б сказати “Інженерний підхід”). Ця установка як основну цінність висуває не пізнання того, “як все є насправді”, а рішення конкретної задачі, в окремому випадку — задовольнити вимоги “замовника”, який має свої власні цілі, часто дуже далекі від канонів власне лінгвістичних досліджень. Це не означає, що результати прикладних досліджень не мають ніякої цінності для теорії мови.

Таким чином, важливою якістю методів ПЛ є оптимізація — опис (модель) проблемної області, при якому ця область зберігає в результуючій уяві тільки ті суттєві якості, які необхідні для певної практичної задачі.

Приклад: категорія часу.

Описова Л — опис граматичної категорії часу (виділення грамам, морфологічних способів виразу грамам, поєднання грамам категорії часу з грамамами інших граматичних категорій), класифікація лексики із значенням часових відношень, класифікація синтаксичних конструкцій.

Теоретична Л (в рамках рівневої моделі мови) — семантика часових відношень, способи вираження на синтаксичному рівні; способи вираження на лексичному рівні; способи вираження на морфологічному рівні.

Прикладний опис буде виглядати зовсім по-іншому:

- складання технічного завдання (визначається замовником);
- аналіз проблемної області (скільки типів часових відношень представлено в проблемній області і які формальні способи вираження темпоральних відношень у цій мові);
- формування метамов, способів опису проблемної області, які сумісні з іншими метамовами, що залучаються;

- використання метамов дає результуюче представлення (модель) проблемної області;
- перевірка результуючого представлення (пояснююча і передбачаюча сила моделі; комп'ютерна реалізація чи експеримент).

Прикладні моделі в цілому орієнтовані на конкретні підмови, а не на всю мову в цілому; вони потребують більшого ступеня формалізації; використовують знання про мову вибірково; прикладні моделі не роблять різниці між власне лінгвістичними і екстралінгвістичними аспектами семантики мовних виразів; у значному ступені роблять грубішим об'єкт, що моделюється, ніж теоретичні моделі, не покладають ніяких сутнісних обмежень на інструмент моделювання.

### **Інші методи**

*Теоретико-множинний метод*, головний елемент якого — лінгвістична матриця. (Зображення на площині співвідношення будь-яких множинних об'єктів.) Примітка: відмінювання і відмінки.

Причини для представлення у формальному вигляді — категорії дистрибуції (синтагматичної) і опозиції (парадигматичної).

*Опозиційний метод*. Відношення між множинами. Два типи опозиції: 1) привативні (між наявністю чи відсутністю якогось елементу). Примітка: 1) дві фонемі з різною протяжністю не можуть знаходитися в одній позиції; 2) градуальні (різний ступінь вираження однієї ознаки). Примітка: голосні високого/низького підйому.

*Дистрибутивний метод*. Дистрибуція — сукупність оточення, де певна мовна одиниця зустрічається в мовленні. Дистрибутивна структура — це класифікація мовних одиниць, яка послідовно встановлюється на основі їх розміщення відносно інших у потоці мовлення. Типи дистрибуції: 1) додаткові: члени не пересікаються, знаходяться у своєму оточенні (примітка: алофони); 2) контрастна (примітка: алофони різних фонем так-ток); 3) еквівалентна: дистрибуція вільного варіювання (примітка: одне й те саме слово, вимовлене в Москві і у Сибіру); 4) часткової еквівалентності (примітка: Бог, але “г” не всюди приглушається до “х”).

*Функтивний метод*. Розглядається через категорію функтора — мовний засіб транспозиції однієї множини мовних одиниць в іншу множину тієї самої мови. Функторення — це якість, чи



знання функтора (примітка: повідомляти + морфемне функторення = повідомлення). Три типи функторів: мовна одиниця (морфема, слово, словосполучення), синтагма (мінімальна модель синтаксичної побудови), операція (послідовне використання операцій) (примітка: запрошення батька — 2 значення).

Методи сучасної ПЛ за Степановим: пізнання; викладення; моделювання; алгебраїчний (метод множин); репрезентативний (семіотичний метод) — відношення між прихованим боком знаку і боком, що спостерігається (примітка: речення і його структурна схема).

### **Питання для самоконтролю**

1. Яке завдання постає перед описовою лінгвістикою?
2. Які типи моделей використовуються в теоретичній лінгвістиці?
3. Що дає оптимізація методам ПЛ?
4. Чим відрізняються теоретична і описова лінгвістика?
5. На що орієнтовані прикладні моделі?
6. Які є типи опозиції в опозиційному методі?
7. На чому ґрунтується дистрибутивний метод?
8. Типи функторів.

*Література [3; 5; 7; 10; 11; 16; 43; 29; 37]*

### ***Тема 9. Комп'ютерна лінгвістика: моделювання спілкування і структури сюжету***

Комп'ютерна лінгвістика (КЛ) зазвичай — широка область використання комп'ютерних інструментів — програм, комп'ютерних технологій організації і обробки даних — для моделювання функціонування мови в тих чи інших умовах, ситуаціях, проблемних областях, а також у сфері використання комп'ютерних моделей мови не лише в лінгвістиці, а й суміжних з нею дисциплінах. Комп'ютерне моделювання мови може розглядатися і як сфера додатків теорії програмування в області лінгвістики. Термін КЛ задає загальну орієнтацію на використання комп'ютерів для вирішення різних наукових і практичних задач, пов'язаних з мовою.

Як окремий науковий напрям КЛ сформувалася у 60-ті роки ХХ ст.

Комп'ютерні програми, які моделюють ті чи інші аспекти функціонування мови, можуть використовувати різні засоби

програмування, таким чином про спільну метамову говорити не доводиться. Але! Це не зовсім так. Існують спільні принципи комп'ютерного моделювання мислення, яке так чи інакше реалізується в будь-якій комп'ютерній моделі. В основі цієї мови лежить теорія знань, розроблена в АІ.

В основному тезисі теорії знань сказано, що мислення — це процес обробки і породження знань. Знання вважаються невідзначеною категорією. Як “процесор”, що обробляє знання, виступає когнітивна система людини.

Два основних види знань: декларативні та процедурні. *Декларативні знання* подаються зазвичай у вигляді сукупності пропозицій (тверджень про що-небудь) (наприклад: тлумачення слів у звичайних тлумачних словниках). Декларативні знання піддаються процедурі верифікації у термінах І/Л. *Процедурні знання* являють собою послідовність (список) операцій, дій, які слід виконати. Це спільна інструкція про дії у певній ситуації (наприклад: інструкції щодо користування побутовими приладами). Процедурні знання неможливо верифіціювати як І/Л, їх можна оцінювати лише за успішністю/неуспішністю алгоритму.

Елементи метамови мають онтологічний та інструментальний аспект — вони одночасно означають деякі реальні сутності когнітивної системи людини і способи представлення цих сутностей деякими метамовами. Онтологічно розділ декларативних і процедурних знань відповідає різним типам знань когнітивної системи людини. Інструментально знання (як онтологічно процедурне, так і декларативне) можна представити як сукупність дескрипцій, описів і як алгоритм-інструкцію. Таким чином, онтологічно декларативне знання про об'єкт дійсності можна представити процедурно як сукупність інструкцій, алгоритмів для його створення, складання чи як алгоритм його типового використання. Правильним буде і зворотне: онтологічне процедурне знання можна представити декларативно.

Дослідники сходяться в тому, що всяке декларативне знання можна представити процедурно, хоча це може виявитися для когнітивної системи дуже неекономним. Декларативне знання суттєво експліцитніше, воно легше, ніж процедурне, усвідомлюється людиною. Таким чином виникає питання про можливість декларативного представлення процедурного знання. Але екс-

плікація імпліцитного процедурного знання часто призводить до його руйнування.

У теорії знань для їх вивчення використовуються різні структури знань — фрейми, сценарії, плани.

Фрейм — це концептуальна структура для декларативного пред'явлення знань про типізовану тематично єдину ситуацію, яка має слоти, пов'язані між собою певними семантичними зв'язками. Фрейм часто подають у вигляді таблиці, рядки якої утворюють слоти. Кожний слот має своє ім'я і зміст. Фрейм може містити вкладені підфрейми і посилання на інші фрейми. Замість таблиці часто використовується предикатна форма подання: фрейм має форму предиката чи функції з аргументами. Існують й інші способи подання фрейму. Наприклад, він може бути у вигляді кортежу такого виду: {<ім'я фрейму <ім'я слота<sub>1</sub>> <значення слота<sub>1</sub>>}, {<СПІВРОБІТНИК> <ПРИЗВИЩЕ> <Петров>, <ПОСАДА> <технік, <ВІК> <32>...}. Зазвичай такий вигляд мають фрейми в мовах представлень знань. Як й інші когнітивні категорії КЛ, поняття фрейму омонімічне: це частина когнітивної системи людини, тобто можна співставити з такими поняттями як гештальт, прототип, стереотип, схема. Інструментальний бік фрейму — структура для декларативного представлення знань.

Тісно пов'язано з категорією фрейму поняття сцени. В граматиці сюжетів сцена постає як частина епізоду чи розповіді. В АІ сцени використовуються в схемах розпізнавання образів, а також у програмах, орієнтованих на дослідження (аналіз, опис) проблемних ситуацій.

