

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Є. В. Купріянов, Н. С. Угольнікова, О. М. Юрченко

СТРУКТУРНА ЛІНГВІСТИКА В ТЕОРІЇ ТА ПРАКТИЦІ

Навчальний посібник
для бакалаврів напряму підготовки
035 «Філологія»
денної форми навчання

Затверджено редакційно-
видавничою радою НТУ «ХПІ»,
протокол № 1 від 15.02.2024 р.

Харків
НТУ «ХПІ»
2024

УДК 81-116
К 91

Рецензенти:

В. А. Широков, акад., д-р техн. наук, проф., директор Українського мовно-інформаційного фонду НАН України;

І. Ю. Шубін, канд. техн. наук, проф., Харківський національний університет радіоелектроніки

Автори:

Є. В. Купріянов, к. філол. н., доц.;

Н. С. Угольнікова, к. філол. н., доц.;

О. М. Юрченко, к. філос. н., доц.;

Купріянов Є. В.

К 91 Структурна лінгвістика в теорії та практиці : навчальний посібник / Є. В. Купріянов, Н. С. Угольнікова, О. М. Юрченко. – Харків : НТУ «ХПІ», 2024. – 256 с.

ISBN 978-617-05-0482-1

Розглянуто теоретико-методичні засади аналізу й опису системно-структурної організації мови та їх застосування на конкретному матеріалі української та іноземних мов. Надано базову термінологію, що використовується в межах структурної лінгвістики як окремому напрямі мовознавства. Всі теми забезпечено прикладами для ілюстрації теоретичного матеріалу, супроводжуються контрольними запитаннями та завданнями для самостійної роботи.

Призначено для студентів напряму підготовки «Філологія» денної форми навчання, які засвоїли курс із загального мовознавства та планують надалі вивчати дисципліни у межах прикладної та комп'ютерної лінгвістики.

Іл. 30. Табл. 25. Бібліогр.: 48 назв.

УДК 81-116

ISBN 978-617-05-0482-1

© Купріянов Є. В., Угольнікова Н. С.,
Юрченко О. М., 2024
© НТУ «ХПІ», 2024

ВСТУП

Навчальний посібник «Структурна лінгвістика у теорії та практиці» має на меті поглибити знання та розвинути навички роботи з мовним матеріалом під час розв’язання теоретичних і практичних завдань. Структурна лінгвістика як окремий розділ мовознавства вивчає мову, аналізуючи її структуру та систему і дає відповіді на такі головні запитання:

1. Які існують наукові підходи до вивчення структури та системи мови?
2. Як співвідносяться система і структура між собою?
3. Як функціонує мова як система і що є результатом її функціонування?
4. Які існують методи аналізу і опису системи та структури мови?

Пропонований посібник призначений для проведення лекцій та практичних занять студентам спеціальності 035 Філологія і рекомендований для тих, хто бажає:

- поглибити теоретичні знання про структурні методи аналізу та опису мови;
- навчитись використовувати отримані знання у розв’язанні лінгвістичних завдань на практиці;
- отримати вміння аналізувати та подавати мовний матеріал у вигляді певної системи та структури;
- опанувати базову термінологію структурної лінгвістики та ознайомитися з доробком провідних лінгвістів.

Весь матеріал посібника розрахований на 16 лекційних та 16 практичних занять та розподілений за 4 тематичними розділами – модулями, кожен з яких включає 3 теми, а одна тема – 3 теоретичні питання. Кожному питанню присвячується окремий підрозділ, що за своєю структурою складається з таких блоків:

- 1) *теоретичний блок*, який містить певну інформацію для вивчення під час лекційних занять;

2) *блок контрольних запитань*, мета якого не лише перевірити засвоєння теоретичних знань, а й сприяти формуванню власних думок і бачення щодо проблеми, яка розглядається в темі заняття;

3) *блок практичних завдань*, що призначені для закріплення теоретичних знань і отримання навичок опрацювання мовного матеріалу на відповідних лінгвістичних завданнях.

Наприкінці подано глосарій, що містить упорядкований за абеткою перелік термінів, використаних в навчальному посібнику. Крім дефініцій також можуть бути присутні перехресні посилання на споріднені терміни, що описані в іншій статті. Такі посилання позначаються ремаркою «Див. також:». Терміни-синоніми мають відсилання на основний термін за допомогою ремарки «Див.». Родові та видові терміни подаються в одній статті. Якщо потрібно знайти видовий термін, наприклад, *породжувальна граматика*, то його слід шукати у статті *Граматика*. Елементами статті є заголовковий термін, дефініція, а також похідні терміни (якщо є) та, за необхідності, приклади використання заголовкових термінів.

Для наочності та зручності засвоєння теоретичного матеріалу в посібнику подаються перехресні посилання на відповідні підрозділи. Також для отримання більш детальної інформації по темі заняття подаються зовнішні посилання на довідкові електронні ресурси. Навчально-методичний посібник розраховано на здобувачів освіти, які засвоїли курс із загального мовознавства та планують надалі вивчати дисципліни у межах освітньо-професійної програми «Прикладна та комп'ютерна лінгвістика» за спеціальністю Філологія.

МОДУЛЬ 1 ІСТОРІЯ СТРУКТУРНОЇ ЛІНГВІСТИКИ

ТЕМА 1 ПЕРЕДУМОВИ ВИНИКНЕННЯ СТРУКТУРАЛІЗМУ

План заняття:

1. Сучасні наукові парадигми у лінгвістиці
2. Структурний підхід у лінгвістиці
3. Основні принципи лінгвістичного структуралізму

Перш ніж переходити до вивчення основних питань структурної лінгвістики, необхідно визначити, що являє собою структурний підхід як метод наукового пізнання (дійсності, певного об'єкта або явища тощо), що зумовило мовознавців на початку ХХ ст. звернутися до структурного підходу до вивчення мови і які саме результати вони хотіли отримати за допомогою цього підходу.

1.1. Сучасні наукові парадигми у лінгвістиці

Сучасний стан лінгвістичної науки (ХХ – ХХІ ст.) можна охарактеризувати одним словом – **поліпарадигматизм**. Як ми можемо побачити, його морфологічну структуру складають префікс *полі-*, корінь *парадигма* та суфікс *-изм*. Отже, що означає це слово і як саме воно характеризує лінгвістичну науку? Так, перша частина *полі-* означає з грецької «велика кількість», «багатоманіття». Остання – це суфікс, що утворює назви суспільно-політичних, наукових та інших напрямків, течій, систем, явищ тощо. А що означає власне корінь **парадигма**? Якщо взяти будь-який лінгвістичний словник, то він нам дасть кілька визначень: 1) клас мовних одиниць, об'єднаний за спільною ознакою або асоціаціями; 2) модель організації класу, репрезентованого тими чи іншими мовними одиницями і 3) у вузькому значенні: сукупність граматичних форм того самого слова, яку ми отримуємо під час відмінювання слів. Очевидно, тут мова йде про **наукову парадигму**, що визначається як *модель організації наукової діяльності, прийняту науковим суспільством*.

Тепер з'ясуємо, що таке **модель наукової діяльності**. Наукову діяльність розуміють як *процес отримання нових знань про той чи інший об'єкт*, а **модель** (ми неодноразово будемо повертатися до цього терміна) у цьому

контексті – *певна система поглядів, що визначає підходи, методи та технічні засоби, які мають використовуватися для отримання нових знань*. Крім цього, така система визначає які аспекти або властивості об'єкта мають бути досліджені. Наприклад, якщо мова йде про описову лінгвістику, то вона має на меті отримати описи мов у певний період їхнього існування, зокрема їхню лексику, фонетику, граматику. Якщо про психолінгвістику, то аспектами вивчення будуть взаємовплив мови, мислення та свідомості.

Підсумовуючи вищесказане, **поліпарадигматизм** позначає *багатоманіття моделей або напрямків наукової діяльності, метою яких є отримання нових знань про мову (сукупності мов) у різних її аспектах із використанням відповідних методів та засобів*. Усі наукові парадигми, що існують у сучасній лінгвістиці, можна об'єднати у чотири групи (див. таблицю 1.1). Перша група включає аспекти, що пов'язані з вивченням мови або групи споріднених мов переважно у їхньому сучасному стані. Предметом другої – порівняння двох або кількох мов, у тому числі діалектів, для виявлення спільних рис у походженні, функціонуванні, словниковому складі тощо. У третій групі об'єднано напрямки лінгвістичної науки, що використовують методи суміжних наук у вивченні мови. Окремим напрямком стоїть прикладна лінгвістика, що має на меті не вивчення мови, а вирішення питань застосування мови у практичній діяльності людини.

Таблиця 1.1. Парадигми сучасної лінгвістичної науки.

«Чиста» лінгвістика	+ порівняння мов	+ суміжні науки	Прикладна лінгвістика
Описова лінгвістика	Порівняльно-історичне мовознавство	Психолінгвістика	Комп'ютерна лінгвістика
Теорія мови	Типологія та універсалії	Соціолінгвістика	Польова лінгвістика
	Контрастивна лінгвістика	Нейролінгвістика	Лінгводидактика
	Ареальна лінгвістика	Етнолінгвістика	Мовна політика

	Діалектологія	Математична лінгвістика	Мовний вплив
		Квантитативна лінгвістика	Міжмовна комунікація

Описова лінгвістика займається описом мовних одиниць різних рівнів, а саме: фонетичного, словотвірного, лексичного та синтаксичного рівнів. Відповідно, виокремлюються відповідні розділи: фонетика, морфологія, лексика та синтаксис. Варто зазначити, морфологія і синтаксис утворюють окремий розділ – граматику. До теорії мови (вона ж теоретичне мовознавство) входять (1) структурна (включаючи європейський структуралізм та американський дескриптивізм), (2) генеративна, (3) формальна, (4) функціональна і (5) когнітивна лінгвістика. Предметом нашої уваги є структурна лінгвістика, окремий напрям у лінгвістичній науці, завдяки якому з'явилися інші напрями (2) – (5).

Контрольні запитання:

1. Які наукові методи / підходи використовуються у різних лінгвістичних науках (таблиця 1.1.) для того, щоб отримати знання про мову?
2. Які саме властивості мовних одиниць (фонема, морфема, лексема, синтаксема) є предметом опису в описовій лінгвістиці?
3. Які знання про мову продукує лінгвістика на стику суміжних наук (див. третій стовпчик таблиці 1.1)?
4. Перелічіть види практичної діяльності, де можуть застосовуватися знання про мову.

Практичні завдання

1. Подайте своїми словами визначення до термінів структурна лінгвістика, генеративна лінгвістика, формальна лінгвістика, функціональна лінгвістика та когнітивна лінгвістика.
2. Визначте, що є спільного і відмінного у таких напрямках, як порівняльно-історичне мовознавство, лінгвістична типологія та контрастивна лінгвістика.

3. Заповніть таблицю про напрямки лінгвістики у її зв'язках з іншими науками, використовуючи додаткові ресурси:

Напрямок	Час виникнення	Засновник напрямку	Які проблеми розглядає?
Психолінгвістика			
Соціолінгвістика			
Нейролінгвістика			
Етнолінгвістика			
Математична лінгвістика			
Квантитативна лінгвістика			

4. Схарактеризуйте зв'язки прикладної лінгвістики з іншими галузями знань:



1.2. Структурний підхід у лінгвістиці

Мова як системно-структурне утворення (див. Тему 3) характеризується внутрішньою організованістю, і її можна аналізувати в кількох аспектах. По-перше, як сукупність елементів, по-друге, як сукупність відношень і, по-третє, як єдину чітку цілу, узгоджену сукупність елементів і відношень. Відповідно, виокремлюють елементарний, системний і структурний (системно-структурний) підходи до вивчення мови як об'єкту пізнання. Розглянемо кожний із них.

Елементарний підхід, який вважається традиційним, як у мовознавстві так і в інших науках. Припустимо, що є певний предмет (рис. 1) певної форми та певного змісту, і завданням дослідника розірвати зовнішню оболонку та розпізнати під нею внутрішній зміст предмета. Так, наприклад, якщо взяти годинник, то його зовнішньою оболонкою є корпус із циферблатом, а його змістом – часовий механізм, який знаходиться у корпусі. Розглянемо приклад із лінгвістики: зовнішньою оболонкою слова є звукова або графічна форма, а змістом – лексичне значення, що закріплене за цією формою. Найповніший опис лексичних значень слів подано у тлумачних словниках. Так само ми можемо взяти будь-яке інше слово і досліджувати його стилістичні, синтаксичні, прагматичні та інші характеристики.



Рис. 1.1. Об'єкт пізнання: зовнішня оболонка (ліворуч) та зміст (праворуч).

Системний підхід. Проте дійсність буває іншою. Переважна більшість предметів, які потребують дослідження, аж ніяк не такі прості і нешкідливі, якими спочатку зображуються на рис. 1.1. Більшість досліджуваних предметів є системними предметами і системними об'єктами. Вони внутрішньо неоднорідні, структурні, постійно змінюються. Якщо більш детально розглянути предмет, наприклад, той самий годинник, як функціонує його механізм, то ми побачимо, що кожна його деталь пов'язана з іншими, що дає йому змогу функціонувати як єдине ціле. Так само ми можемо спостерігати в будь-якій мові, де одиниці різного рівня не існують окремо одна від одної, а перебувають у певних відношеннях, або зв'язках, та залежать одна від одної.

У мові як системі розрізняють **парадигматичні** («вертикальні») та **синтагматичні** («горизонтальні») відношення.

Парадигматичні відношення утворюють ієрархію елементів, і як приклад можна назвати: рід-вид, синонімія, антонімія та омонімія. Такі відношення засновано на групуванні елементів за певною спільною ознакою і одночасно на протиставленні іншим елементам. На фонологічному рівні ми також можемо спостерігати парадигматичні відношення між фонемами на основі їхнього протиставлення, наприклад, за способом артикуляції: проривні, носові, фрикативні, якщо це приголосні; або голосні переднього і заднього ряду тощо. На морфемному рівні парадигматичні відношення включають **аломорфію** (*нос-ити* – *нош-у*), **дублетність** (*сині* – *синій*), **полісемія** (*baker* – *chooser*), **омонімія** (*worker* – *boiler*), **синонімія** (*бунтар* – *бунтівник*; *житомирянин* – *житомирець*), **антонімія** (*зайти* – *вийти*).

Синтагматичні зв'язки (їх ще називають лінійними) характеризують сполучуваність мовних елементів, тобто здатність поєднуватися з іншими елементами у більш складний лінгвістичний об'єкт. Синтагматичні відношення фонем характеризують, наприклад, сполучуваність голосних із приголосними. На морфологічному рівні – як реалізується той чи інший варіант кореня слова, залежно від того, з якою морфемою він сполучається, наприклад: *сказ-а-ти* – *скаж-у*, на відміну від *вез-у*. На синтаксичному рівні – вибір відповідної форми присудка залежно від підмета, а також форми додатка (залежного підмета) пор.: *мама розказала казку* – **мама розкажу казкою*.

Структурний (або системно-структурний) підхід має на меті виявити структурні елементи та встановити зв'язки між ними у межах досліджуваної системи. Спирається на комплекс загальнонаукових, експериментальних, природнонаукових та статистичних методів. У центрі уваги лежить не вивчення поодиноких елементів, а в першу чергу структуру об'єкта та місця елементів у ній. Про системно-структурний характер мови наголошував основоположник

структурного підходу у лінгвістиці Фердинанд де Сосюр¹ у своїй роботі «Курс із загальної лінгвістики»: **мова** – це система, елементи якої утворюють ціле, а значущість одного елемента виникає тільки від одночасної наявності інших. Основними моментами системно-структурного підходу є такі:

- визначення структурних елементів, що складають об'єкт / систему як єдине ціле;
- дослідження системних зав'язків між елементами та закономірності їхнього сполучення.

Контрольні запитання:

1. Як співвідносяться поняття *парадигма* і *парадигматика*; *синтагма* і *синтагматика*?
2. Визначте різницю між поняттями *фон*, *фонема* і *алофон*; *морф*, *морфема* і *аломорф*. Наведіть приклади. Що є спільним між алофонією та лексичним змістом слова у словосполученні або реченні? (контекст)
3. Що саме складає парадигму одиниць фонологічного, словотвірного, лексичного та синтаксичного рівнів?
4. Наведіть приклади елементарних синтагматичних відношень, що можуть мати місце на фонетичному, словотвірному та синтаксичному рівнях.
5. Наведіть приклади складних елементів мови, які можна побудувати на синтагматичних відношеннях.

Практичні завдання

1. Нижче перелічено характеристики парадигматичних і синтагматичних відношень, але у змішаному порядку. Визначте їх належність до певного типу відношення та подайте їх у таблиці:

характерним є взаємовиключення («або – або»), угруповання чи асоціація, «горизонтальні», лінійність (послідовність), засновані на виборі, ніколи не зустрічаються у мовленні, неспостережувані, характерна спільна зустрічальність («і – і»), засновані на комбінаториці (сполучуваності

¹ Ще раніше про системний характер мови стверджував Вільгельм фон Гумбольдт: «Для того, щоб людина могла досягнути хоча б одне-єдине слово..., вся мова повністю і у всіх своїх взаємозв'язках вже має бути закладена в ній».

елементів), спостережувані, реалізуються у мові й мовленні (у тексті), «вертикальні».

Синтагматичні відношення	Парадигматичні відношення
<i>Властивості</i>	
1.	1.
2.	2.
3.	3.
4.	4.
5.	5.
6.	6.

2. Розподіліть у таблиці подані нижче характеристики відношень у системі мови:

парадигматичні (в основному), включення та входження, не можуть поєднуватися один з одним, ієрархічні, синтагматичні, можуть поєднуватися один з одним.

Відношення між одиницями мовної системи	
між одиницями різних рівнів	між одиницями одного рівня

3. Наведіть елементи парадигми для різних мовних одиниць англійської та української мов:

- фонеми [a]
- дієслова *to work, працювати*
- лексеми: *to go, ходити*;
- речення: *The workers build a house* та *Робочі будують дім*.

1.3. Основні принципи лінгвістичного структуралізму

У мовознавчій літературі виокремлюють доструктурну лінгвістику (кінець XIX ст. – початок XX ст.) і власне структурну лінгвістику (починаючи з 20-х рр. XX ст.). Перша ставила за мету аналіз поодиноких лінгвістичних фактів та їхні **експліцитні** (чітко виокремлювані, безпосередньо спостережувані,

виражені наявним чином) властивості – фонетичні, семантичні та інші. Матеріальний прояв таких властивостей пояснювалися впливом певних історичних, соціальних та психологічних чинників. Представниками дотрунтурної лінгвістики (XIX ст.) є **молодограматики** – *німецька мовознавча школа, що запропонувала природно-наукові принципи для верифікації фактів мови*. Основні роботи присвячено порівняльно-історичному мовознавству та індоєвропеїстиці. Характерними рисами молодогограматизму є:

- дослідження фізіологічних і акустичних умов утворення звуків, завдяки чому фонетика стала спеціальною науковою дисципліною;

- використовуючи фонетичний підхід до вивчення історичних пам'яток, лінгвістам вдалося отримати повну картину розвитку індоєвропейських мов та уточнити в цих мовах фонетичні закони;

- наукові відкриття в галузі граматики – низки фонетичних процесів (**перерозкладення й опрощення, аналогічне вирівнювання форм**), які дали більш глибоке розуміння природи флексії, а також засобів вираження граматичного значення, таких як порядок слів, службові слова, наголос та інтонація.

Головною рисою молодогограматизму було те, що створювана ученими емпірична база, тобто мовленнєвий матеріал, прості спостереження за особливостями індивідуального мовлення, ототожнювалася зі самою мовою, що робило неможливим розрізняти об'єкт і предмет дослідження. Лінгвісти цього напрямку прагнули атомізувати факти, іншими словами досліджувати окремі фонетичні явища або окремі граматичні форми. Кожен окремий мовний факт вивчався з погляду до індивідуальних особливостей уживання слів та форм, зокрема їхня семантика, яка отримувала психологічну інтерпретацію. У цілому можна констатувати, що молодогограматизм характеризується зростанням кола об'єктів дослідження, багатоаспектністю вивчення мовних фактів, прагненням надати цим фактам психологічну інтерпретацію, історичним спостереженням за ними та відсутністю єдності теоретичних поглядів на мову.

Безперечно, що молодограматика зробили великий внесок у розвиток мовознавчої науки, зокрема в галузі порівняльно-історичного мовознавства. Проте молодограматизм як науковий напрямок на початку XX ст. почав зазнавати кризи через певні недоліки, головними з яких є:

- атомізм, що виявився у недостатньому врахуванні системної взаємозумовленості мовних явищ;

- індивідуально-психологічне розуміння мови, що вело до заперечення реальності спільної для певного народу мови, до оголошення загальної мови науковою фікцією;

- посилена увага до **діахронічного** (історичного) опису мови і водночас зневажливе ставлення до опису мови з погляду **синхронії** (сучасного стану).

Найбільш плідною наприкінці XIX ст. видалася критика молодограматизму з боку Миколи Крушевського та Івана Бодуена де Куртене, а на початку XX ст. – Фердинанда де Сосюра. Зокрема учені запропонували ідеї опису та вивчення системних закономірностей у мові, тобто системного опису мови безвідносно до її історичного розвитку. Так, швейцарський лінгвіст Фердинанд де Сосюр зробив першу спробу встановити взаємозв'язки між фактами мови, угрупувати їх та синтезувати в єдине ціле. На основі ідей де Сосюра та інших учених щодо системного, або більш точно системно-структурного опису мови, виник новий напрям – **лінгвістичний структуралізм**. Розвитку нового напрямку сприяли також відкриття в галузі природничих наук, наприклад: таблиця періодичних елементів Дмитра Івановича Менделєєва, закон Грегора Менделя про розщеплення спадкових ознак батьків та їх нащадків тощо. Завдяки ним вводиться поняття **дискретної структури**, тобто *матеріальної структури, яку складають окремі елементи*. Крім цього, за аналогією до того, що матеріальну оболонку фізичних та хімічних явищ складають атоми та молекули, лінгвісти вперше вводять поняття фонем та морфем, які складають матеріальну оболонку мови.

Отже, **лінгвістичний структуралізм** можна визначити – як *напрямок, що розглядає мову як чітко структуровану систему, елементи якої перебувають*

між собою у певних відношеннях та набувають свого місця у системі завдяки цим відношенням. В основі поглядів на мову та методи її дослідження лежить:

- розуміння мови як знакової системи з чітко виокремленими структурними елементами (тобто одиницями мови, їх класами, групами тощо);
- прагнення до чіткого формального опису системи мови, як це відбувається у точних науках.

Засновником структурного напрямку лінгвістики вважається швейцарський лінгвіст Фердинанд де Сосюр, основні ідеї якого він виклав у своїй праці «Курс із загальної лінгвістики» (1916 р.). Варто зазначити, що ідею системності мови ще раніше висловив німецький мовознавець Вільгельм фон Гумбольдт: у мові немає нічого одиничного, кожний факт мови треба розглядати в його зв'язках. Проте основні положення системно-структурного опису було висвітлено де Сосюром то його послідовниками Альбером Сеше та Шарлем Балі. Основні положення системно-структурного підходу до дослідження мовних явищ спрямовано на: 1) виокремлення як об'єкта досліджень такої організації елементів у сукупності, що репрезентує цільну єдність; 2) виявлення внутрішніх законів об'єднання елементів у ціле, механізму інтеграції; 3) встановлення залежностей та відношень, за рахунок яких елементи об'єднуються в єдине ціле (підсистеми або мікросистеми). Висвітливо основні принципи лінгвістичного структуралізму (ПЛС).

ПЛС 1. *Справжньою та основною реальністю є не окремий факт певної мови, а мова як система.* Це означає, що кожний елемент мови існує лише через його відносини до інших елементів у складі системи; система не підсумовується з елементів, а, навпаки, визначає їх.

ПЛС 2. *Кістяк, структуру системи створюють позачасові відносини.* Такими є відношення опозиції елементів поза текстом та відношення взаємного розташування (дистрибуції) елементів у тексті. Відношення у межах системи домінують над елементами.

ПЛС 3. *Стає можливим позачасове квантитативне («алгебраїчне») вивчення системи мови, засноване на відносинах, а не на особливості елементів*

або їх матеріальності; можливе застосування чітких, математичних методів у лінгвістиці.

ПЛС 4. *Мова є система особливого роду – знакова система, що існує, з одного боку, об'єктивно, тобто поза психікою людини, у міжособистісному спілкуванні людей, з іншого боку – ця система існує й у психіці людей.*

ПЛС 5. *Подібно до мови так само організовані деякі інші системи, що діють у людських суспільствах, – фольклор, звичаї та ритуали, відносини спорідненості тощо; всі вони можуть вивчатися, подібно до мови, лінгвістично, зокрема бути досить строго формалізуватися (семіотика).*

Отже, вимогами до структуралістського підходу до вивчення, опису та репрезентації мови є:

- застосування більш чітких методів дослідження мови, аналогічних методам логіки, математики, фізики та інших точних наук, які виключають звернення до позамовних та позасистемних чинників, таких як психологічні, соціологічні, естетичні, географічні тощо;

- визнання цілісності мови та характеристика кожного окремого її елемента шляхом указівки на його зв'язок з іншими елементами мовної системи, на його місце у системі, а не шляхом перелічення його властивостей;

- опис мови як складної системи, що має багаторівневу будову та розрізнення у ній рівня (рівнів) знакових та рівня (рівнів) незнакових одиниць;

- розповсюдження методів опису **знакових одиниць** на опис **незнакових одиниць мови** (наприклад, фонем);

- розрізнення парадигматичних відношень (відношення між одиницями вищого та нижчого рівнів) та синтагматичних відношень (відношення сполучуваності між одиницями, що входять до складу певної конструкції);

- перевага віддається синхронічному аспекту опису (опису мови у її певному етапі розвитку), а не до діахронічного аспекту (зіставлення різних етапів розвитку мови).

Контрольні запитання:

1. Визначте сильні сторони молодогограматизму як лінгвістичного напрямку

2. Які саме чинники привели до кризи молодогограматизму?
3. Порівняйте предмет дослідження у молодогограматичному і структурному напрямках лінгвістики.
4. Мову визначають як знакову систему? А що є власне мовним знаком у системі?
5. Що таке граматичні класи слів? На якому ґрунті слова об'єднуються у різні граматичні класи?
6. Що таке лексичні класи слів? Як ви розрізняєте поняття *лексичний склад мови* та *лексична система мови*?