Сценарій являє собою концептуальну структуру для процедурного представлення знань про стереотипну ситуацію чи стереотипну поведінку. Елементами сценарію є кроки алгоритму чи інструкції. Сценарій можна подати не лише у вигляді алгоритму, а й у вигляді сітки, вершинам якої відповідають деякі ситуації, а дугам — зв'язки між ситуаціями. Деякі дослідники залучають для комп'ютерного моделювання інтелекту категорію скрипту. Скрипт — це певна загальноприйнятна послідовність причинних зв'язків.

В результаті прикладання сценарію до конкретної проблемної ситуації формується план. Він використовується для процедурного представлення знань про можливі дії, які ведуть до досяг-

нення певної мети. План співвідносить мету з послідовністю дій. В схемах АІ план виникає в результаті плануючої діяльності відповідного модуля — модуля планування. На відміну від сценарію, план пов'язаний з конкретною ситуацією, конкретним виконавцем і переслідує досягнення певної мети.

Ще одне важливе поняття — модель світу. Вона є сукупністю певним чином організованих знань про світ, які притаманні когнітивній системі чи її комп'ютерній моделі. У загальнішому вигляді — частина когнітивної системи, яка зберігає знання про устрій світу, його закономірності і т. д. У системах АІ модель світу є особливим блоком, до якого, залежно від вибраної архітектури, можуть входити загальні знання про світ (у вигляді простих пропозицій), деякі специфічні факти (найвища вершина в світі — Еверест), а також цінності та їх ієрархії. У різних формалізмах модель світу може мати різну структуру: предикатну, сітьову, фреймову та ін.

Деякі напрями КЛ-області, які безпосередньо пов'язані з оптимізацією когнітивної функції мови. Розглянемо дві сфери комп'ютерного моделювання, в яких використовуються знання про функціонування мовної системи: моделювання спілкування, структури сюжету.

### **Моделювання спілкування**

Проблематика КЛ часто пов'язується з моделюванням спілкування, зокрема, із забезпеченням спілкування людини з ЕОМ природною чи обмеженою природною мовою. Комп'ютерні моделі спілкування часто використовуються для вивчення самого процесу спілкування.

Вивчення існуючого досвіду експлуатації комп'ютерних систем дозволило дослідникам по-новому подивитися на функції і структуру комунікації. Розглянемо такі питання: Що забезпечує природність спілкування? Які умови зв'язності бесіди? Коли спілкування виявляється успішним? В яких випадках виникають комунікативні невдачі і чи можна їх уникнути? Які стратегії спілкування використовують учасники комп'ютерної взаємодії при досягненні своїх комунікативних цілей?

Однією із найцікавіших комп'ютерних моделей діалогу, яка викликала значні теоретичні дискусії, була програма Джозефа Вейценбаума “Еліза” (перший варіант 1966 р.), програма-

імітатор, метою якої було не моделювання мислення в точному смислі, а моделювання мовленнєвої поведінки — обмежені програмні ресурси, лінгвістичний аналіз і синтез були зведені до мінімуму.

Програма “Еліза” була використана групою дослідників на чолі з М. Макгайром для вивчення структури діалогу й особливостей електронно-мовленнєвої комунікації. В експерименті з “Елізою” розмовляли протягом години 24 досліджуваних. 15 учасників (62 %) були впевнені, що їм відповідала людина, 5 досліджуваних (21 %) виявили певну невпевненість і лише четверо досліджуваних (17 %) були абсолютно впевнені, що спілкувалися з ЕОМ.

З лінгвістичної точки зору алгоритми програми “Елізи” містять мінімум лінгвістичної інформації. По-перше, це комплекс ключових слів, які актуалізують деякі стійкі комунікативні форми (шаблони), по-друге, здатність відносно нескладно трансформувати попереднє висловлювання (наприклад, програма має можливість переробити попереднє твердження в загальне питання).

**Висновки.** 1. Прояв найважливішої особливості комунікації на електронній мові: електронно-мовленнєвий дискурс дуже терплячий відносно збоїв і помилок — він надлишковий і стійкий до перешкод. Таким чином, людина схильна надавати сенс тому, що часто його не має (досліджувані самі надають сенс діалогові, самі забезпечують його зв’язність, самі приписують партнеру комунікативні інтенції).

2. Досліджувані намагалися встановити, хто перед ними — комп’ютер чи людина (тобто рольові характеристики партнера) як можна раніше: 22 учасники з 24 з’ясували для себе ситуацію не більше, ніж за 5 реплік. Визначення ролей в комунікації належить до метарівня спілкування, оскільки це складає одну з передумов успішної комунікації, що зберігає спілкування від багаточисленних комунікативних невдач. Цю якість електронно-мовленнєвої комунікації можна назвати принципом пріоритету метакомунікативних параметрів ситуації спілкування.

3. Існують різні типи комунікативної взаємодії між людьми. Успішна взаємодія між людиною і програмою типу “Еліза” можлива лише в ситуації, коли відбувається так зване “асоціативне спілкування”, при якому репліки діалогу пов’язані не стільки логічними зв’язками типу “причина — наслідок”, “посилка —

висновок”, а асоціаціями. Програма “Еліза” не володіла асоціативною базою, а одна з типових стратегій “уникнення нерозуміння”, реалізована у програмі “зміна теми бесіди”.

4. Неуніверсальність правил комунікативної взаємодії. Кожний тип комунікації обслуговується своїм набором за простими правилами, які забезпечують зв'язність дискурсу, його усвідомлення для учасників. Із експериментів М. Макгайра з програмою “Еліза” випливає, що крім асоціативного способу спілкування, виділяється ще “рішення задач”, “задавання питань” і “уточнення розуміння”.

#### **Моделювання структури сюжету**

Комп'ютерне моделювання структури сюжету — перспективний напрям КЛ. Вивчення структури сюжету належить до проблематики літературознавства (в широкому сенсі), семіотики і культурології. Комп'ютерні програми моделювання сюжету базуються на трьох базових формалізмах представлення сюжету — морфологічному і синтаксичному напрямках, а також на когнітивному підході.

Ідеї морфологічного устрою структури сюжету мають місце у відомих працях В. Я. Проппа “Про руську чарівну казку”. Дослідник спостерігав, що при численності персонажів і подій “чарівної казки” кількість функцій персонажів обмежена, і запропонував апарат для опису цих функцій. Ідеї Проппа лягли в основу комп'ютерної програми TALE, яка моделює породження сюжету казки. В основу алгоритму програми TALE покладено послідовність функцій персонажів казки. Фактично функції Проппа створювали чимало типізованих ситуацій, упорядкованих на основі аналізу емпіричного матеріалу. Можливості поєднання різних ситуацій у правилах виявились типовою послідовністю функцій у тому вигляді, в якому це вдається встановити з текстів казок. У програмі типові послідовності функцій описувались як типові сценарії зустрічей персонажів.

У подальшому система була ускладнена за рахунок введення моделі світу казки, географія якого складається зі звичайного, проміжного (середнього) й іншого світу. Сам світ складається з локусів, пов'язаних між собою певними зв'язками. Стосунки поєднують не лише локуси всередині світу, але й локуси різних світів.

*Звичайний світ* складається з таких локусів: місце проживання героя, місце отримання завдання, місце дарування чарівних предметів, які допомагають виконати завдання.

*Середній світ* відокремлює світ героїв від світу антигероїв.

*Інший світ* включає місце проживання антигероя, місце битви між героєм і антигероєм, і локус-місце нагороди чи мети, якої прагне досягти герой. Локуси пов'язані зв'язками переходу, які представляють можливі послідовності розгорання сюжету.

Суто “морфологічного” підходу до структури сюжету казки недостатньо. “Морфеми” казкового сюжету повинні не лише певним чином поєднуватися між собою, а й мати специфічні обмеження на сполучення. Фіксація одного типового порядку послідовності функцій персонажів казки суттєво обмежує наявні можливості сполучення. Адекватніше вирішення цієї проблеми дає синтаксичний підхід до структури сюжету.

Теоретичну основу синтаксичного підходу до сюжету тексту склали “сюжетні граматики”, чи “граматики розповіді” (*story grammats*). Вони з'явились у середині 70-х років у результаті переносу ідей породжуючої граматики Хомського на опис макроструктури тексту. Якщо найважливішими складниками синтаксичної структури в породжуючій граматиці були дієслівні та іменні групи, то в більшості сюжетних граматик як базові виділялись експозиція, подія та епізод. В теорії сюжетних граматик широко обговорювались умови обмеження, які визначили статус послідовності. Але виявилось, що лінгвістичними методами це зробити неможливо. Багато обмежень мають соціокультурний характер. Сюжетні граматики суттєво розрізняються набором категорій у дереві породження, допускали досить обмежений набір правил модифікації розповідної структури: в переважній більшості випадків ці правила були взяті з тієї ж породжуючої граматики. Потенціал варіювання структури сюжету забезпечується в першу чергу трансформаціями пересування і опущення.