Практичні завдання

1. Наведіть у таблиці нижче характерні риси молодогограматизму та структуралізму:

№	Молодограматизм	Структуралізм
1		
2		
3		
4		
5		

2. Визначте місце голосних: українських [a], [o], [y], [e], [u], [i], німецьких [i:] та [ö], а також англійських [ɑ:], [ə:], [ə], [ɔ], [ʌ] та [æ] за вказаними ознаками:

Піднесення	Ряд					
	Високий		Середній		Низький	
	Нелаб.	Лабіал.	Нелаб.	Лабіал.	Нелаб.	Лабіал.
Високе						
Середнє						
Низьке						

3. Визначте місце елементів у лексичній системі української мови:

бик, бичок, вовк, вовчиця, вовчєня, корова, собака (чол. р., жін. р.), пес, тапір, тєлиця, тєля (чол. р.), щєня.

‘adult’		‘young’	
‘male’	‘female’	‘male’	‘female’

4. Визначте місце слів на позначення родинних відносин у лексичній системі англійської мови:

spouse, mother, father, sister, sibling, son, son-in-law, daughter, daughter-in-law, step-son.

‘close relative’		‘distant relative’	
‘male’	‘female’	‘male’	‘female’

ТЕМА 2 ШКОЛИ СТРУКТУРАЛІЗМУ ТА ЇХ РОЛЬ У СУЧАСНІЙ ЛІНГВІСТИЦІ

План заняття:

1. Празька лінгвістична школа
2. Копенгагенський структуралізм
3. Американський дескриптивізм

У лінгвістиці проведено численні дослідження мови та вироблено багато концепцій щодо її складу, будови та функціонування. Але в цьому розділі ми розглянемо лише ті концепції, які мали вирішальне значення для сучасної науки та які застосовуються сьогодні на практиці, наприклад у сфері комп'ютерної лінгвістики під час розроблення різних систем опрацювання природної мови. Провідними школами структурної лінгвістики, що зробили значний внесок у розвиток сучасної лінгвістики, є празький лінгвістичний гурток, копенгагенська школа структуралізму та американський дескриптивізм. Розглянемо нижче їхні напрацювання.

2.1. Празька лінгвістична школа

Празька школа, або функціональний структуралізм, – один із провідних напрямків лінгвістичного структуралізму, що склався в Чехословаччині та Австрії у другій половині 20-х рр. ХХ ст. та організаційно оформився у 1928 р. Керівником празької школи став Вілем Матезіус, а його постійними членами – чеські мовознавці Богуміл Трнка, Богуміл Гавранек і Ян Мукаржовський. Пізніше долучилися Володимир Скалічка, Йозеф Вахек, Микола Трубецькой, Роман Якобсон та Сергій Карцевський. Представники осередку додержувалися структурного підходу та визначали мову як *систему засобів вираження, тобто як функціональну систему, що має цільову спрямованість*. Звідси й друга назва школи – функціональний структуралізм. Розглянемо нижче основні ідеї, наукові проблеми та досягнення празької лінгвістичної школи (ПЛШ), які залишаються актуальними у наш час.

ПЛШ 1. *Інтерпретація мови як функціональної системи, тобто як системи засобів вираження, призначених для досягнення певної мети.* Аналіз

мовної діяльності як засобу спілкування показує, що найпростішою метою мовця є висловлення. Тому до лінгвістичного аналізу слід підходити з функціональної точки зору. Жодне явище у мові не може бути зрозуміле без урахування системи, до якої ця мова належить. У ряді випадків функціональний підхід сприймається як *орієнтація на ту роль, яку певна одиниця грає у складі більшого цілого* (або у складі одиниці вищого рангу). У розумінні пражських структуралістів термін **функція** використовується тоді, коли мова йде про значення (функція слова, функція речення) або про структуру смислових одиниць (функція фонем). Так, наприклад, слово виконує свою функцію в складі речення, реалізуючи своє значення і формуючи зміст речення у цілому. Візьмемо для порівняння такі речення зі словом *albatross*:

- (1) *The large wandering **albatross**, for example, can live for up to 80 years.*
- (2) *The national debt is an **albatross** around the president's neck.*
- (3) *I flew up high like an **albatross**, and then I saw you on the ground.*
- (4) *Her own supporters see her as an **albatross** who could lose them the election.*

Аналізоване слово у системі англійської мови може реалізовувати два значення. Перше: «a large white bird with long, strong wings that lives near the sea, found especially in the areas of the Pacific and South Atlantic oceans», а друге: «Something or someone you want to be free from because that thing or person is causing you problems». Яке з двох значень реалізує лексема у поданих реченнях? Розглянемо контекстні оточення:

- (1) *The large wandering **albatross**...*
- (2) *The national debt is an **albatross**...*
- (3) *I flew up high like an **albatross**...*
- (4) *Her own supporters see her as an **albatross**...*

У контекстах (1) і (3) *albatross* є самостійним словом, а в контекстах (2) і (4) воно зустрічається одночасно із прийменником *as* та неозначеним артиклем *an*, і крім цього воно утворює стійкі сполучення *to be an albatross (for smb.)*, *to see smb. as an albatross*. Отже, у контекстах (1) і (3) аналізоване слово реалізує

значення (1), а в (2) та (4) – друге значення. Варто зазначити, що для реалізації цих значень та, відповідно, здійснення комунікативної мети, використовуються різні структури речень. Щодо практичного аспекту, комп'ютерна система перекладу аналізує структуру речень перед тим, як визначити значення слова і передати зміст усього речення мовою перекладу.

ПЛШ 2. *Уведення поняття **фонологічних опозицій**, визначення **фонем** як «пучка» ознак, розроблення типології фонологічних опозицій.* Використовуючи системно-функціональний підхід до аналізу звукової системи мови, празькі структуралісти наочно показали, що у мові важливими є не звуки і значення самі по собі, а відношення між звуками й значеннями, так що матеріальний зміст фонологічних елементів є менш суттєвим, ніж їхній взаємозв'язок усередині системи. Микола Трубецькой увів поняття опозиції і на його основі довів, що два різні звуки можуть в одній мові виступати як різні фонем, а в іншій – як одна. **Фонологічні опозиції** інтерпретуються як *звукове протиставлення, яке може диференціювати значення слів*. Усе залежить від системи протиставлень (опозицій), яка має місце в кожній мові. Головним є не те, що звуки утворюють систему, а те, що у фонологічних дослідженнях потрібно виходити із системи. Так, наприклад, за допомогою опозиції звуків англійської мови, можемо розрізняти слова. Порівняйте звуки <i> та <i: > для розрізнення слів *grin* (посміхатися) та *green* (зелений); або <θ> та <s> для розрізнення лексем *think* (думати) та *sink* (тонути). У практичному плані, розроблена празькими структуралістами система фонологічних опозицій використовується, наприклад, у комп'ютерних системах перевірки вимови та у системах набору текстів за допомогою голосу.

ПЛШ 3. *Вироблення системи граматичних опозицій.* Поняття і методи, опрацьовані на фонологічному матеріалі, згодом були застосовані пражцями для дослідження інших рівнів мови. Так, наприклад, Роман Якобсон застосував методику опозиційного аналізу в дослідженні граматичних категорій та обґрунтував положення про бінарність граматичних опозицій. Граматичні значення виявляються у протиставленнях (наприклад, значення однини,

протиставлене значенню множини). Саме граматичні опозиції (протиставлення) утворюють системи, які називають граматичними категоріями. **Граматичну категорію** можна визначити як *низку протиставлених один одному однорідних граматичних значень, систематично виражених формальними показниками*. Так, наприклад, в українській мові граматичними категоріями є відмінок, рід (чоловічий, жіночий, середній), стан (активний, пасивний) і число (однина, множина і двоїна). В іншій сучасній лінгвістичній літературі виокремлюють tense, number (число), gender (рід), noun classes (класи іменників) та locative relations (локативи). Кожна граматична категорія має відповідні граматичні значення, які називаються **грамемами**. Різні грамеми однієї категорії виключають одна одну і не можуть бути виражені разом одночасно. Так, наприклад однина та множина є грамемами категорії числа, і одна з них має бути реалізована обов'язково. Грамема може бути репрезентована низкою форм, спільних за значенням елемента, але різних за значеннями інших категорій. Так, наприклад: *робиш, робив, робила, робитимеш, робив би, роби* – грамеми дієслова *робити*, що об'єднуються за спільним граматичним значенням другої особи однини, але відрізняються за іншими значеннями способу, часу, виду, стану і особи. Система граматичних опозицій є основою для створення систем перекладу, навчання, але у першу чергу – це перевірка правопису (спел-чекери).

ПЛШ 4. *Теорія актуального членування речень.* Помітним внеском Празької лінгвістичної школи в синтаксичну теорію є вчення Вілема Матезіуса про актуальне членування речення. Функціональний синтаксис передбачає вивчення засобів і способів об'єднання слів у речення в межах певної ситуації мовного спілкування (комунікативної ситуації). Актуальне членування речення з'ясовує метод його включення у предметний контекст, з урахуванням якого воно виникло. Учений висловлює думку про принципову відмінність між формальним членуванням речення (виділення підмета й присудка), яке показує його граматичну структуру, і поділом речення на **тему** і **рему**, який виявляє його «функціональну перспективу». **Актуальне членування речення** являє

собою членування речення за змістом на дві змістові частини: перша називає *предмет мовлення* (тема) й *те, що про нього говориться* (рема), тобто якусь інформацію про цей об'єкт. Друга частина є головною. Наведемо деякі приклади:

Традиційно у нашому місті відбувається етап чемпіонату України з автомобільного слалому; де темою є: *Традиційно у нашому місті*, а ремою – *відбувається етап чемпіонату України з автомобільного слалому*;

Italians are deciding whether to choose their most right-wing government; тема: *Italians are deciding*; рема: *whether to choose their most right-wing government*;

У наведених вище прикладах ми можемо побачити, що така побудова речення забезпечує зв'язність тексту. З одного боку, кожне наступне речення будується на основі попереднього, а з другого – містить у собі нову інформацію. Варто зазначити, що не кожне речення можна розкласти на тему і рему. Іноді зустрічаються речення, що не містять теми, а лише рему: *Прийшло літо. Стало тепло.*

ПЛШ 5. Питання типології мов і мовних союзів. Серед оригінальних ідей пражців – теорія Миколи Трубецького щодо мовних контактів і мовного союзу, тобто *об'єднання мов за подібністю, що зумовлена контактуванням, взаємовпливами*, й типологічні ідеї Володимира Скалічки, котрий обґрунтував необхідність структурного порівняння мов незалежно від їхніх генетичних зв'язків. Наявність одного з типологічних явищ (флексії, аглютинації тощо) не виключає наявності інших, однак вони можуть неоднаково комбінуватись у різних мовах, а це якраз і характеризує відповідні мови. Можна говорити про суто флективний, суто аглютинативний тип і т. д. як про деякі ідеальні еталони, яким різною мірою відповідають конкретні мови. Володимир Скалічка вказав також на те, що морфологічна структура – не єдина основа для типології. Він один із перших проводив зіставні дослідження фонологічних структур різних мов. Лінгвістична типологія мов включає **фонологічну** (подібність або різниця мов за звучанням), **морфологічну** (подібність або різниця мов за утворенням

граматичних форм) та **синтаксичну** (подібність або різниця мов за ознаками семантичної й формальної будови головних членів речення) класифікації.

ПЛШ 6. *Теорія асиметрії мовного знаку Сергія Карцевського*, основна ідея якої зводиться до того, що «знак і значення не покривають один одного повністю. Їх межі не збігаються в усіх точках: один і той самий знак має декілька функцій, одне й те саме значення виражається декількома знаками. Будь-який знак є «омонімом» і «синонімом» одночасно. Непаралельність плану вираження і плану змісту, яка полягає в тому, що:

1) план вираження є лінійним і дискретним, а план змісту має кумулятивну властивість, тобто характеризується структурною глобальністю і часовою безперервністю;

2) один експонент (позначувальне) може мати декілька позначуваних, тобто існують відношення «одне – декілька» (полісемія) і «декілька – одне» (явища омонімії, синонімії й **синкретизму**); **синкретизм** розуміють *явище, коли в одній формі одночасно об'єднано кілька значень або компонентів значень* (наприклад, *бабах* може означати «удар» і звук «звук удару»);

3) позначувальне і позначуване характеризуються автономністю розвитку. План вираження може змінюватися при незмінності плану змісту (*бт>чела* → *бджола*) і навпаки (*міщанин* «житель міста» → «назва соціального стану людей»).

Контрольні запитання:

1. Перелічіть усі відомі вам ознаки фонем, за якими вони утворюють фонологічні опозиції. Наведіть приклади.

2. За якими формальними показниками можна виокремити граматичні класи слів (іменник, прикметник, дієслово тощо)?

3. Які граматичні опозиції можуть утворювати слова у межах тих чи інших граматичних класів (іменник, прикметник, дієслово тощо)? Наведіть приклади.

4. Чи існують, на вашу думку, формальні ознаки, за якими можливо виокремити в реченні тему і рему? Можна за приклад узяти будь-яку відому вам мову.

5. Які одиниці мовного рівня мають план вираження і план змісту? Наведіть приклади.

Практичні завдання

1. За рахунок чого можна розрізняти схожі за звуковим складом слова у таких логопедичних скоромовках, і які диференційні ознаки реалізують фонemi у цих словах:

1. У дворі собача будка, в Петрика співуча дудка. Рекс почує звуки дудки, зразу вибігає з будки (Ю. Рібцун).