Використання сюжетних граматик у комп'ютерному моделюванні виявилось невдалим. Синтаксичний компонент сюжету, який описується граматиками, відображає зовнішні особливості тексту. Спроба використати граматики сюжетів для породження сюжету призводить до того, що породжуються тексти, які не відповідають інтуїтивному уявленню про розповідь. Основний висновок дискусії про недоліки сюжетних граматик звівся до необ-

хідності опису сюжету із залученням категорій “мета”, “проблема”, “план” тощо. Це призводить до метамови, яка враховує лише зовнішні особливості сюжету, чого явно недостатньо. Необхідно звернутися до когнітивних станів персонажів.

**Когнітивний підхід до сюжету.** На початку 80-х років В. Ленерт у працях зі створення комп’ютерного генератора сюжетів запропонувала оригінальний формалізм емоційних сюжетних одиниць, який виявився сильним засобом представлення структури сюжету. При тому, що він був насамперед розроблений для системи АІ, цей формалізм використовувався в чисто теоретичних дослідженнях. Сутність підходу Ленерт полягала в тому, що сюжет описувався як послідовна зміна когнітивно-емоційних станів персонажів. Отже в центрі уваги формалізму Ленерт стоять не зовнішні компоненти сюжету — експозиція, подія, епізод, мораль, — а його змістові характеристики. В цьому відношенні формалізм Ленерт частково виявляється поверненням до ідей Проппа.

Кожна афективна сюжетна одиниця являє собою бінарне співвідношення між деякими подіями, які оцінюють персонажі, і когнітивно-емоційними станами персонажів (у різних комбінаціях — подія і стан; подія і подія і т. д.). Всього виокремлюють 5 типів бінарних відношень: мотивація, актуалізація, припинення першої дії іншою, еквівалентність, а також каузальний зв’язок між персонажами. Кожна сюжетна одиниця отримує назву, наприклад: УСПІХ, НЕВДАЧА, НАПОЛЕГЛЕВІСТЬ, ПРОБЛЕМА тощо.

В. Ленерт брала участь у проєкті, в якому формалізм, призначений для машинної реалізації, використовується для виявлення сюжетних сувертків євангельського тексту, пояснюючих, чому тексти Євангелія здійснили величезний вплив на мільйони людей і на їхню культуру. Вдалося виявити кілька таких сюжетних мотивів, які повторюються у фольклорі й міфах, і займають центральне місце в сюжетній структурі Євангелій (наприклад, “повстання проти несправедливої влади та успіх повстання, який інтерпретується як таємне блаженство, яке досягається ціною власної смерті”).

### **Питання для самоконтролю**

1. Як розглядається комп’ютерне моделювання мови?
2. Зазначте відмінності між двома основними видами знань.



3. Розкажіть про онтологічний та інструментальний аспекти.
4. Які структури знань використовуються в теорії знань для вивчення?
5. Як представляють фрейми?
6. Що є елементами сценарію?
7. Як відбувається моделювання спілкування?
8. Як відбувається моделювання структури сюжету?
9. Чому використання сюжетних граматики у комп'ютерному моделюванні виявилось невдалим?
10. Що лежить в основі когнітивного підходу до сюжету?

*Література [15; 27; 33; 51; 54; 55]*

### ***Тема 10. Квантитативна лінгвістика: дешифровка тексту, авторизація тексту***

#### **Статистична обробка експериментальних даних**

Сучасна лінгвістика стає більш експериментальною. Вимога коректності використання статистичних процедур змушує відкидати частину матеріалу, спотворюючи структуру об'єкта. Статистика вивчає варіюючу ознаку, наприклад, довжину речення: одне, два, три і більше слів у реченні — градації ознаки “довжина”. Лінгвістичні ознаки можуть бути:

- кількісні, тобто співставляються з певним мірилом: кількість морфем у слові, висота основного тону);
- якісні: наприклад, актив/пасив залогу; ударність (можна задати 3 градації — наголошений, ненаголошений, напівнаголошений), можна дослідити одну характеристику — частоту зустрічання градацій цих ознак;
- розгалужені — з ними виникає проблема: наприклад, ознака “ритмічна структура слова” — розгалужена ознака, тому що передусім це довжина слова, а потім — місце наголошеного складу.

Якщо ми хочемо за спостереженнями над вибіркою зробити певний висновок про генеральну сукупність, то вибірка повинна досить чітко представляти генеральну сукупність. У випадку кількісної вибірки, вона має зробити висновок про генеральні параметри з заданою точністю. Якісна достовірність вибірки встановлюється, коли структура вибірки повторює структуру генеральної сукупності. Коли структура відома, точніше відомі фактори, які впливають на параметр що вивчається, вибірку

будують відповідно до цих розподілень. Якщо генеральна структура вибірки невідома, то вибірку будують за випадковим законом, тобто за допомогою будь-якого датчика випадкових чисел.

Побудова розподілення необхідна для виявлення його форми, наприклад, визначення типу мови при дешифровці. Генеральна сукупність у статистиці — множина об'єктів, які організовані не просто за ознакою, що вивчається, а й обов'язково за однорідною ознакою.

*Дешифровка (Д)* — це дослідження повідомлень чи текстів для виявлення інформації, яка представлена способом, не відомим досліднику. Об'єктами лінгвістичної Д є тексти мовами будь-яких народів чи їхні записи у знакових системах: ноти, картографічні документи, математичні формули. Але ні вивчення нуклеотидних ланцюгів, ні медична діагностика, ні робота детектива не є лінгвістичною Д. До лінгвістичних Д належать дослідження історичних письмен і читання тайнопису (криптограм). У першому випадку незрозумілість тексту для дослідника зумовлена розривом традиції, у другому — спеціальним зменшенням автором чи шифрувальником кола адресатів. У текстах, що записані шифрами і кодами, елементи замінені іншими, раніше обумовленими, які зберігають природні закономірності побудови.

Як “означення” *шифр* має чи номінанту, звук чи елемент писемності даної мови. Наприклад, найпростішим шифром є означення кожної букви алфавіту порядковим номером. Позначене в цьому шифрі — двозначне число номінант — буква українського алфавіту, порядковий номер якої позначений цим числом: а-01, б-02. На відміну від шифру *код* має позначене в прямому сенсі, тобто являє собою умовне позначене для існуючого в мові значення, що кодується. Прикладом може бути “офенська мова”, жаргон бродячих торговців центральної Росії ХІХ–ХХ ст.: “Обтыривай маз! Дулец-то яманный!” — говорив коробейник своєму напарнику і той розумів: “Обвішуй, приятель! Мужик простуватий!” Цей код розвинувся як соціальний діалект російської мови. Кодуванню, заміні піддалися лише ті елементи, які мають лексичне значення, а граматика і слово, утворюючи морфеми, залишилися російські.

Задачі, які вирішує криптографія і дешифровка історичних систем письма, майже протилежні, і методи першої можуть досить обмежено використовуватися у другій. У *криптографії* передбачається, що початковий, незакодований текст написаний відомою мовою, причому зазвичай ця мова або добре відома, або коло кандидатів на цю роль з числа добре відомих мов досить невеликий. Мета криптографії полягає у відтворенні початкового тексту по зашифрованому тексту, в якому різні статичні розподілення елементів тексту можуть дуже сильно відрізнятися від відповідних розподілень елементів в початковому тексті.

Під час *дешифровки* історичних систем письма невідома мова, якою написаний текст, але сам текст не був спеціально оброблений з метою ускладнити читання тексту; текст записаний відповідно до норм орфографії даної мови. В кінці 50-х років спеціалісти в області дешифровки історичних систем письма стали надавати велике значення машинній обробці текстів, що вивчалися — “машинна дешифровка”. Використання обчислювальної техніки під час дешифровки дає можливість здійснити дуже громіздку обробку: розбивка нерозчленованого тексту на блоки, співставлення прямих і зворотних словників, вияв формальної граматики. Невідомий текст вводиться в обчислювальну машину в цифровій транскрипції, яка складається вручну.

Основним методом вивчення невідомих текстів вважають *позиційну статистику*. Звичайно при дешифровці широко використовуються такі прості статистичні методи, як співставлення частот знаків алфавіту, частотні словники і т. д., тобто такі прийоми, які дають частотні характеристики знаків, морфем, слівформ у тексті, не пов’язуючи частотні характеристики зі структурою тексту. Під *позиційною статистикою* тексту розуміють частотні характеристики, які пов’язують одиниці тексту з деякою структурою тексту, з положенням, яке займають ці одиниці в тексті. Такого роду частотні характеристики дають змогу визначити закономірності, що існують між різними елементами тексту. Наприклад, на основі частот знаків тексту залежно від положення можна виявити характерні позиції знака в середині слова (ініціальна, медіальна, фінальна) і намітити підходи для виявлення морфології. Можна аналогічним чином дослідити сполучення морфем між собою, визначити закономірності поєднання слів у середині речення. Це можливо тому, що текст роз-

глядається як структура зв'язків, діючих для близьких елементів. Для текстів малих обсягів позиційна статистика стає єдиним способом отримання надійної інформації про структуру тексту.

Відповідно до методики дешифровки, запропонованої Ю. В. Кнорозовим, припускається, що невідомий текст записується деякою, поки для нас невідомою, але натуральною мовою, яку не піддавали спеціальним перетворенням, таким як зашифровка з метою ускладнити читання цього тексту. Припускається, що в тексті існує певна структура зв'язків, до того ж у більшості випадків зв'язки тексту знаходяться на невеликій відстані і близькі елементи — в деякому співвідношенні. Тому основний метод — позиційна статистика.

### **Опис етапів**

Розбиття непереривного тексту (написаного без прогалин) на окремі блоки, відповідно до основних словоформ на основі виділення в тексті стійких знакосполучень. Якщо знаєш напрям письма, а його часто підказує орієнтація знаків, і видно, що він поділений на блоки, можна ввести певну систему координат. Це означає, що позицію кожного блоку можна визначити стосовно початку тексту, а позицію окремого знака в блоці — стосовно до початку і кінця цього блоку.

Наступна задача — аналіз морфології слова, до якої належить розбиття кожного блоку на окремі частини, їх класифікація з метою виділення постійних і перемінних частин, відповідних кореневим і службовим морфемам. Важливою і необхідною задачею також є з'ясування структури речення. На основі отриманих даних можна ставити задачу щодо виявлення класів блоків як за морфологічною, так і за синтаксичною структурою. Потім, використовуючи і факти порівняльного мовознавства і додаткові нетекстові дані (археологічних, історичних, філологічних та інших відомостей), потрібно конкретизувати граматику невідомої мови і тільки потім переходити до вивчення смислу папок. Під час дослідження структури словоформи, використовується прийом “оточення”: для кожного елемента тексту вказуються групи з  $n$ -елементів тексту, які “оточують” в тексті зліва і справа елемент, що досліджується. Такі оточення дають змогу виявити зв'язки між елементом, що досліджується та іншими “близькими до нього” частинами тексту.

*Комбінаторний метод включає як дешифровку зсередини (аналіз комбінацій мовних елементів), так і дешифровку ззовні (залучення даних різного характеру: місцезнаходження надпису, предмет, на якому він нанесений, супроводжуючі зображення, відомості про людей і їх сусідів).*

*Атрибуція (А.) те саме, що й авторизація* — встановлення належності анонімного твору чи у випадку містифікації. В широкому розумінні А. — встановлення авторства твору на основі композиції, способів текстоутворення, почерку, мови й позатекстових відомостей про його походження та історію. Текст, авторство якого не підлягає визначенню, називається анонімним. Якщо автор ставить своє ім'я під чужим твором — це плагіат. Інколи автор видає твір під псевдонімом.

Проблема встановлення авторства (атрибуції) анонімних і псевдонімних статей давно не завдання філології. За наших часів рішення цього завдання пов'язано як з прикладними історико-філологічними дисциплінами, так і з природничо-технічними науками, такими як статистика та теорія ймовірностей, теорія комунікацій та ін. Природно, що з розвитком обчислювальної техніки, з'явилась можливість використовувати більш ефективні методи для вирішення цієї проблеми, враховуючи різні лінгвостатичні параметри. Основна мета формально-кількісних методів — надати процесу атрибуції об'єктивнішого характеру.

Формально-кількісні методи пов'язані з аналізом синтаксичних структур, сполученням граматичних класів, довжиною речення, насиченням словника. Наприклад, завданням досліду є встановлення авторства Прянікова щодо деяких статей, чи навпаки, відхилення гіпотези про те, що автором суперечливих статей є Пряніков. Для цього необхідно: по-перше, визначити формально-граматичні ознаки стилю Прянікова; по-друге, порівняти встановлений інваріант з граматичними параметрами дослідних анонімних статей. Методики проведення аналізу можуть будуватися, наприклад, на основі розподілення частин мови на перших трьох і останніх трьох позиціях речень. За допомогою спеціальних програм з бази даних творів для кожної статті отримують частотні таблиці частин мови для визначених позицій речення. Однак перед тим як використовувати ті чи інші

формальні характеристики для атрибуції тексту, необхідно визначити підсвідомі особливості мови автора, тобто елементи авторської манери, присутні обов'язково в його будь-якому тексті. Отже бажано знайти такі характеристики, які можна виділити у будь-якого автора, але значення цих характеристик відрізнялись би у різних авторів. Також слід визначити інваріанти для всіх текстів конкретного жанру, незалежно від авторства.

### **Області використання**

Останнім часом завдання визначення авторства тексту має більше значення і для криміналістики. Типове завдання — ідентифікація автора анонімного тексту за його мовними параметрами (визначення авторства різних анонімних листів, які містять погрози, шантаж).

Лінгвістична експертиза важлива при розв'язанні суперечок, пов'язаних з політичними проблемами. Анонімна чи псевдоанонімна інформація частіше розповсюджується під час передвиборчих кампаній з метою дискредитації конкурентів. У зв'язку з розвитком електронних мереж та збільшенням інформації, яка поширюється з їх допомогою, загострюється проблема дотримання авторських прав.

В юридичній психолінгвістиці сформувався напрям, пов'язаний з лінгвістичними методами ідентифікації особистості за анонімним текстом. Загальний метод полягає в тому, щоб на основі досліджень формальних (лінгвістичних, текстових) і неформальних (змістовних, смислових) характеристик тексту визначати окремі психологічні риси людини, якщо можливо, то його повний психологічний портрет.

### **Конкретні розробки**

Нині існує програмний комплекс “Атрибуція” для лінгвістичного аналізу, який дозволяє в діалоговому режимі здійснювати граматичний та синтаксичний розбір літературних текстів, використовуючи багаточисленні лінгвістичні характеристики, наприклад, такі як частини мови: відмінок, рід, число, тип речення та ін. Програмний комплекс складається з двох частин: модуль “Граматичний аналіз” і модуль “Синтаксичний аналіз”. Вони дають можливість формалізувати 69 лінгвістичних параметрів. Принцип роботи обох модулів однаковий: вхідною інформацією є літературний текст в електронному вигляді. Кожний

модуль насамперед виділяє ціле речення, а потім дозволяє працювати з кожним словом речення, пропонуючи у простих ситуаціях (наприклад, сполучник) свій варіант значення, але вирішальне право на прийняття рішення має користувач, який має бути спеціалістом-філологом. На виході отримуємо структурований, у вигляді таблиці, файл з повним синтаксичним чи граматичним розбором. Модуль “Грамматичний аналіз” орієнтований на виокремлення таких лінгвістичних параметрів, як частина мови, відмінок, число, рід та ін. Загальна кількість параметрів дорівнює 46.

У модулі “Синтаксичний аналіз” аналізуються структура і граматичне значення синтаксичних одиниць, визначається тип конструкції. Список параметрів включає: тип речення (розповідне, окличне, питальне), частину мови для підмета, кількість членів для присудка.

За допомогою цього програмного комплексу опрацьована 31 стаття Ф. М. Достоєвського.

### **Зображення знань в системах штучного інтелекту (ШІ)**

Комп’ютерна лінгвістика як окрема прикладна дисципліна виокремлюється за інструментом — за використанням комп’ютерних засобів для опрацювання мовних даних. Спільної метамови немає — комп’ютерні проги, які моделюють ті чи інші аспекти функціонування мови, і використовують різні засоби програмування. Існують спільні принципи комп’ютерного моделювання мислення, які так чи інакше реалізуються в будь-якій комп’ютерній моделі.

Теорія знань, розроблена в ШІ, є важливим розділом когнітивної науки. Головний тезис: мислення — процес обробки і породження знань. При цьому “знання” — невизначена категорія. “Процесор” — когнітивна система людини. В когнітивній науці — два види знань: декларативні (“знання чого?”) і процедурні (“знання як?”). Декларативні знання — сукупність тверджень (примітка: тлумачення слів у тлумачному словнику). Піддаються верифікації в термінах “істина — неправда”.

#### **Питання для самоконтролю**

1. Лінгвістичні ознаки.
2. Що є об’єктом лінгвістичної дешифровки?
3. Що таке шифр?

4. Чому позиційна статистика вважається основним методом вивчення невідомих текстів?
5. Що таке атрибуція і авторизація?

*Література* [28; 32; 33; 36; 40; 42; 48; 49; 59; 66; 68]

### **Тема 11. Генеративна (породжуюча) граматики**

Генеративна (породжуюча) граматики (ГГ) одна з основних напрямів сучасної лінгвістики, створена Хомським у США у другій половині 50-х років ХХ ст.

ГГ займається аналізом мовних даних, кінцевий результат дослідження — відповідь на питання про те, яким чином людина засвоює мову. ГГ — це не теорія мови, а теорія засвоєння мови.

**Суть глибинної структури (ГС):** два рівні глибинної структури і поверхневої структури (ПС), пов'язані за допомогою трансформацій.

За час свого існування генеративізм пройшов кілька етапів. Найбільш значні з них такі:

**1. Стандартна теорія (СТ).** Термін “стандартна теорія” об'єднує моделі з книги Хомського “Синтаксичні структури” та “Аспекти теорії синтаксису”. Схема СТ передбачає перехід від правил структур, які належать до ГС, і далі шляхом трансформації до ПС.

Щоб мати уявлення про мову як механізм зародження незліченної множини речень за допомогою кінцевого набору граматичних засобів, Хомський запропонував поняття:

- ГС: глибинної структури (прихованої від безпосереднього сприйняття, і яка походить із системи рекурсивних правил, тобто таких, що можуть використовуватися багаторазово);
- ПС: поверхневої граматичної структури (що безпосередньо сприймається);
- трансформацій (які описують перехід від глибинних структур до поверхневих).