2. Заплетена в Катрусі кіска. Готова йти на пашу кізка. Напаслась кізонька й брикає, а поруч дівчинка стрибає. Розвеселилась дуже кізка і розплелась в Катрусі кіска. (Ю. Рібцун).

3. Каже Килинка: – Калинка – не малинка! – Так, так, Килинко, гірка калинка! (Г. Бойко).

4. За рецептом на базарі Дятел вибрав окуляри, натягнув собі на ніс, полетів трудитись в ліс. У осінню хмуру пору хитрий жук зашивсь під кору, дятел глянув, сів на сук, носом тук – і згинув ЖУК. Ось які-то окуляри вибрав дятел на базарі. (М. Стельмах).

2. Визначте, за якими лінгвістичними ознаками можна об'єднати подані іменники у відповідні лексико-граматичні розряди:

книга, бібліотека, автомобіль, світлофор, метро, вулиця, людина, штурман, сміливець, красуня, хатня робітниця, комаха, Василь, Карпати, золото, нафта, сіль, доброта, молодість, темрява, чорнота, білизна, біг, плавання, сон, зліт.

book, library, car, traffic light, subway, street, man, navigator, brave, beauty, domestic worker, insect, Basil, Carpathians, gold, oil, salt, kindness, youth, darkness, blackness, underwear, running, swimming, dream, take off.

Buch, Bibliothek, Auto, Ampel, U-Bahn, Straße, Mensch, Navigator, Draufgänger, Schönheit, Hausmädchen, Insekt, Vasyl, Karpaten, Gold, Öl, Salz, Freundlichkeit, Jugend, Dunkelheit, Schwärze, Leinen, Laufen, Schwimmen, Schlaf, Abheben.

3. Порівняйте наведені нижче дієслівні форми польської мови та їхні українські відповідники. Вкажіть формальні ознаки, які відображають граматичне значення особи, числа, часу? Як виявляється у тому й іншому випадку значення чоловічого роду?

Czytałem	Я читав
Czytałam	Я читала
Czytałeś	Ти читав
Czytałaś	Ти читала
Czytał	Він читав
Czytała	Вона читала
Czytaliśmy	Ми читали
Czytaliście	Ви читали
Czytali	Вони читали

4. Наведіть власні приклади, які відображають асиметрію мовного знаку, дотримуючись таких умов:

- а) одна форма знаку включає кілька значень;
- б) одне й те саме значення може бути відображено кількома формами;
- в) дві однакові форми мають різний зміст.

5. Побудуйте речення відомою вам мовою, у яких:

- а) тему / рему складають: одне слово; два-три-чотири слова; складно-підрядне речення;
- б) можуть бути кілька тем і рем.

2.2. Копенгагенський структуралізм (глосематика)

Копенгагенський структуралізм – лінгвістична течія структуралізму, яка мала на меті створення універсальної лінгвістичної теорії; трактує мову як абстрактну структуру й описує її суто формальними способами без звертання до її субстанцій (реального змісту і звучання). Інші назви цього наукового напрямку – датський структуралізм, або **глосематика** (від грецької *glossema*, род. відм. *glossematos* «слово»). Тривалий час був найвпливовішим і найавторитетнішим структуральним напрямом. Виник у 1931 р. внаслідок

об'єднання датських мовознавців у лінгвістичне товариство, так званий Копенгагенський осередок. Засновником осередку був Луї Єльмслев – директор Інституту лінгвістики і фонетики при філософському факультеті Копенгагенського університету. Крім нього, до осередку входили Вігго Брьондаль, Ганс Ульдалль, Кнут Тогебю, Ганс Сьоренсен.

Глосематика прагне побудувати певну універсальну лінгвістичну теорію, спираючись на те, що ця назва поєднує у своїй структурі слова *глоса* та *математика*. Отже, мова сприймається як система відношень, абстрагованих від «субстанції», від матеріальної природи мовних одиниць. Мета теорії полягає в тому, щоб вивчити структуру відношень у мові. При цьому мова інтерпретується як мережа функцій, а її елементи – як пучки функцій. У глосематиці визначено лише три типи залежностей:

1) **Інтердепенденція** (пор. англ. interdependence «взаємозалежність») – двостороння залежність між елементами в тексті чи системі. Умовне позначення $x \leftrightarrow y$. Тобто елемент x наявний тоді й тільки тоді, коли є елемент y . Приклади елементів, які не існують один без одного: голосні та приголосні (залежність у системі); категорії відмінка та числа іменників в українській та латинській мові (залежність у тексті). Наведемо додаткові приклади, в яких зустрічається взаємозалежність елементів:

П1. Якщо в реченні є підмет та присудок, то вони обов'язково мають узгоджуватися один з одним: Лінгвістика сформувалася в середині XIX століття.

П2. Якщо в мові є категорія однини, то в мові також є категорія множини і навпаки: *стіл – столи; столи – стіл* (ми не розглядаємо випадки *singularia tantum* і *pluralia tantum*);

П3. Якщо в мові є непрямі відмінки, то існує й умовний «називний» відмінок (прямий відмінок), і навпаки. Як приклад, можна навести відмінки у слов'янській та німецькій мовах

П4. Якщо в слові два склади, то у ньому є два голосних звуки, або навпаки (*різ-ний, вер-теп, би-ти*).

2) **Детермінація** (порівн. англ. *determination* «визначення») – *одностороння залежність* – має місце між двома елементами, один з яких передбачає існування іншого, але протилежне невірно. Умовне позначення $x \rightarrow y$. Тобто елемент x наявний тоді, коли є елемент y , а не навпаки. Приклади детермінації на рівні словотвору: суфікс передбачає наявність основи, але не навпаки, бо відомі безсуфіксальні основи (*око, ліс* та інших.). Детермінація також характерна для прийменників української мови *з, для, біля*, які вимагають родового відмінка іменників (*з вікна, для дому, біля дороги*). Протилежне твердження не є вірним, бо форми родового відмінка іменників поєднуються і з іншими прийменниками, а також вживаються без прийменника. Наведемо ще приклади, які ілюструють цей тип залежності:

П1. Приголосний у межах складу передбачає голосний, а не навпаки. (*брати*).

П2. Якщо в системі відмінків є давальний, то в ній обов'язково має бути і називний; зворотне є необов'язковим. Таке твердження є справедливим для слов'янських та німецької мов. Але не для англійської, тому що в ній давального відмінка немає.

П3. У структурі речення наявність прямого додатка передбачає наявність присудка, тоді як зворотне твердження є необов'язковим. Порівняйте речення: *я бачу море. Я бачу. Я думаю.*

3) **Консте́ляція** (пор. англ. *constellation* «сузір'я») – *вільна залежність*, що має місце між двома елементами, кожен з яких може існувати без іншого та не обумовлює іншого. Умовне позначення $x - y$. Прикладом такої залежності є узгодження форми дієслова: у теперішньому часі ця форма узгоджується тільки із особою виконавця дії (*я пишу, ти пишеш, він, вона, воно пише*), а у минулому часі – лише із родом актанта (*хлопець писав, дівчина писала, перо писало*). Консте́ляція є також у разі прийменників з подвійним керуванням відмінками. Аналогічно, в українській мові прийменники *в, на, про* поєднуються зі знахідним та прийменниковим відмінками (*у гору, в горі, об стіл, про стіл*), а *за* – зі знахідним і орудним (*за будинок, за будинком*).

Додатковим прикладом констеляції є сполучання дієслів із різними частинами мови:

- *говорити швидко* (прислівник);
- *говорити, не перестаючи* (дієприслівник);
- *говорити швидше* (прикметник у порівняльному ступені);
- *говорити на рівних* (прийменникова конструкція).

Позитивним у копенгагенському структуралізмі було прагнення опрацювати точний метод аналізу, який спирається на дані математичної логіки. Глосематики розширили поняттєвий апарат мовознавства, запропонували цінні методологічні принципи, які ми застосовуємо сьогодні у комп'ютерній лінгвістиці. Науковці вважають, що глосематика може бути корисною для створення формальних універсальних мов (мов-посередників, метамов) для машинного перекладу і для створення типологічної класифікації мов. Ця теорія отримала продовження і розвинена Хомським під час побудови теорії універсальної граматики.

Контрольні запитання:

1. У чому полягає різниця між празьким лінгвістичним осередком та копенгагенським структуралізмом?

2. У чому полягає сутність універсальної лінгвістичної теорії?

3. Чи можливо, на вашу думку, виробити універсальну теорію мови? Обґрунтуйте свою думку.

4. Що таке універсальна мова, на вашу думку? Які в неї можуть бути обмеження?

Практичні завдання

1. *Спробуйте виявити залежності (інтердепенденція, детермінація та констеляція) у наведеному нижче тексті:*

Садок вишневий коло хати, // Хрущі над вишнями гудуть, // Плугатарі з плугами йдуть, // Співають ідучи дівчата, // А матері вечерять ждуть.

Сем'я вечерея коло хати, // Вечірня зіронька встає. // Дочка вечерять подає, // А мати хоче научати, // Так соловейко не дає.

Поклала мати коло хати // Маленьких діточок своїх; // Сама заснула коло їх. // Затихло все, тільки дівчата // Та соловейко не затих.

1а. Які формальні ознаки обумовлюють наявність тих чи інших залежностей?

2. З якими елементами пов'язані виділені дієслова у поданому реченні? З якими іншими частинами мови вони можуть поєднуватися в реченні? Наведіть конкретні приклади:

Jeremy Hunts **tells** the BBC he **needs** to **show** the world the numbers add up, but warns difficult decisions on tax and spending lie ahead.

2а. Проаналізуйте відповідні дієслова в інших мовах, наприклад, українській.

3. З якими відмінками поєднуються прийменники **an, auf, für, hinter, in, über, neben, unter, vor, zwischen**? Продемонструйте на конкретних прикладах.

2.3 Американський дескриптивізм

Американський дескриптивізм є одним із напрямків американського мовознавства, що виник і активно розвивався у 30 – 50-х рр. XX ст. у загальному руслі структурної лінгвістики. Дескриптивна лінгвістика склалася під впливом ідей Леонарда Блумфілда, який застосовував до індіанських мов строгі методи порівняльно-історичного мовознавства, що засновані на визнанні регулярності фонетичних змін та фонетичних відповідників між спорідненими мовами. Прийшов до необхідності вироблення нової методики опису цих мов, які були здебільшого безписемними та не мали ані граматик, ані словників,.

Лінгвістична концепція Леонарда Блумфілда вперше була викладена у статті «Низка постулатів для науки про мову» (1926). Основною ж роботою вченого є книга «Мова» (1933). У роботі викладаються основоположні принципи лінгвістичного опису, оскільки ученого цікавили принципи саме лінгвістичного опису, а не конкретні прийоми та методи вивчення мови. Звідси й назва – дескриптивна лінгвістика. Науковець прагнув зробити лінгвістику точною наукою. На формування концепції дескриптивізму вирішальною мірою вплинула дослідницька практика вчених США, які вивчали мови і культуру

американських індіанців. Ці мови могли описуватися лише синхронно, дослідники не мали жодних даних про їх історію. Великі труднощі викликало членування текстів на слова. Незрозумілими виявлялися численні граматичні і лексичні значення. Складність розуміння текстів також була зумовлена культурними особливостями аналізованих мов. У результаті, лінгвіст змушений постійно звертатися з питаннями до його носія-інформанта. У процесі дослідження незнайомих мов, коли значення мовних форм лінгвісту невідомі, встановлення і розрізнення одиниць мови виникла необхідність визначення формального критерію – сполучуваність одиниць, їхнє місце відносно інших одиниць. Такий критерій отримав назву дистрибуції (англ. distribution – розподіл).

Дескриптивна лінгвістика не ставила завдання створити загальну лінгвістичну теорію, яка б пояснювала явища мови та їх взаємозв'язку, але розробляла методи синхронного опису та моделювання мови (хоча важливість історичних досліджень не заперечувалася). Детально розроблялися проблеми рівнів дистрибутивного аналізу, рівнів структури мови та відповідних їм основних одиниць – фонем, морфем та (іноді) конструкції (або речення). Було поставлене питання про проміжні рівні, наприклад, морфонематичному та його одиниці – **морфонемі**.

У межах пропонованої теорії американські дескриптивісти виокремлюють одиниці мови за таким принципом:

- Слово як основна одиниця мови, як правило, не виділялося і трактувалося як особливо тісно спаяний ланцюжок морфем, з'єднаний у межах речення з іншими такими ж ланцюжками.

- Одиниці, більші за речення, не розглядалися, оскільки вони вважалися не належними до структури мови.

Дистрибутивний аналіз здійснюється у напрямку «знизу вгору», спочатку на фонологічному, а потім на морфологічному. Методика аналізу в дескриптивній лінгвістиці включає на фонологічному та морфологічному рівнях одні й ті самі основні етапи та операції (див. рис. 2.1):

1-й етап – членування тексту на мінімальні для цього рівня сегменти (фони, морфи), встановлення їх дистрибуції та підведення на цій основі певних одиниць структури мови (фонем, морфем) як їх варіантів (алофонів, аломорфів);

2-й етап – встановлення дистрибуції самих одиниць структури мови та об'єднання їх у дистрибутивні класи;

3-й етап – побудова деякої моделі мови на даному рівні її структури.