Одній ГС може відповідати більше однієї ПС (наприклад, пасивна конструкція *Указ підписується президентом* виводився за допомогою пасивізації з тієї самої ГС, що й активна конструкція *Президент підписує указ*) і навпаки (так, неоднозначність *Відвідування родичів може бути стомлюючим* описується як



результат співпадіння ПС, походять з двох різних ГС, в одному з яких родичі є тими, хто відвідує когось, а в іншому — тими, кого хтось відвідує).

Модель “Аспектів”, чи стандартна теорія, викладена в книзі Хомського “*Аспекти теорії синтаксису*” (*Aspects of the Theory of Syntax*, 1965, рос. пер. 1972) і являє собою спробу введення у формальну модель семантичного компонента — правил семантичної інтерпретації, які приписують значення ГС-ам. Введено п/п-е мовної компетенції (системи процесів походження мовних висловлювань) і використання мови (performance), прийнята т. д. гіпотеза Катца-Постала щодо зберігання смислу при трансформації, у зв’язку з чим виключено поняття факультативної трансформації, і введено апарат синт-ознак, які описують лексичне поєднання.

Розширена стандартна теорія, чи “лексикалізм”: отримала розповсюдження в 70-х роках ХХ ст. На відміну від попереднього етапу, особлива увага надається лексиці й семантиці інтерпретації мовних одиниць. Відповідно до гіпотези Каца-Постала, трансформації не змінюють значення висловлювання; все, що має семантичну інтерпретацію, має бути представлено на рівні ГС. Сам Хомський виділив 3 групи явищ, які заперечують положення породжуючої семантики: непродуктивність деяких мовних процесів (номіналізації), структурний характер глибинного представлення і здатність трансформацій до змін значень мовних виразів.

**2. Теорія управління і пов’язування (70-ті роки, *Лекції про управління і зв’язування (Lectures on Government and Binding*, 1981, GB-теорія за першими буквами).** Основна зміна — відмова від специфічних правил, які описують синтаксичні структури конкретних мов, і заміна їх деякими універсальними обмеженнями. Всі трансформації замінені однією універсальною трансформацією переміщення. Виділені деякі модулі (X-штрих-теорія, теорія обмеження, теорія пов’язування, теорія управління, теорія відмінку, тета-теорія), кожна з яких відповідає за свою частину граматики, діє за своїми принципами та має параметри, які налаштовуються, визначають конкретно мовну специфіку. Оскільки поняття принципів та параметрів збереглося і на наступному етапі розвитку генеративізму, інколи говорять

про теорію принципів та параметрів як про особливу стадію, яка охоплює 2 та 3 етапи генеративізму.

Важлива особливість теорії принципів і параметрів — **модульність**, тобто ділення граматики на незалежні блоки (**модулі**), кожний з яких визначає один аспект мовної компетенції людини і включає один чи кілька принципів. Такі блоки також називаються “теоріями”. Виділимо найважливіші з них:

1) **теорія семантичних ролей** (“тета-ролей”): кожна семантична роль предикату (діюча особа, об’єкт дії) у будь-якій конститутції повинна бути заповнена (в деякому випадку вона може бути заповнена “пустою категорією”, тобто синтаксичною одиницею, яка не має звукового виразу);

2) **теорія відмінків**: задає обмеження на відмінкове оформлення підмету і додатків;

3) **теорія пов’язування**: набір універсальних обмежень на поведінку “пов’язаних” слів чи словосполучень, тобто які позначають той самий предмет, що й інше слово чи словосполучення, яке займає більш пріоритетну структурну позицію в тому самому реченні (наприклад, зворотний займенник, який позначає той самий предмет, що й підмет);

4) **теорія відмежування**: визначає заборони на синтаксичні переміщення;

5) **теорія управління**: має умови, за яких певна синтаксична позиція (наприклад, позиція підмету чи прямого додатку) може бути заповнена фонетично невираженим елементом.

**3. Мінімалістична програма** (основні положення викладені Хомським у книзі *The Minimalist Program*, 1995). Суть: мінімізація мовних уявлень і опис їх взаємодії з іншими когнітивними системами, постулюючи в “мовному апараті людини дві головні підсистеми: **обчислювальна і лексикон**. *Обчислювальна підсистема* породжує мовні вирази і подає накази системам реалізації. *Лексикон* утримує всю конкретну інформацію конкретної мови”. Можна говорити про дві системи реалізації: **артикуляторно-перцептивну і концептуально-інтенціональну**.

Отже, ця програма передбачає в мовному апараті людини дві головні підсистеми — **лексикон і обчислювальну систему**, а також два інтерфейси — **фонетичний і логічний**.

Загальна схема:

**Правила базової граматики**

↓

**На виході: базові (глибинні) структури → Семантична інтерпретація (значення)**

↓

**Трансформаційні правила**

↓

**На виході: поверхневі структури**

↓

**Фонологічні правила**

↓

**На виході: звукове представлення**

**Модель Мельчука**

Основа теорії — природна мова — це особливого виду перетворювач, який виконує переробку заданих смислів у відповідні їм тексти і заданих текстів у відповідні їм смисли. Він має на увазі багато-багатозначне відображення множини смислів у множину.

Повна модель: дійсність більшого смислу і більшого тексту, і більшого ніж мовлення, що звучить.

Модель: інформація передається через послідовність мовленевих сигналів (акустичних / візуальних); послідовність сигналів, які несуть інформацію, спрямовується від мовця / який пише до слухача / який читає через певний канал зв'язку. Слухач бере з сигналу ту / майже ту інформацію, яку мовець мав на увазі, завдяки тому, що обидва володіють одним і тим самим кодом / правила відповідності між мовними сигналами і мовленевою інформацією. Мельчук бере інформацію, послідовність мовленевих сигналів і код, і називає все це смислом (сміслами), текстом і відповідністю між смислом і текстом.

**Текст** — достатньо деталізована фонетична транскрипція мови, знаходиться на самій поверхні, а в глибині — позамова дійсність.

**Смисл** — конструктивний семантичний / смисловий запис (пучок відповідностей між реальними рівнозначними висловлюваннями, фіксуємо за допомогою спеціальних символів).

Положення моделі:

1. МСТ — це діюча / динамічна модель, прилад, який має вхід (завдання) та вихід (результат) + схеми переробки завдання в результат.

Існує дві логічно незалежні частини:

- опис відповідностей між смислами та текстами;
- опис механізму, який реалізує ці відповідності у процедурі дійсного переходу від заданого смислу та відповідністю текстам і навпаки.

Опис відповідностей включає три компоненти:

1) інвентар елементарної одиниці смислу — семантика і правила їх комбінування у складні смислової одиниці — семантичного представлення;

2) інвентар елементарних одиниць текстур морфів та правила їх комбінування у складні текстові — словоформа, словосполучення, речення;

3) правила, які зіставляють будь-які смислові одиниці у відповідні текстові одиниці.

Правила, які встановлюють відповідності, мають таку логічну структуру: якщо взяти будь-які елементарні смисли (семи) та скомбінувати їх яким-небудь способом, то для їх вираження можна взяти будь-які елементарні тексти (морфи) та скомбінувати їх будь-як.

2. МСТ задана зовсім формально, тобто за допомогою однозначних та логічних послідовних формулювань, які не вимагають залучення додаткової інформації.

3. МСТ — це змінний прилад на відміну від приладу Хомського. Це не генератор текстів, а транслятор смислу тексту, тобто вміння переробити змістовно БУДЬ-ЯКЕ смислове завдання у відповідні тексти + витягнути з заданого тексту його смисл.

## **5 рівнів представлення МСТ**

### **(на прикладі перетворення смислу в текст)**

1. Сем Уж — зображення змісту зв'язного фрагмента мови без розділу на фрази та слова у вигляді сем-представлення:

2 компоненти:

- сем-граф, вершини яких позначені семами / символами комплексів семи, а дуги позначаються зв'язками;
- відомості про комплексну організацію смислу, що розглядається, + логічні акценти.

## 2. Синтаксичний рівень має два підрівні:

а) глибинний синтаксис являє собою зображення висловлювання у вигляді послідовності глибинно-синтаксичних представлень фраз (ГСП), які складаються з 5 компонентів:

- глибинна синт-структура фрази — це дерево залежності, вузли якої представляють лише знаменні словоформи фрази у вигляді узагальнених лексем (реал-лексеми даної мови; лекс-форми; фраземи — фразеологічні звороти; фіктивні лексеми — символи для нульових висловів);
- відомості про комплексні організації фрази (тема / рема);
- відомості про просодичний характер фрази (інтонація);
- відомості про тотожність денотатів тих чи інших іменних груп (які потім замінюються на місці);
- відомості про угруповання слів (наприклад: я знаю лише РЯ);

б) поверхневий синтаксис — зображення висловлювання у вигляді послідовності поверхнево-синтаксичних представлень, які складаються з 5 компонентів:

- поверхнево-синтаксичної структури — дерево залежності максимально близьке текстовій фразі, тобто вузли складають всі словоформи фрази і лише їх (без порожніх, фіктивних лексем), а гілки представляють типові поверхневі синтаксичні відношення даної мови (наприклад, відтінкові);
- ост. — як у ГСП.

## 3. Морфологічний рівень складається з двох підрівнів:

а) глибинна морфологія — зображення фрази у вигляді її глибинно-морфологічного представлення (ГМП) двох компонентів:

- ланцюжок ГМП всіх словоформ фрази (лінійно упорядкована послідовність);
- відомості про просодичні характеристики фрази (інтонація, пауза, наголос).

Наприклад: ЛЕВ<sub>од. речення</sub>, ИДТИ<sub>несов, прош, изъав, невозвр, ед, муж</sub>;

б) поверхнева морфологія зображення фрази у вигляді її поверхневої морфології представлення (ПМП), в якому є два компоненти:

- послідовність ПМП;
- відомості про просодичні характеристики фрази.

ПМП — це відповідний словоформі ланцюжок морфем, супра-морфем та аналогічних морфемам сімейств значущих морфем-

них операцій. Супраморфеми — сімейства морфемно значущих просодичних явищ. Аналогічні сімейства — чергування, конверсії.

Наприклад: {лев}+{од, речен}  
{розглядати}+{іва}+{прош}+{ми}+{ся}.

4. Фонологічний рівень — фонемна транскрипція фрази — послідовність символів, які представляють фонемні і просодеми.

5. Фонетичний рівень — фонетична транскрипція фрази — послідовність символів, які представляють звуки мови та просодичні явища.

### **Ідеалізована картина роботи МСТ**

Завдання надходить на вхід чергового перетворювача-генератора, який видає для цього завдання багатьох можливих представлень наступного рівня. Отримані результати фільтруються селектором чергового перетворювача, який потім по черзі опрацьовує залишені варіанти, будуючи для кожного з них багато альтернативних представлень.

### **Питання для самоконтролю**

1. Чим займається генеративна граматики?
2. Які є незалежні блоки або теорії поділу граматики?
3. Яка суть мінімалістичної програми?
4. Які положення моделі Мельчука?
5. Назвіть рівні представлення МСТ.

*Література* [6; 8; 10; 12; 16; 20; 53; 57; 58; 62]

### **Тема 12. Когнітивна лінгвістика**

Когнітивна лінгвістика (КЛ) вивчає мову як прояв ментальної діяльності людини, як засіб пізнання світу, його категоризації, концептуалізації та оцінки.

Лідери КЛ — Кубрикова, Баранов, Караулов, Карасик, Болдарев, Нікітін; зарубіжні КЛ — Чейф, Шенк, Лакофф, Тальми.

Постулати КЛ:

1) нерозривний зв'язок між мовою та мисленням (мовні структури завжди розглядаються крізь призму мислення, тісного зв'язку між мисленням та психологією);

2) міждисциплінарність підходів (когнітивістика — федерація наук => потрібно враховувати дані культурології, етимології);

3) постулат про примате когнітивного (при лінгво-когнітивному дослідженні когнітивний аспект важливіший за лінгвістичний, когнітивна близькість важливіша, ніж відмінності у мовленнєвих структурах; важливіше сутнісне представлення знань ментальні одиниці — концепт, фрейм, сценарій, модель);

4) постулат про нерелевантності більшого-меншого лінгвістичного і екстралінгвістичного знання (традиційна лінгвістика їх розмежовувала; когнітивістика вважала, що у цьому немає сенсу: часто, щоб зрозуміти смисл висловлювання, необхідна лінгвокультурологічна інформація — наприклад: *Він проплив стометрівку за 45 сек — потрібно знати, що це світовий рекорд*);

5) постулат про тенденцію до економії зусиль (реальне втілення у словах бідніше, ніж думка);

6) постулат про множинність втілення когнітивної структури у мові (один зміст може бути виражений по-різному — наприклад: синоніми);

7) постулат про неоднорідності плану змісту у мові (традиційна структурна лінгвістика — наказувалось знайти однотипні одиниці — набір різних ознак — наприклад: дочка, холостяк);

8) постулат про множинність методик семантичного опису (традиційна семантика — може створити одну ідеальну методику та на її основі описати всю лінгвістику, в ідеалі — вся мова — наприклад: лексико-семантичні групи; у реальності різні групи потребують різних методик опису, і ці різні методики взаємодоповнюють одна одну);

9) постулат про значимість нестандартних застосувань (лінгвісту цікаві всі види незвичайного застосування слова — помилки, мовна гра — наприклад: імпузантний);

10) постулат про когнітивну обов'язковість (передбачає можливість і вимоги верифікації даних однієї когнітивної науки за допомогою методів та прийомів інших когнітивних наук).

### **Структура когнітивних категорій**

Для того щоб отримати, опрацювати, зберегти, добувати інформацію, людський мозок організовує її в категорії, і на відміну від лінгвістичних категорій у чітко визначених межах, а належність до будь-якої категорії визначається в межах необхідних і достатніх характеристик якого-небудь об'єкта, і всі члени категорії рівні між собою. Ментальні категорії навпаки не мають чіткої межі, але мають внутрішню будову. Вони не суперечать суво-

рій належності (елементи не рівні, можуть взагалі залишати її і входити в нову категорію), але! у кожній групі є свій прототипічний член, який пов'язаний відносинами з іншими членами категорії.

Віднесення до якої-небудь категорії не завжди залежить від точної наявності певної ознаки, типових якостей у всіх її членів. Категорія побудована у вигляді ієрархічного ланцюжка / ланцюжка послідовно пов'язаних членів (члени однієї категорії, які знаходяться на одній периферії, можуть не мати необхідні та достатні характеристики / загальні якості з прототипом).

Радіальні категорії більш складні, мають кілька ієрархічних ланцюжків, які розходяться від прототипу. Кінці цих ланцюжків не можуть бути пов'язані, відбувається послідовна зміна характеристики, наприклад: doubt → question → challenge → invite → request → ask → communicate → convey → transfer → move.

Основний принцип відношення від ядра до периферії — розширення / спрощення знань / зміна однієї з характеристик. Виділення ядра категорій мотивується ставленням людини до членів категорії, так як наша свідомість побудована нашим сприйняттям, і людина взаємодіє з ними, означає більше, ніж об'єктивний опис словника.

Одні категорії є членами більш складних категорій, і члени якої-небудь категорії не є зашитими в інші категорії (немає визначеного місця в категорії), а можуть входити у кілька ланцюжків різних категорій (що призводить до двозначності, неясності, накладання, збігів) — різнорідно мотивовані члени категорії.

Пізнання та використання мови включає в себе звертання до ментальних просторів (організовується наслідком досвіду людського сприйняття та розширюється за допомогою процесів образного співвідношення / зіставлення — метафора, метонімія, зрощення).

*Метафора* — це зіставлення однієї області з іншою.

*Метонімія* — це коли частина замінює ціле / тісно пов'язане з яким-небудь поняттям (еліпсис, фонологічна редукція).

*Зрощення* — це зіставлення двох областей для утворення третьої (мото + готель = мотель, мова = реальність + мислення).

Когнітивна лінгвістика розглядає мову як невід'ємну частину пізнання, яка відображає взаємодію культурного, психологічно-



го, комунікативного і функціонального аналізу, та яке може бути сприйняте (зрозуміле) з точки зору концептуального представлення реальності і внутрішнього аналізу.

### **Обчислення ілокутивних актів (ІА) (Дж. Серль)**

Теорія мовленнєвих актів (МА) виокремлює 3 рівні (аспекти) аналізу МА:

1. МА як власне мовлення людини — локутивний акт.

2. У процесі мовлення людина одночасно здійснює якусь дію, яка має якусь позамовну мету: він запитує чи відповідає; інформує, стверджує чи попереджає; назначає когось кимось; критикує когось за щось та ін. МА, який розглядається з точки зору його позамовної мети, як ілокутивний акт. У змісті висловлювання на ілокутивному рівні виокремлюють дві основні складові: ілокутивну функцію ( $F$ ) — узагальнення та цілісна характеристика висловлювання як засобу здійснення ІА та пропозицію ( $P$ ) — об'єктивне, тобто позбавлене істинної оцінки, змісту, вираженого реченням, що представляє його у вигляді формули  $F(P)$ . Так, зміст висловлювання у прикладі, коли мати говорить сину *Сідай за уроки!* розкладається на пропозиціональні частини “ти сідаєш за уроки” (із знятим установленням, тобто без істинної оцінки) та ілокутивну функцію спонукування “*Ти сідаєш за уроки?*”, має той же самий пропозиційний зміст, але іншу ілокутивну функцію — функцію питання.

3. За допомогою мовлення людина досягає певних результатів, виконуючи ті чи інші зміни в оточуючій дійсності, і насамперед — у свідомості свого співрозмовника. МА, якщо розглядати в аспекті його реальних наслідків, виступає як перлокутивний акт.