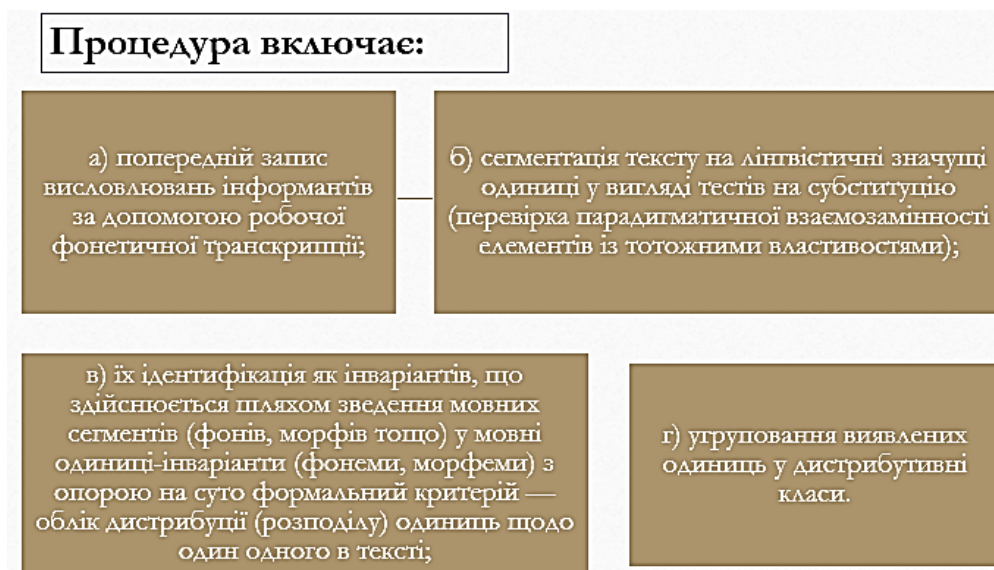


Рис. 2.1. Схема дистрибутивного аналізу.

Опис мови мав завершуватися побудовою загальної моделі структури мови, що відбиває взаємодію різних рівнів. Цим, на думку дескриптивістів, вичерпувалися завдання мікролінгвістики, чи власне лінгвістики, що становить ядро науки про мову (макролінгвістики).

Фонетичні дослідження вони відносили до передлінгвістики, а дослідження значень – до металінгвістики. Дескриптивісти визнавали теоретичну можливість побудувати повний опис мови виключно на основі даних про дистрибуцію її форм. Через це дескриптивну лінгвістику часто називають дистрибутивною лінгвістикою. У дескриптивній лінгвістиці були:

– докладно розроблені методи фонологічного аналізу, що найлегше піддаються формалізації через відсутність у одиниць фонологічного рівня прямого зв'язку зі значенням.

– створено описи фонологічних систем багатьох мов, у тому числі раніше не описаних, досліджувалися супрасегментні (просодичні) явища – тон, наголос, явища стику.

– виділено та описано різні типи морфем.

Багато дескриптивістів розглядали синтаксис як просте продовження морфології. Подібно до того, як все в морфемі зводиться до складових її фонем, слова і конструкції є можливим описувати через їх морфеми і класи морфем. Велику увагу дескриптивна лінгвістика приділяла мові лінгвістичної науки (метамову), проблемам лінгвістичної термінології. Так, були створені тріади термінів для позначення одиниць мови, одиниць структури мови та їх варіантів: фон – фонема – алофон, морф – морфема – аломорф тощо.

Проте спрощене розуміння мови, обмеженість проблематики, абсолютизація дистрибутивного аспекту мови вже в кінці 50-х – початку 60-х рр. привели до кризи дескриптивної лінгвістики, до різкої критики «лінгвістики без сенсу», до появи теорій, що широко звертаються до семантики, в розробці яких взяли участь багато колишніх дескриптивістів, – трансформаційно-породжувальна граматики, теорія компонентного аналізу, різні теорії синтаксичної семантики тощо.

Попередня школа більше приділяє увагу правилам мови, відповідно до яких виводяться залежності. А у випадку дескриптивної лінгвістики, формується поняття корпусу тексту, з якого отримують правила (граматику) і сам тезаурус (лексику). Це поклало початок ідеї корпусної лінгвістики. Крім того, принципи, розроблені американським дескриптивізмом, становлять основу роботи систем дешифрування.

Контрольні запитання:

1. Визначте різницю між європейським структуралізмом та американським дескриптивізмом.

2. Як ви розумієте термін **дистрибуція** у контексті дистрибутивного аналізу?

3. За яким принципом визначається, що перед нами: різні фонemi чи варіант тієї самої фонemi?

4. Як ви розумієте термін **дистрибутивний клас**? За яким принципом такі слова як *bee, girl, war, street* об'єднуються в один дистрибутивний клас?

Практичні завдання

1. Подано український текст такого змісту:

Я глянув і побачив будинок. І сходинок. Я побачив там двері. А біля будинку стоять два хлопчики. І грають на трубі. Тільки вони нічого не грають. Бо вони дуже великі та сірі. Бо вони – статуї. Я стояв на сходах.

Нижче подано варіанти цього самого тексту різними мовами:

Польська:

Spojrzałem i zobaczyłem dom. I kroki. Zobaczyłem drzwi. A przed domem stoją dwaj chłopcy. I grają na trąbce. Tylko, że oni nie grali w nic. Bo to są bardzo duże i szare. Bo to są posągi. Stałem na schodach.

Португальська:

Olhei e vi uma casa. E os passos. Eu vi uma porta. E havia dois rapazes à porta de casa. E eles estão a tocar trompete. Só que eles não tocam nada. Porque são muito grandes e cinzentos. Porque são estátuas. Eu estava de pé nos degraus.

Шведська:

Jag tittade och såg ett hus. Och stegen. Jag såg en dörr. Två pojkar stod utanför huset. Och de spelar trumpet. Men de spelar ingenting. För att de är väldigt stora och grå. För att de är statyer. Jag stod på trappan.

Виокреміть із цих текстів усі можливі частини мови ЧМ (іменники, прикметники, дієслова, частки тощо). Впишіть їх у таблицю:

Мова	ЧМ1	ЧМ2	ЧМ...
Українська			
Польська			
Португальська			
Шведська			

2. Які, на вашу думку, формальні показники допомагають вам виокремити частини мови у наведених вище текстах?

ТЕМА 3 ДОСЛІДЖЕННЯ МОВИ З ПОГЛЯДУ СИСТЕМИ ТА СТРУКТУРИ

План заняття:

1. Мова як системно-структурне утворення
2. Конститутивні та неконститутивні одиниці мови
3. Основні напрямки вивчення системи мови

У цій темі порушено деякі теоретичні питання, що пов'язані з будовою мови як системи. Поняття мови як системно-структурного утворення є ключовим для створення відповідних моделей мови та алгоритмів обробки тексту за цими моделями.

3.1. Мова як системно-структурне утворення

Перш ніж переходити до детального розгляду цього питання, необхідно з'ясувати, що являє собою мова з погляду структурного аспекту її вивчення. З огляду на це варто зазначити, що лінгвістична література пропонує різні визначення. В узагальненому вигляді можна навести таке: **Мова** – це основний засіб, який використовують люди у суспільстві для спілкування, зберігання інформації та оформлення думки. Тут слід зазначити, що мова – це не стихійно утворена сукупність одиниць (знаків), а чітко структурована система, у якій кожен елемент має строго певне місце й установлює зв'язки різного типу з іншими елементами у межах цієї системи. Отже, надалі **мова** буде визначатися як *складно структурована система взаємопов'язаних її елементів, яка служить для комунікації, зберігання й передавання інформації*. Коли говорять про структурованість системи, це означає, що елементи:

- взаємно розрізняються за своїми характеристиками;
- об'єднуються один із одним у групи та класи на основі спільності їх характеристик та протиставляються елементам з інших класів / груп;
- пов'язані один із одним відношеннями різного типу, про що йтиметься нижче.

Розглянемо тепер словосполучення *структурована система*, що використано у поданій дефініції. Це словосполучення передбачає розгляд мови

як *структури* та *системи*. Що є структурою, а що – системою в мові, немає єдиної думки серед учених навіть у сучасній лінгвістиці. Про це викладено у наступній темі. Але поки вважатимемо, що **структура** – *це сукупність елементів і підсистем мови, які організовуються єдине ціле з допомогою різних відношень*. А **система** мови – *це сукупність цих відношень, завдяки яким різноманітні елементи утворюють мову як єдине ціле*. При цьому мова – не просто структурована, а складно структурована система, тобто має багаторівневу організацію, і кожен рівень репрезентований різними одиницями. Традиційно вважається, що структура включає чотири рівні, наведені у таблиці 3.1. На кожному з них елементи мови об'єднуються у класи, категорії, групи та утворюють один із одним зв'язки. Основним поняттям мовної системи є поняття мовних одиниць. **Одиниці мови** – *це її постійні елементи, кожен із яких виконує свою певну функцію і має своє особливе призначення*.

Таблиця 3.1. Мовні рівні, що складають структуру мови.

Рівень	Одиниці	Тип одиниць
Фонетико-фонологічний (ФФН)	фонеми	Стройові
Морфологічний (МРФ)	морфеми	Стройові
Лексико-семантичний (ЛСМ) ²	слова	Номінативна, комунікативна
Синтаксичний (СНТ)	словосполучення	Номінативна, комунікативна
	речення	Комунікативна

Стройові одиниці мови є матеріалом для побудови номінативних та комунікативних одиниць; до стройових належать фонеми й морфеми, а також форми слів та форми словосполучень. Слово є основною **номінативною одиницею** мови, що виконує функцію найменування та повідомлення про предмети, процеси і властивості. Слово як одиниця мови посідає своє певне місце у мовній системі, де всі одиниці взаємопов'язані, однаково важливі для мови. **Комунікативна одиниця** – *відрізок мовлення, що може самотійно*

² На рівні лексики з'являється значення та зміст, тобто семантичний компонент. Тому лексичний, так і синтаксичний рівні мови мають семантичний підрівень: лексико-семантичний, синтактико-семантичний.

передавати повідомлення. До комунікативних одиниць також зараховують лексеми, синтагми, і навіть синтаксеми – висловлювання, абзац, мовний відрізок. У межах мови одиниці утворюють відношення різного типу, як показано на малюнку 3.1. Основними є парадигматичні (червоні стрілки), синтагматичні (символ «+») та ієрархічні (коричнева стрілка) відношення.

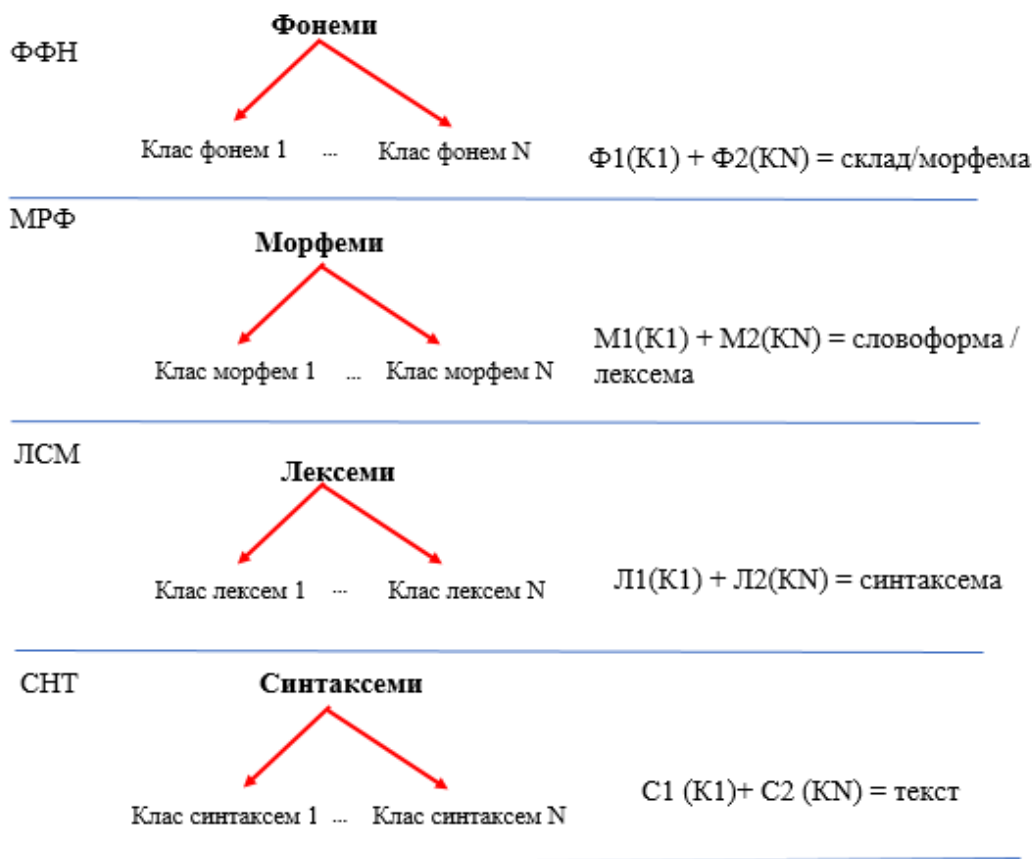


Рис. 3.1. Відношення у системі мови.