### **Основні поняття теорії ілокутивних актів**

*Контекст вимовлення* — контекст, в якому шляхом вимовлення здійснюється ІА. Так як одне й те саме речення може бути вимовлено в різних контекстах, ілокуція буде різна. Наприклад, *Я повернусь* — може бути обіцянка/погроза. Разом з контекстом буде змінюватись і його пропозиційний зміст, так як у різних контекстах воно буде вказувати на різних мовців та різні часові уривки. Контекст вимови складається з п'яти компонентів: 1) мовця; 2) слухача; 3) часу; 4) місця; 5) деякі інші характерис-

тики, пов'язані зі здійсненням МА (психологічний стан, наміри, бажання, вірування та ін. мовця і слухача) — світ вимовляння.

Можливі світи — це ті, які могли б бути справи, це такий світ, в якому об'єкти дійсного світу (дійсні об'єкти) мають інші властивості, і/чи такий світ, в якому є об'єкти, відмінні від дійсних (безліч можливих світів позначається  $W$ ).

Дійсні світи — те, як є справи насправді (безліч можливих світів містить елемент  $w_0$ , який є дійсним світом).

З кожним можливо діючим світом пов'язана  $U$ -предметна область світу, яка містить всі об'єкти, які належать цьому світу.

Об'єднання предметних областей усіх можливих світів є універсумом міркувань ілокутивної логіки. Ця безліч містить усі колишні, теперішні та майбутні індивіди, існування яких можливе.

### Формальні якості ілокутивних сил (ІС)

Їх можна розділити на сім компонентів:

1. Ілокутивна мета — задум, притаманий цьому типу акта (наприклад, мета ствердження — сказати, як ідуть справи, наказу — намагатися змусити кого-небудь робити щось). Існує п'ять ілокутивних цілей:

а) **асертивна** — сказати, як ідуть справи (асертиви *стверджувати, заявляти, признавати, повідомляти*);

б) **комісивна** — мовець зобов'язує сам себе зробити що-небудь (комісиви: *обицяти, клястися, погрожувати, присягати, ручатися*);

в) **директивна** — змусити іншого зробити щось (*примушувати, вимагати, спокушувати, уполити, радити*);

г) **декларативна** — змінити обставини зовнішнього світу за допомогою вимовляння (*назначати, нарікати, об'являти, проголошувати*);

д) **експресивна** — висловити почуття чи установки (*вітати, дякувати, співчувати, хвалити, жаліти*).

2. Спосіб досягнення ілокутивної мети. У більшості випадків ілокутивної мети можна досягти різними засобами, так що при цьому зміняться і формальні якості ІС вимовляння. Деякі ІС мають особливий спосіб досягнення своєї мети. Наприклад, віддаючи наказ, мовець досягає директивної ілокутивної мети тим, що апелює до свого службового/іншого положення, що дає йому

владу над слухачем. Умовляючи слухача зробити щось, мовець досягає тієї самої мети шляхом свого приниження перед слухачем. Такий спосіб досягнення ілокутивної мети, який обмежує умови її досягнення чи зобов'язання її досягнути, називається особливим засобом досягнення.

3. Інтенсивність ілокутивної сили. Більшість ілокутивних цілей така, що їх можна досягнути з більшою/меншою інтенсивністю. Наприклад: *а) порадити, щоб слухач вийшов з кімнати* ~ *б) наказати, щоб він вийшов; а) констатувати, що іде дощ* ~ *б) клястися, що іде дощ*. ІА (а) слабше, ніж (б). В ілокутивній логіці, такі факти репрезентують приписуванням більших/менших чисел різним ступеням інтенсивності.

4. Умови пропозиційного змісту. Деякі ІС накладають обмеження на пропозиційний зміст. Ці обмеження називаються умовами пропозиційного змісту даних ІС. Наприклад, багато пропозицій, які належать до майбутнього з точки зору моменту часу *t*, задає деяку умову пропонованого змісту (наприклад, умова пропонованого змісту для передвіщання).

5. Попередні умови. Наприклад, попередня умова розпорядження, наказу, прохання складається з того, що пропозиційний зміст є фізично можливим, тобто слухач здатний реалізувати у майбутньому лінію дій, репрезентовану пропонованим змістом.

6. Умови щирості. Якщо мовець висловлює той чи інший психічний стан, то разом з тим він висловлює і всі ті психологічні стани, до яких його зобов'язують перші. Наприклад, якщо мовець висловлює співчуття у зв'язку з деяким станом справ, він разом з тим висловлює і віру в наявний стан справ, тому що він не може співчувати, якщо в нього немає цієї віри.

7. Інтенсивність умов щирості. Психічний стан виявляється у *РА* з більшою/меншою інтенсивністю залежно від відповідної ІС. Наприклад, мовець, який дає урочисту обіцянку щось зробити, виявляє сильніший намір, ніж мовець, який просто приймає на себе обов'язок зробити це. Інтенсивність умов щирості виявляється цілим числом. Інтенсивність ІС залежить від інтенсивності умов щирості та інтенсивності, пов'язаної з засобом досягнення.

Кожна  $F$  є упорядкованою послідовністю цих семи елементів. Але! не всяка вільно взята упорядкована послідовність цих елементів є  $F$ , так як різні компоненти можливої  $F$  пов'язані визначеними логічними відносинами, які накладають обмеження на можливі комбінації цих елементів. Наприклад, всі  $F$  з директивною ілокутивною метою повинні мати такий пропорційний зміст, який репрезентує майбутню лінію дій слухача з попередньою умовою, яка полягає в тому, що слухач має змогу реалізувати цю лінію і умову щирості, яка складається з того, що мовець хоче, щоб слухач зробив певну дію.

### **Питання для самоконтролю**

1. Що визначають постулати КЛ?
2. Яку будову мають ментальні категорії?
3. Рівні аналізу МА.
4. Назвіть основні поняття теорії ілокутивних актів.
5. Які якості мають ІС?

*Література [4; 8; 10; 34; 52; 62; 66; 68]*

### **ПИТАННЯ ДО КУРСУ**

1. Зв'язок і співвідношення теоретичної і прикладної лінгвістики. Моделювання і оптимізація як основні методи прикладної лінгвістики.
2. Основні типи задач, які вирішує прикладна лінгвістика.
3. Когнітивний напрям у гуманітарних науках. Процедурне і декларативне знання. Способи його формалізації.
4. Поняття про фрейми, сценарії, плани, моделі світу.
5. Прототипи і їх різновиди як поняття когнітивної лінгвістики.
6. Моделювання комунікації (принципи, досягнення).
7. Моделювання сюжетів: ідеї В. Я. Проппа і сюжетні граматики. Гіпертекст і людська свідомість. Специфіка гіпертекстового представлення знань.
8. Типи гіпертекстів. Комп'ютерні засоби створення і підтримки гіпертекстів.
9. Природа продукування і розуміння мовлення. Опора на характеристики мовлення у прикладному моделюванні.

10. Проблеми лінгвістичної атрибуції тексту. Кількісні прийоми атрибуції.
11. Текстові корпуси і корпусна лінгвістика. Основні поняття корпусної лінгвістики.
12. Текстові корпуси і корпусна лінгвістика. Моделювання предметної області текстовим корпусом. Основні вимоги до текстового корпусу.
13. British National Corpus ([www.natcorp.ox.ac.uk/using/index.html](http://www.natcorp.ox.ac.uk/using/index.html)). Принципи організації.
14. Національний корпус української мови. Принципи організації, стан, можливості використання.
15. UNILEX-T як програма обробки текстів і текстових корпусів.
16. Інформаційно-пошукові системи. Інформаційно-пошукові мови.
17. Структурні компоненти словника і словникової статті (зони). Машинні словники.
18. Частотні словники (структура, використання у прикладних цілях).
19. Частотний словник як відображення мовної “картини світу”.
20. Комп’ютерні версії асоціативних словників, можливості їх використання у прикладних цілях.
21. Термінологія, терміносистеми і їх обслуговування прикладною лінгвістикою.
22. Галузеві тезауруси і їх призначення. Структура тезаурусу за теоретичною і прикладною лінгвістикою С. Е. Нікітіної.
23. Лінгвістичні термінологічні словники (О. С. Ахманової, Ж. Марузо, Н. В. Подольського, “Англо-русский словарь по лингвистике и семиотике” А. Н. Баранова, Д. О. Добровольського та ін.).
24. Машинний переклад як проблема прикладної лінгвістики.
25. Комп’ютерна лінгводидактика (принципи, засоби).
26. Поняття мовної політики. Засоби мовної політики. Закони про мови.
27. Політична лінгвістика. Методика контент-аналізу.
28. Основні поняття нейролінгвістичного програмування.
29. Лінгвістичні експертизи, їх різновиди, комп’ютерна підтримка.
30. Законодавча база лінгвістичної експертизи. Права і обов’язки експерта.