За рахунок цих відношень забезпечується об'єднання елементів, а отже, цілісність мови як системи. Однак об'єднання може відбуватися по-різному. Так, **парадигматичні відношення** в межах кожного рівня поєднують за певними параметрами елементи мови у різні класи, групи / категорії. У той самий час за цими параметрами відбуваються протиставлення елементам іншого класу. Так, наприклад, на ФФН за параметром «участь голосу та шуму» приголосні фонери $\langle b \rangle$, $\langle n \rangle$, $\langle ж \rangle$, $\langle ш \rangle$, $\langle з \rangle$, $\langle с \rangle$ об'єднуються в клас дзвінких приголосних та клас глухих приголосних. Своєю чергою, ті самі приголосні протиставляються один одному за дзвінкістю та глухістю. Схематично це можна зобразити так, як показано в таблиці 3.2.

Таблиця 3.2. Клас приголосних за параметрами «Участь голосу і шуму».

Параметр	Значення параметра	Клас	Елементи класу
Участь голосу і шуму	Дзвінкість	Дзвінкі приголосні	б, ж, з
	Глухість	Глухі приголосні	п, ш, с

Кількість класів, як і кількість параметрів, за якими елементи мови розподіляються за класами, залежить від конкретної мови (пор., наприклад, українська, англійська або німецька мови). На малюнку 3.1. сказано подається як: «Клас фонем 1 ... Клас фонем N», де N – кінцева кількість класів. Так само можна сказати і про кількість класів та груп на інших трьох рівнях мови. Як бачимо, у словосполученні «парадигматичні відношення» присутній термін **парадигма**, який ми розглядали у Темі 1. І цей термін означає «клас мовних одиниць, об'єднаний за спільною ознакою». Звідси й походить назва цього типу відношень, мета якого угрупувати елементи мови.

Якщо подивитися на малюнок 3.1., можна також побачити, що елементи одного рівня служать свого роду матеріалом для побудови елементів наступного, вищого рівня. Для цього вони об'єднуються в лінійні ланцюжки. Наприклад, на рівні ФФН фонemi <q>, <u>, <i> утворюють склад *qui*. Об'єднання фонем також дає морфеми, які є одиницями більш високого рівня – морфологічного. З морфем і складів утворюються слова, зі слів – словосполучення та речення. Із речень – текст. Побудова лінійних ланцюжків відбувається завдяки **синтагматичним відношенням**. Слід зазначити, що цей тип відношень заснований на правилах, що визначають, які саме одиниці мають поєднуватися одна з одною. Так, фонема <q> поєднується з <u>. Ще одним прикладом є утворення словосполучень в англійській мові, що складаються з кількох прикметників. За правилом прикметники у цих словосполученнях повинні дотримуватись у такому порядку:

- 1) артикль або інший детермінант (*a, the, his*)
- 2) оцінка, судження (*good, bad, terrible, nice*)
- 3) розмір (*large, little, tiny*)
- 4) вік (*new, young, old*)

- 5) форма (*square, round*)
- 6) колір (*red, yellow, green*)
- 7) походження (*French, lunar, American, eastern, Greek*)
- 8) матеріал (*wooden, metal, glass*)
- 9) призначення (*sleeping, writing, dressing*)

Порівняйте: *a huge round wooden table, mother's nice long green French woolen dress.*

Крім парадигматичних та синтагматичних зв'язків, у системі мови також наявні ієрархічні відношення. **Ієрархічні відношення** – це насправді *відношення входження простішої одиниці у складнішу*. Такі відношення характеризують будову різних мовних одиниць. Перехід від одиниці нижчого рівня до одиниці вищого рівня здійснюється, як правило, в результаті комбінування, тобто реалізації синтагматичних властивостей елементів нижчого рівня. Наведемо деякі приклади у таблиці 3.3.

Таблиця 3.3. Приклади входження простих одиниць до складних.

Тип одиниць нижчого рівня	Одиниці нижчого рівня	Наступний тип одиниці вищого рівня	Наступна одиниця вищого рівня
фонеми	<в>, <о>, <д>, <а>	морфеми	<вод><а>
морфеми	<вод>, <а>	лексема	<вода>
лексеми	<вода>, <тече>	синтаксема	<тече вода>
синтаксеми	<тече вода>...	текст	сам текст

Елементи цієї структури утворюють єдність і цілісність мови, що легко зрозуміти, звернувши увагу на їхній зв'язок. Кожний ступінь може бути потенційно вищим по відношенню до наступного і, у той самий час, нижчим до попереднього. У межах кожного рівня мовної структури (фонетичного, морфемного, лексичного та синтаксичного) наявна власна система, оскільки всі елементи певного рівня виступають як члени системи. Системи окремих ярусів мовної структури, взаємодіючи друг з одним, утворюють загальну систему мови. Для наявності системи обов'язковими є такі умови:

- цілісність;

- упорядкованість;
- дискретність.

Говорячи про цілісність мовної системи, мають на увазі, що у кожному рівні компоненти об'єднані у системи. Так, перший рівень – система фонем, другий – система морфем, третій – система лексем, а далі – система частин мови та система синтаксичних одиниць. Упорядкованість системи характеризує здатність її елементів взаємодіяти та утворювати взаємозв'язки один із одним. Як приклад, можна навести лексичну систему, в якій існують різні впорядковані підсистеми: семантичні поля, лексико-семантичні групи (ЛСГ), синонімічні ряди, антонімічні пари, опозиції на кшталт «питоме – запозичене», «старе – нове», «активне – пасивне». Дискретність означає, що кожна система складається з окремих ділянок (підсистем), які у свою чергу мають досить складну внутрішню організацію (пор. будь-яку підсистему мови).

Контрольні запитання:

1. Як співвідносяться один із одним поняття словоформа, парадигма та лексема? Наведіть приклади
2. Кожний елемент набуває своїх властивостей лише в межах системи. Визначте, які саме властивості набуває мовний елемент «к» у певній підсистемі мови: фонетичний, морфемний та лексичний?
3. Синтагматичні відношення між мовними одиницями також утворюються на лексичному рівні, що зумовлює такі явища як *лексична сполучуваність* та *семантична сполучуваність*. Дайте власне визначення цих понять і наведіть приклади.
4. Які парадигматичні відношення між словами наявні на лексичному рівні? Наведіть власні приклади.
5. Що являють собою парадигматичні відношення на рівні синтаксису? Наведіть власні приклади.

Практичні завдання

1. *Утворіть парадигми різного типу (словотвірну, словозмінну і лексичну) для поданих нижче слів українською мовою:*

блат, болт, гриб, кіт, народ, м'яч, полон, піп, піт, кермо, світло, спирт, **страх**, сором, хміль, джміль, шторм, шум, щит.

Словотвірна: сукупність однокореневих слів.

Однокореневі форми слова *страх*: страхіття, страховисько, страхота, страхітливість; страшний, безстрашний; страшити, устрашати, устрашити; страхітливо;

Словозмінна: парадигма відмінювання / дієвідмінювання, тобто усіх словозмінних форм

Наведемо словозмінну парадигму слова *страх*.

Відмінок	Однина	Відмінок	Множина
Називний	Страх	Називний	Страхи
Родовий	Страху	Родовий	Страхів
Давальний	Страху, страхові	Давальний	Страхам
Знахідний	Страх	Знахідний	Страхи
Орудний	Страхом	Орудний	Страхами
Місцевий	Страху, страхові	Місцевий	Страхах
Кличний	Страху	Кличний	Страхи

Лексична: сукупність слів, об'єднаних за якоюсь семантичною ознакою (синоніми, антоніми, семантичні групи тощо);

Синоніми до слова *страх*: боязкість, боязнь, жах (підсил.), моторошність, настрашка (розм.), острах, обава (діал.), острашка (розм.), пострах (розм.), побоювання.

Антоніми до слова *страх*: відвага, героїзм; геройство, мужність, хоробрість.

Для виконання цього завдання рекомендується використовувати такі ресурси:

- «Словники України онлайн»: <https://lcorp.ulif.org.ua/dictua/>
- Електронний онлайн словник «Горох»: <https://goroh.pp.ua/>
- Словник антонімів онлайн: <https://slovnyk-antonimiv.com/>

2. Розподіліть слова у завданні 1 по відповідних семантичних групах. Наприклад, слова *спирт* і *нім* входять до групи «Назви речовин, субстанцій».

Семантична група	Слова
Назви речовин, субстанцій	спирт, піт
...	

Щоб виокремити групи, необхідно попередньо визначити лексичне значення всіх слів, для чого можна використати такі ресурси:

- Тлумачний словник СУМ 11 онлайн: <http://sum.in.ua/>
- Електронний онлайн словник «Горох»: <https://goroh.pp.ua/>

3. Для утворення прикметників чоловічого роду використовуються суфікси -ов-, -ев-, -ан-, -н-, -ін-, -інн-, -ин- і, відповідно, закінчення -ий та -ій, як це показано нижче:

кавуновий, барабанний, ватний, верхній, вечірній, винний, зовнішній, вчорашній, задній, марний, олівцевий, кіношний, крайній, літній, мідний, новорічний, нинішній, осінній, передній, післязавтрашній, тодішній, хвойний, шоколадний.

3а. Користуючись правописом, визначте правила (алгоритм) утворення прикметника, які зумовлюють вибір того чи іншого суфікса і закінчення.

3б. За якими параметрами (напр.: основа слова, семантика, кількість складів, частина мови тощо) здійснюється вибір суфікса і закінчення? Обґрунтуйте свій вибір і надайте у таблиці параметри та їхні можливі значення:

Параметр	Значення параметра	Суфікс	Закінчення

3.2. Конститутивні та неконститутивні одиниці мови

Як було сказано вище, мова є комплексом систем, що пояснює її складність у плані організації, існування в ньому різних її підсистем, елементів, за рахунок яких мова формується і функціонує як єдине ціле. Вище ми

розглянули одиниці мови за своїм призначенням. Однак їх можна ще класифікувати за іншим критерієм – за їхнім місцем та роллю в організації системи та структури мови. Виходячи з цього, одиниці мови поділяються на **конститутивні** та **неконститутивні**.

Конститутивні одиниці (КО) – *це базові одиниці у системі мови*. Вони відтворювані та обчислювані, мають властиві їм формальні показники, а також відносно постійний статус у системі мови. До них належать фонemi, морфеми, словоформи, слова, словосполучення, речення. Будучи відтворюваними, класи таких одиниць суттєво різняться і між собою у системній ієрархії мови. Сукупність КО одного порядку, їх матеріальна та ідеальна (для двосторонніх одиниць) природа, виконувані ними функції, закономірні парадигматичні та синтагматичні відношення між ними, значимість цих одиниць, і формують структурний рівень мови, що входить в ієрархію інших мовних рівнів. Рівні мови утворюють у сукупності її **стратифікацію** (поділ, розчленування чогось на шари, рівні). Рівень мови є одним із головних компонентів структури мови. Своєрідність КО виявляється не лише у відношеннях між собою, а й у зв'язках з конститутивними і неконститутивними одиницями інших рівнів. Для КО справедливими є такі твердження:

КО 1. Їм притаманна ієрархічність відносин, яка є однією з головних умов утворення структури мови. КО вищого рівня утворюється у вигляді участі одиниць нижчого рівня; нижча КО, таким чином, входить до складу вищої. Наприклад, фонема входить до складу морфеми, морфеми утворюють слова, слова – словосполучення, словосполучення – речення), як це було показано у таблиці 3.3.

КО 2. Утворена за допомогою нижчих одиниць КО вищого порядку є новотвором, що наділений властивими тільки йому значенням і функціями. Так, наприклад, функцією фонем є розрізнення вищих елементів (морфеми і слова), а також засіб для побудови цих самих елементів. А утворені від фонем морфеми виконують вже інші функції: словотвірну, формоутворюючу

і смислові. Далі – слова, що утворюються від морфем, виконують номінативну та комунікативну функції.

КО 3. Характерними для КО є інваріантні та варіантні відношення. Варіантні спостерігаються на всіх рівнях структури мови: варіюються її основні одиниці та їх об'єднання. Спільним для всіх варіантів є наявність інваріанту. Наприклад, на фонетичному рівні варіантом фонemi є її реалізація у тому чи іншому фонетичному оточенні (алофони).

На словотвірному рівні варіантність виявляється у вживанні тієї самої морфеми у кількох матеріальних різновидах. При цьому варіації можуть зазнавати морфеми: 1) кореневі (*ніс* – *носовий*, *рух* – *русі* – *рушати*); 2) префіксальні (*збирати* – *зібрати*, *відмикати* – *відімкнути*, *підбирати* – *підібрати*), 3) суфіксальні (*дубок*, *нігтик*, *слинявчик*); 4) закінчення (*співа* – *співає*).

На лексичному рівні варіантність полягає: 1) вираження у слові різних значень залежно від контекстного оточення; 2) вираження того самого значення різними лексичними засобами (синонімія: *ходити* – *вештатися*, *шкитильгати*; територіальні діалектизми: *картопля* – *картона*, *бульба*; соціальні або професіональні діалектизми: *човен* – *дуб* «човен, видовбаний із суцільного дерева», *байдак* «великий човен для плавання по морю», *калабуха* «невеликий човен»).

Варіація на синтаксичному рівні зумовлена: 1) взаємодією лексичної семантики слова або його граматичного значення з морфемним складом (*натиснути клавішу* / *натиснути на клавішу*); 2) дією закону смислової або структурної аналогії (*смикати мотузку* / *смикати за мотузку* – за аналогією до *брати щось і брати за щось*); 3) вплив процесу синтаксичного перерозкладання, тобто зміни зв'язків та відносин між частинами складеного словосполучення (*подивитись у очі сина* / *подивитись у очі синові*).

КО 4. КО властиві парадигматичні та синтагматичні відношення, про які детально висвітлено в першому підрозділі цієї теми (див. рис. 3.1).

Проте мовні рівні не складаються лише зі стройових, фундаментальних одиниць. Існують також **неконститутивні одиниці (НКО)**, що не мають зазначених вище функцій, але виконують внутрішньо- та міжрівневі функції. Класифікація та характеристика можливих НКО подані у таблиці 3.4.

Таблиця 3.4. Характеристика і класифікація НКО.

Характеристика	Види НКО
Утворюються безпосередньо в мовленні за прийнятими у мові правилами та моделями	фонетичне слово, синтагма, члени речення, елементи актуального членування речення
Виступають елементом тієї чи іншої КО	склад, інтонація, різні види наголосу
Утворюються на базі КО та виконують властиві КО окремі функції	фразеологізми, різноманітні лексикалізовані сполучення слів ³
Утворюють нові одиниці з новими функціями	моделі словотвору, різного роду основи слова

Є підстави вважати, такі одиниці можуть утворювати підрівні мови. Розглянемо деякі НКЄ. **Синтагма** – частина фрази, що виділяється ритміко-мелодійними засобами та вимовляється на одному диханні. Відокремлюється від подібних одиниць короткими паузами (/). Виділення мовних тактів обумовлено змістом, значенням, яке вкладає мовець у своє висловлювання. Наприклад: *З дому навпроти / лунав сміх; З дому / навпроти / лунав сміх; Вона розважала його / малюнками дітей, Вона розважала його малюнками / дітей.* У мовному такті можна назвати більш дробові одиниці – фонетичні слова.

Фонетичне слово – фрагмент мовленнєвого потоку, об'єднаний одним словесним наголосом. Пауз між фонетичними словами не може бути. Розглянемо фразу: *Садок вишневий / коло хати.* У цій фразі два мовленнєвих акти, кожен із яких складається з двох фонетичних слів. До складу фонетичного слова (*коло хати*) у другому такті входить ненаголошений прийменник. Фонетичне слово – це значеннєве слово зі службовими словами (сполучниками, прийменниками та частками), наприклад: *а ти, на нозі, йшов би.*

³ Лексикалізовані сполучення слів – це поєднання слів, які функціонально прагнуть стати одним словом. Наприклад: *Новий Рік, білий квиток.*

Своєю чергою, фонетичне слово поділяється на **склади**, а склади – на звуки. Отже, звук, склад, фонетичне слово, такт і фраза – це лінійні відрізки звукового потоку. Їх називають сегментними одиницями. Більші одиниці поділяються на дрібніші. Об'єднання дрібних одиниць у більші здійснюється завдяки інтонації та наголосу. Інтонація і наголос – це суперсегментні одиниці, оскільки вони накладаються на сегментні.

Контрольні запитання:

1. Визначте спільні та відмінні риси між поняттями *іменника* та *підмета*, *дієслова* та *присудка*, *прикметника* і *означення*?
2. Перелічіть основні типи морфем. Які значення вони можуть передавати? Що таке внутрішня форма слова?
3. Конститутивні одиниці вищого рівня не являють собою суму значень одиниць нижчого рівня. Подайте приклади, що ілюструють таке твердження.
4. Дайте визначення поняття *інтонації* та надайте приклади, як вона може міняти зміст речення з тією самою синтаксичною конструкцією.

Практичні завдання

1. У наведених нижче словах виокреміть суфікси та визначте, які значення реалізує суфікс у складі того чи іншого слова. Для зручності виконання цього завдання можна використовувати онлайн словник за адресою: <https://www.merriam-webster.com/>

children, bookish, worked, Chinese, reading, usefully, worker, easiness, childhood, better, impossible, chosen; unspoken, government, closest, wisdom, writes, upbringing; available, hairy, rooms, speechless, recommendation, fashionable, discover, pleasantly.

Слово	Суфікс	Значення, що реалізує суфікс у складі слова

2. Наведіть характеристику суфіксів із завдання 1. Для зручності можна використовувати онлайн словник за адресою: <https://www.merriam-webster.com/>

Суфікс	Тип суфікса	Що позначає слово, з яким сполучається суфікс	Частина мови, до якої належить слово	Яке значення набуває слово після додання суфікса
-ish (2) *	СЛТ**	істота	іменник	щось характерне або притаманне істоті
-es (1)	СЛЗ***	істота/неістота	іменник	множина предметів
-es (2)	СЛЗ***	дія	дієслово	граматичне значення третьої особи однини Present Simple

* номер значення, що може реалізовувати суфікс;

**СЛТ – словотвірний суфікс;

***СЛЗ – словозмінний суфікс;

3. Розбийте поданий нижче текст на речення, а самі речення – на лексичні синтагми. Вкажіть, за якими ознаками може відбуватися поділ на синтагми. Порахуйте кількість фонетичних слів у кожній синтагмі.

Я серед нив

Мої дні течуть тепер серед степу, серед долини, наливої зеленим хлібом. Безконечні стежки водять мене. по нивах, а ниви котять та й котять зелені хвилі та хлюпають ними аж у краї неба.

На небі – сонце, серед нив я. Більше нікого. Іду. Гладжу рукою соболину шерсть ячменів, шовк колосистої хвилі. Тихо пливе блакитними річками льон. Так тихо й спокійно в зелених берегах, що хочеться сісти на човен і попливти. А там ячмінь хилиться та тче з тонких вусів зелений серпанок. Волошки дивляться в небо. Тепер пішла пшениця. Біжить за вітром, немов табун лисиць, і блищить на сонці хвилясті хребти. А я все йду.

Урешті зупиняюся. Мене спиняє біла піна гречок, запашна її легка, наче збита крилами бджіл. Просто під ноги лягла співуча арфа й гуде на всі струни. Стою й слухаю.

Повні вуха маю того дивного гомону поля, того шелесту шовку, того безупинного, як текуча вода, пересипання зерна.

3.3. Основні напрямки вивчення системи мови

У середині ХХ століття відбувається перехід атомарного підходу до системного, а точніше – системно-структурного підходу до вивчення мови (див. підрозділ 1.3). Новий підхід означає перехід від фрагментарного уявлення про мову до розуміння її єдності та цілісності, принципів організації, устрою мови, що вивчається. Також спостерігалися спроби створити загальну теорію системи мови, розроблялися нові методи та системно-структурні дослідження. У сучасній лінгвістиці можна виділити такі напрямки у вивченні системно-структурної організації мови: інвентарно-таксономічний, функціональний і власне структурний. Крім них існують також стратифікаційний і типологічний напрямки, проте ми не будемо на них зупинятися.

У **парадигмі інвентарно-таксономічного підходу** мова як система являє собою набір, або інвентар, що використовується для побудови складніших лінгвістичних об'єктів (фонема -> морфема -> лексема -> синтаксема). Цей напрямок дав науковцям значні результати, зокрема:

- розроблено теорії, методики опису мовних елементів та диференційних ознак цих елементів;
- виявлено принципи поєднання елементів мови у відповідні групи, наприклад, граматичні класи, частини мови, парадигматичні ряди тощо з урахуванням їх диференціальних ознак;
- описані відношення, за допомогою яких елементи мови утворюють систему (за рахунок таких відношень відбувається не тільки угруповання одиниць у класи, але також їхнє поєднання один з одним);
- визначено мінімальні елементи мови, а також підсистеми /рівні мови, на яких знаходяться ці елементи (див. рис. 3.1).

Таким чином, маючи вихідну базу мінімальних одиниць, можна побудувати необмежену кількість лінгвістичних об'єктів (з фонем – морфемами, з морфем – словами, зі слів – синтаксеми: словосполучення та речення). При цьому побудова таких об'єктів ґрунтується на певних правилах поєднання вихідних одиниць, як про це йшлося у підрозділі 3.1. У світлі даного підходу також слід

зазначити, що система мови зображується у вигляді класифікаційних схем, у яких всі елементи мови перебувають у складних ієрархічних відносинах один із одним.

Відповідно до **функціонального підходу**, мова як система розглядається з погляду виконуваних нею функцій, головним чином комунікативної. Таке бачення системи мови зустрічається у працях Празького лінгвістичного гуртка (див. розділ 2.1). Основним критерієм виокремлення системних елементів при такому підході є функціональний, основними принципами якого є:

- мова є системою засобів вираження, спрямованою на досягнення комунікативної мети;
- система мови існує як взаємодія та взаємозалежність лексичного, граматичного та фонетичного аспектів мови;
- зміст висловлення визначає його синтаксичну структуру з використанням лексичних та граматичних засобів та відповідного фонетичного оформлення.

Описи конкретних підсистем мови (граматичних або фонологічних) базуються на відносинах типу «засіб – мета». Так, наприклад, функція одиниць нижчого рівня полягає у тому, щоб бути засобом побудови одиниць вищого рівня. Наприклад, рівень морфем містить, з одного боку, результати поєднання фонем, а з іншого, – засоби для побудови слів.

Одним із головних досягнень функціонального структуралізму є теорія актуального членування речення, а також система фонологічних та граматичних опозицій, які використовувалися лінгвістами у межах інвентарно-таксономічного підходу. У сфері функціонального підходу було також розроблено принципи семантичного моделювання структури речення, що привело до виникнення окремого напрямку – функціонального синтаксису. Результатом застосування функціонального підходу в лексикології є розробка лексичних мінімумів, а також укладання різноманітних словників: частотних, граматичних, асоціативних та інтегральних. І, нарешті, сам підхід перетворився

на окремий напрямок – функціональна лінгвістика, яка вивчає мовленнєві акти та відношення, які безпосередньо пов'язані з ними.

Власне структурний підхід передбачає, що мовна система базується на структурі, тобто на мережі відношень: а) між різними рівнями системи мови; б) у межах певного рівня мови. Ці відношення не залежать від природи елементів, між якими вони встановлюються. Наслідуючи цю ідею, копенгагенські структуралісти (див. підрозділ 2.2) розробили універсальну лінгвістичну теорію (глосематику), яка потім удосконалювалася в процесі подальшого розвитку лінгвістики. Саме завдяки цьому аспекту вдалося не тільки виділити структуру мови як самостійний об'єкт дослідження, але й розробити математичні моделі, здатні описувати цю структуру. Пізніше, у другій половині 50-х рр. XX століття поняття структури широко використовувалося в теорії породжувальних граматик Ноама Хомського, про які мова піде в наступних темах. Під породжувальною граматикою слід розуміти певну автоматичну систему, яка генерує безліч граматичних речень будь-якою мовою. У складі цієї системи структура є невід'ємним елементом граматики, а також певним шаблоном (тобто синтаксичною конструкцією), відповідно до якого обираються елементи та будується речення.

Контрольні запитання:

1. Які мінімальні та максимальні одиниці, а також типи зв'язків між ними виокремлюють в інвентарно-таксономічному напрямку?
2. Які існують основні принципи виокремлення системних одиниць у функціональному напрямі?
3. У чому полягає перевага структурного підходу до опису мови?
4. У чому полягає різниця в інтерпретації поняття системи у таксономічному, функціональному та структурному підходах?

Практичні завдання

1. *Визначте, які із запропонованих ознак є означальними для угруповання слів за певними частинами мови:*

Ознака	Відповідь: так / ні	Обґрунтуйте відповідь та надайте приклади
Семантичні класи слів		
Формальні класи слів		
Мають єдність лексичного значення		
Мають спільне категоріальне значення		
Мають спільні морфологічні категорії		
Граматичні класи слів		
Мають спільну систему парадигм словозміни		
Мають спільні словотвірні форманти		
Мають спільні синтаксичні функції		

2. Визначте, за якими семантичними ознаками можна об'єднати подані нижче дієслова стану у відповідні групи:

admire, adore, appear, believe, belong, care for, consist, contain, depend, deserve, desire, detest, dislike, enjoy, expect, feel, forget, hate, hear, include, involve, know, lack, like, look, love, mean, notice, possess, prefer, realize, recognize, remember, resemble, respect, see, smell, sound, taste, think, understand, want, wish.

Сформулюйте назви ознак та розподіліть наведені дієслова у таблицю:

Ознака 1:	Ознака 2:	Ознака 3:	Ознака 4:	Ознака 5:
_____	_____	_____	_____	_____

3. Морфемі в словах виконують різну функцію: утворюють нові слова (словотворча функція) й змінюють форму того самого слова (формотворча функція). Проаналізуйте морфемний склад слів та розподіліть морфемі по відповідних стовпчиках для кожного слова.

Приклад: знесилений

Словотворчі морфемі			Формотворчі морфемі		
префікс	суфікс	закінчення	префікс	суфікс	закінчення
з-	Ø	-ий	не-	-ен	-ий

молочниця, дочитати, золотистий, позолотити, перемалював, синюватий, золотитися, театральний, квіткою, зробити, зеленесенький, обійти, безхмарний, надламати, найкращий, собачий, зареєструвати, куряча, дослухати, виспівати, тепліший, зроблений, тесак, викрадач.

МОДУЛЬ 2 СТРУКТУРНА ЛІНГВІСТИКА ЯК ДИСЦИПЛІНА

ТЕМА 4 СТРУКТУРНА ЛІНГВІСТИКА ТА ПРЕДМЕТ ЇЇ ВИВЧЕННЯ

План заняття:

1. Структурна лінгвістика: її предмет, завдання і методи
2. Структура і система: співвідношення понять
3. Граматика і словник як дві підсистеми мови

Сучасна структурна лінгвістика як наукова дисципліна надає підґрунтя для репрезентації природної мови у комп'ютерних системах, а також розробку алгоритмів аналізу та систематизації мовного матеріалу. Цей розділ знайомить зі структурною лінгвістикою, її предметом, проблемами та методами.