## СПИСОК ЛИТЕРАТУРЫ

### Основна

1. Бакулов А. Д., Леонтьева Н. Н., Шаляпина З. М. Отечественные системы машинного перевода // Искусственный интеллект: Справочник. Кн. 1. Системы общения и экспертные системы. — М., 1990.
2. Баранов А. Н. Автоматизация лингвистических исследований: корпус текстов как лингвистическая проблема // Русистика сегодня. — 1998. — № 1–2.
3. Баранов А. Н. Введение в прикладную лингвистику. — М., 2000.
4. Бовтенко М. А. Компьютерная лингводидактика. — М., 2005.
5. Герд А. С. Предмет и основные направления прикладной лингвистики // Прикладное языкознание. — СПб., 1996.
6. Герд А. С. Научно-техническая лексикография // Прикладное языкознание. — СПб., 1996.
7. Звегинцев В. А. Теоретическая и прикладная лингвистика. — М., 1968.
8. Кибрик А. Е. Прикладная лингвистика // Очерки по общим и прикладным вопросам языкознания. — М., 1992.
9. Леонтьева Н. Н. Построение единиц типа “Ситуация” на основе семантического словаря “РУСЛАН” // Проблемы прикладной лингвистики. — М., 2004.
10. Марчук Ю. Н. Основы компьютерной лингвистики. — М., 2000: Новое в зарубежной лингвистике: Вып. XII. Прикладная лингвистика. — М., 1983.
11. Новое в зарубежной лингвистике. Вып. XIV. Проблемы и методы лексикографии. — М., 1983.
12. Новое в зарубежной лингвистике. Вып. XXIII. Когнитивные аспекты языка. — М., 1988.
13. Новое в зарубежной лингвистике: Вып. XXIV. Компьютерная лингвистика. — М., 1989.
14. Попов Э. В. Общение с ЭВМ на естественном языке. — М., 1982.
15. Потапова Р. К. Речевое управление роботом. — М., 1989.

16. *Потапова Р. К.* Речь: коммуникация, информация, кибернетика. — М., 1997; *Прикладное языкознание* / Под ред. А. С. Герда. — СПб., 1986.
17. *Субботин М. М.* Гипертекст. Новая форма письменной коммуникации / ВИНИТИ. — Т. 18. — М., 1994. Сер. Информатика.
18. *Филиппович Ю. Н., Прохоров А. В.* Семантика информационных технологий: опыты словарно-тезаурусного описания. — М., 2002.
19. *Фрэнсис У. Н.* Проблемы формирования и машинного представления большого корпуса текстов // Новое в зарубежной лингвистике: Вып. XIV. Проблемы и методы лексикографии. — М., 1983.
20. *Щерба Л. В.* Опыт общей теории лексикографии // Щерба Л. В. Языковая система и речевая деятельность. — М., 1974.

#### *Додаткова*

21. *Андрющенко В. М.* Концепция и архитектура машинного фонда русского языка. — М., 1989.
22. *Апресян Ю. Д., Богославский И. М., Иомдин Л. Л. и др.* Лингвистическое обеспечение системы ЭТАП-2. — М., 1989.
23. *Белоногов Г. Г., Кузнецов Б. А.* Языковые средства автоматизированных информационных систем. — М., 1983.
24. *Блюменау Д. И.* Проблемы свертывания научной информации. — Л., 1982.
25. *Виноград Т.* Программа, понимающая естественный язык. — М., 1976.
26. *Виноград Т.* Работа с естественными языками // Современный компьютер. — М., 1986.
27. *Виноградов В. В.* Проблема авторства и теории стилей. — М., 1961.
28. *Войскунский А. Е.* Моделирование мышления // Речевое общение: проблемы и перспективы. — М., 1983.
29. *Герд А. С.* Предмет и основные направления прикладной лингвистики // Прикладное языкознание. — СПб., 1996.
30. *Гончаренко В. В., Шингарева Е. А.* Фреймы для распознавания смысла текста. — Кишинев, 1984.

31. *Городецкий Б. Ю.* Актуальные проблемы прикладной лингвистики // Новое в зарубежной лингвистике: Вып. XII. — М., 1983.
32. *Городецкий Б. Ю.* К созданию Машинного фонда русского языка (определение, применения, актуальные проблемы) // Машинный фонд русского языка: идеи и суждения. — М., 1986.
33. *Городецкий Б. Ю.* Компьютерная лингвистика: моделирование речевого общения // Новое в зарубежной лингвистике: Вып. XXIV. Компьютерная лингвистика. — М., 1989.
34. *Городецкий Б. Ю.* О лингвистическом подходе к теории информационных языков // Исследования по структурной и прикладной лингвистике. — М., 1975.
35. *Звегинцев В. А.* Теоретическая и прикладная лингвистика. — М., 1968.
36. *Исаев И. А.* Опыт автоматизации лексикографических исследований. Система DIALEX // Слово Достоевского. — М., 1996; *Искусственный интеллект*: В 3 кн. — М., 1990. — Кн. 1. Системы общения и экспертные системы. Кн. 2. Модели и методы; *Использование ЭВМ в лингвистических исследованиях*. — К., 1989; *Лингвистические проблемы автоматизации редакционно-издательских процессов*. — К., 1986.
37. *Маккьюн К.* Дискурсивные стратегии для синтеза текста на естественном языке // Новое в зарубежной лингвистике. Вып. XXIV. Компьютерная лингвистика. — М., 1989.
38. *Мальковский М. Г.* Диалог с системой искусственного интеллекта. — М., 1986.
39. *Марусенко М. А.* Атрибуция анонимных и псевдонимных текстов методами прикладной лингвистики // Прикладное языкознание. — СПб., 1996.
40. *Марчук Ю. Н.* Проблема машинного перевода. — М., 1983.
41. *Марчук А. Г., Холюшкин Ю. П., Загоруйко Ю. А.* Разработка новых методов и информационных технологий представления и обработки археологических и этнографических данных // Информационные технологии в гуманитарных исследованиях. — Новосибирск: Изд-во. НГУ, 2004. — Вып. 7.
42. *Мельников Г. П.* Системология и языковые аспекты кибернетики. — М., 1978.



43. *Минский М.* Фреймы для представления знаний // Психология машинного зрения. — М., 1978.
44. *Москович В. А.* Информационные языки. — М., 1971.
45. *Мошкович Ж. Г.* Автоматическая лексическая система Унилекс-2. — М., 1989.
46. *Никитина С. Е.* Семантический анализ языка науки: на материале лингвистики. — М., 1987.
47. *Олкер Х. Р.* Волшебные сказки, трагедии и способы изложения мировой истории // Язык и моделирование социального взаимодействия. — М., 1987.
48. *Панков И. П., Захаров В. П.* Информационно-поисковые системы // Прикладное языкознание. — СПб., 1996.
49. *Пиотровский Р. Г.* Инженерная лингвистика и теория языка. — Л., 1979.
50. *Пиотровский Р. Г., Бектаев К. Б., Пиотровская А. А.* Математическая лингвистика. — М., 1977.
51. *Поликарпов А. А.* Теоретические проблемы прикладной лексикологии // Вестн. МГУ. — 1989. — № 5. Сер. 9. Филология.
52. *Попов Э. В.* Экспертные системы. — М., 1987.
53. *Поспелов Д. А.* Ситуационное управление. Теория и практика. — М., 1986.
54. *Поспелов Д. А.* Логико-лингвистические модели в системах управления. — М., 1981.
55. *Прикладные аспекты лингвистики.* — М., 1989.
56. *Проблемы прикладной лингвистики.* — М., 2004.
57. *Промышленные системы машинного перевода.* — Вып. 20. — М., 1991.
58. *Протченко И. Ф.* Словари русского языка. — М., 1996.
59. *Реферирование в общественных науках: Теория и методика.* — М., 1982.
60. *Рождественский Ю. В., Марчук Ю. Н., Волков А. А.* Введение в прикладную филологию. — М., 1998.
61. *Севбо И. П.* Сквозной анализ как шаг к структурированию текстовых знаний // НТИ. Сер. 2. — 1989. — № 2.
62. *Скороходько Э. Ф.* Лингвистические проблемы обработки текстов в автоматизированных ИПС // Вопросы информационной теории и практики. — М., 1974. — № 25.

63. *Субботин М. М.* Новая информационная технология: создание и обработка гипертекстов // НТИ. — 1988. — № 5. Сер. 2.
64. *Федоров А. В.* Основы общей теории перевода. — М., 1968.
65. *Филиппович Ю. Н., Филиппович А. Ю.* Специальность “Компьютерная лингвистика и семиотика” // Интеллектуальные технологии и системы. — М., 2003.
66. *Шаляпина З. М.* Автоматический перевод: эволюция и современные тенденции // Вопросы языкознания. — 1996. — № 2.

## ***ЗМІСТ***

Пояснювальна записка .....	3
Зміст самостійної роботи з дисципліни “Основи інформатики та прикладної лінгвістики” .....	4
Питання до курсу .....	60
Список літератури.....	62

Відповідальний за випуск *А. Д. Везеренко*  
Редактор *А. А. Тютюнник*  
Комп’ютерне верстання *О. М. Бабаєва*

Зам. № ВКЦ-4592

Формат 60×84/16. Папір офсетний.

Друк ротаційний трафаретний.

Наклад 30 пр.

Міжрегіональна Академія управління персоналом (МАУП)

03039 Київ-39, вул. Фрометівська, 2, МАУП

ДП «Видавничий дім «Персонал»

03039 Київ-39, просп. Червонозоряний, 119, літ. ХХ

*Свідоцтво про внесення до Державного реєстру  
суб’єктів видавничої справи ДК № 3262 від 26.08.2008*

Надруковано в друкарні ДП «Видавничий дім «Персонал»