4.1. Структурна лінгвістика: її предмет, завдання і методи

Як було зазначено вище (див. підрозділи 1.3 та 3.3), структуралізм як нова наукова течія у мовознавстві виник унаслідок кризи молодогограматизму, що була зумовлена:

1) фрагментарним підходом до вивчення мовного матеріалу, тобто вивчали, досліджували окремі мовні явища без спроби встановити взаємозв'язки між ними та звести їх у єдину систему;

2) розглядом індивідуально-психологічних особливостей мовців, які відображалися у їхньому мовленні, і це часто вело до ототожнення мови з мовленням окремих носіїв;

3) зосередженням лише на історичному аспекті дослідження мовних фактів (діахронія) і в той самий час нехтування сучасного стану мови (синхронія).

Крім цього, велику роль у становленні структуралізму в мовознавстві зіграли досягнення у галузі природничих наук, зокрема у хімії. Так, наприклад, учені зацікавились питанням, з яких елементів (молекул, атомів) складається та чи інша речовина, і як вони утворюють між собою структуру цієї речовини. До того ж, методи та підходи, що використовувалися в цих науках, довели свою ефективність. Застосування методів математики та законів логіки (типу «А передбачає В», «Якщо А еквівалентне В, і В еквівалентне С, то А еквівалентне С» тощо) дало можливість візуалізувати зв'язки між елементами мови та

показати, що сама мова являє не стихійно сформовану сукупність одиниць, а певний об'єкт, у якому чітко простежується його системно-структурна організація: його складники та їхні взаємозв'язки, а також утворювані в межах об'єкта класи, групи і категорії. Історію розвитку структурної лінгвістики можна поділити умовно на два етапи: (1) класична структурна лінгвістика, що охоплює період 20-40-х рр. XX ст., та (2) сучасна структурна лінгвістика. Такий поділ ґрунтується на таких моментах:

- які існували теоретичні погляди на структурні особливості системи мови;
- які використовуються методи дослідження протягом того чи іншого етапу;
- на які задачі були орієнтовані теоретичні дослідження у галузі структурної лінгвістики.



Рис. 4.1. Етапи розвитку структурної лінгвістики.

Класична структурна лінгвістика бере свій початок у 1920-ті рр., а її основи закладено Фердинандом де Сосюром. На той момент вона являла собою нову течію у мовознавстві під назвою «структуралізм», точніше «лінгвістичний структуралізм». Однак, можна стверджувати, що саме тоді вона почала формуватися як окремий розділ лінгвістичної науки зі своєю теоретичною базою та методами, які досі використовуються сучасними лінгвістами (див.

модуль 5). Класичну структурну лінгвістику репрезентують європейські школи структуралізму (підрозділи 2.1 і 2.2): Празький лінгвістичний гурток і Копенгагенський структуралізм, а також американський дескриптивізм (підрозділ 2.3). У межах нового напрямку (або парадигми) учені вважали завданням лінгвістики *розробку теоретико-методологічної бази для об'єктивного опису мови*, наближуючи таким чином лінгвістику до числа точних наук, що оперують математичними, логічними і статистичними методами. Основними рисами такого опису є:

- 1) репрезентація мови як системної структури, у якій все взаємопов'язано та взаємозумовлено;
- 2) віддання переваги синхронії, оскільки мова на цьому етапі усталена, і, відповідно в ній можна чітко простежити системні та структурні риси;
- 3) прагнення отримати об'єктивний опис, виключаючи історичні, соціальні та психологічні чинники;
- 4) уявлення про структуру мови як певного утворення, що складається з кількох пов'язаних між собою рівнів.

Виходячи з вищесказаного, провідні структуралісти прагнули виробити процедури лінгвістичного аналізу, що дають у результаті об'єктивне уявлення про мову, її організацію та функціонування. Структурна лінгвістика переважно зосереджувалася на тих аспектах вивчення мовної структури, які можна підраховувати, класифікувати: фонетичний, морфемний та лексичний склад, синтаксичний устрій тощо. При цьому мовознавці намагалися виробити такий науковий опис мовного матеріалу, щоб він був непротирічним, вичерпним та максимально простим, як у глосематиці, предметом якої є опис структури мови через мережу різних типів відношень (див. підрозділ 2.2). Розглянемо методи класичної структурної лінгвістики, які досі використовуються дослідниками та становлять методологічну базу сучасного мовознавства.

Метод опозицій, що полягає у протиставлянні одних мовних одиниць іншим за їхніми властивостями або виконуваними функціями. Такий метод вперше застосував провідний представник Празької лінгвістичної школи

Микола Трубецькой у науковій праці «Основи фонології», що вийшла у світ 1939 р. У ній учений розробив систему фонологічних опозицій, за якою усі фонemi певної мови протиставляються одна одній на основі їхньої смислорозрізнявальної функції. За рахунок таких опозицій отримуємо об'єднання фонем у різні класи, що загалом формують фонологічний рівень мови. Пізніше такий метод почали застосовувати для усіх інших рівнів мови. Наприклад, на морфемному рівні, точніше на підрівні словотвору, відбувається об'єднання словоформ у різні граматичні класи і категорії за рахунок протиставлення їхніх граматичних значень (пор. наприклад, розподілення іменників за категорією числа, роду; дієслів за категорією доконаності / недоконаності тощо).

Метод дистрибутивного аналізу, алгоритм реалізації якого описано в підрозділі 2.3, розроблено у 1930-40-і рр. представниками американського дескриптивізму. Його головною особливістю є виявлення порядку розташування (аранжування) і розподілу (дистрибуції) окремих частин висловлювання одна відносно одної. Основними прийомами є сегментація (членування) тексту та дистрибутивний аналіз виявлених одиниць. При цьому застосовується техніка субституції (заміщення), під час якої дослідник підставляє сегменти одного висловлення в інші та перевіряє їх прийнятність для інформанта. Після здійснення процедури сегментування тексту на елементарні одиниці (на фонологічному рівні є фони, на морфологічному – морфи) проводиться процедура ідентифікації, тобто встановлюється, які є варіантами однієї й тієї ж одиниці (алофонами чи аломорфами), а які – різних. Фінальний етап процедури – розподіл виявлених елементів по дистрибутивних класах.

Метод компонентного аналізу. Значні успіхи при аналізі фонологічних та граматичних проблем стимулювали окремих структуралістів адаптувати й застосувати випробувані методи до формального опису семантичної сторони мови. Зазначена обставина привела до введення поняття **семи** представником Празького лінгвістичного гуртка Володимиром Скалічкою для позначення

мінімального компонента значення. Метод дослідження семантичної структури отримав назву **метод компонентного аналізу**, і його метою є розкладання значення на мінімальні семантичні складники. Основною гіпотезою стало припущення, що значення будь-якої одиниці мови складається зі сем і його можна описати за допомогою обмеженої і відносно невеликої кількості семантичних ознак. Сукупність самих мовних одиниць, об'єднаних спільністю лексичного значення, одержало назву **семантичного поля**. Загальна семантична ознака, за якою слова об'єднуються в поля, виражається **інтегральною семою**. Проте крім інтегральної слова мають також **диференційні семи**, якими ці слова відрізняються одне від одного. Наприклад, для дієслів *йти*, *бігти*, *їхати*, *плисти*, *летіти* тощо інтегральною є сема «переміщення у просторі», а диференційними – «швидкість», «спосіб», «середовище» пересування.

Трансформаційний метод, який детально описаний у темі 11, полягає в тому, що основу синтаксичної системи будь-якої мови утворює набір елементарних типів речень, які називаються **ядерними**. Зазвичай як такі розглядалися прості нерозповсюджені речення активного стану, з яких можливо побудувати різні речення (**трансформи**). Перетворення ядерних речень на складніші відбувається за трансформаційними правилами, які передбачають, з одного боку, перебудову структури речення, а з другого – їх з'єднання між собою. Розповсюджені правила трансформації включають:

- перетворення стверджувального речення у питальне і навпаки (Петро прийшов → Прийшов Петро?);
- перетворення стверджувального речення у заперечне і навпаки (Петро прийшов → Петро не прийшов).
- перетворення активної конструкції у пасивну і навпаки (Микитюк буде організовувати концерт → Концерт буде організовуватися Микитюком);
- перетворення речення у словосполучення і навпаки (Микола виконує завдання → Виконання завдання Миколою).

За допомогою трансформаційного методу ілюструвалася відмінність елементів, репрезентованих у однакових зовнішньо синтаксичних конструкціях. Так, зазначалося, що при зіставленні речень «Він пише другові листа» і «Він пише другові годину» – перше допускає трансформу «Лист другові пишеться ним», тоді як друге не може бути перетвореним на «Година другові пишеться ним».

У другій половині ХХ ст. відбувається поява нових напрямів. Це було пов'язано з появою після другої світової війни нових галузей: машинний переклад, штучний інтелект, інформаційний пошук, питально-відповідальні системи тощо. Внаслідок чого структурна лінгвістика набуває «технічного» статусу. Не зважаючи на те, що класична структурна лінгвістика на той період часу вже вичерпала себе, наведені вище методи досі використовуються у сучасних мовознавчих студіях. Перейдемо до розгляду другого етапу – сучасної структурної лінгвістики.

Сучасна структурна лінгвістика характеризується зміною класичної структуралістської парадигми на нову, а саме на **генеративізм** або генеративну лінгвістику, в основі якої лежить уявлення про **породжувальну модель мови**. Така модель являє собою скінчений набір правил, що дають змогу задати, або «породити», усі можливі правильні речення мови. Іншими словами, генеративна лінгвістика не описує мову, як це відбувалося у класичній структурній лінгвістиці, а являє собою процес моделювання мови. Відповідно, метод моделювання є характерною рисою сучасної структурної лінгвістики. Зародженню генеративізму в сучасній структурній лінгвістиці сприяло виникнення нових на той момент комп'ютерних галузей, пов'язаних з автоматизованим опрацюванням мови, а також математичної лінгвістики.

Породжувальна модель мови репрезентує структуру речення (синтаксичну структуру), яка розкладається на **безпосередні складники**. Варто зазначити, що метод розкладання на безпосередні складники (метод аналізу БС) запропонував основоположник американської дескриптивної лінгвістики Леонард Блумфілд, який потім активно використовували генеративісти,

зокрема Ноам Хомський у теорії породжувальних граматик (див. тему 12). Алгоритм розкладання речення на БС включає такі кроки:

- 1) вихідним для аналізу визнається ціле речення, тобто певне висловлювання, яке репрезентує ту чи іншу синтаксичну конструкцію;
- 2) сама конструкція розкладається на дві частини (бінарний принцип), які називаються група підмета і група присудка;
- 3) потім послідовно розкладаються група підмета і група присудка на два елементи;
- 4) якщо у той чи іншій групі залишається більше двох елементів, то повторюється крок (3).

Наведемо приклад побудови породжувальної моделі, яка може генерувати речення (S) типу «Little boy plays football». За наведеним алгоритмом при першому поділі ми отримуємо як безпосередні складники групу підмета (NP) = «Little boy» та групу присудка (VP) = «plays football», а при другому поділі кожної з цих груп виокремлюються безпосередні складники: прикметник (A) = «Little», іменник (N) = «boy» та дієслово (V) = «plays», додаток (O) = «football». Отже, модель безпосередніх складників для такого типу речень можна подати як на малюнку 4.2.

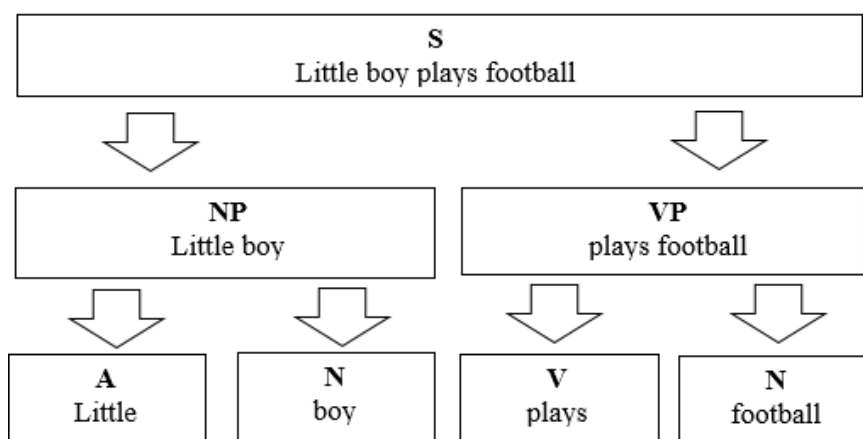


Рис. 4.2. Модель безпосередніх складників.

Такий формалізм, заснований на моделюванні синтаксичних структур отримав назву **генеративний синтаксис** і став основою теорії породжувальних граматик Ноама Хомського. Якщо порівнювати попередні напрями, що існували в структурній лінгвістиці, породжувальна (генеративна) граматика

розв'язує зовсім інші завдання. Вона не призначена для збирання, опису та узагальнення мовних даних. Кінцевим результатом досліджень, що виконуються в межах породжувальної граматики має стати відповідь як людина засвоює мову, як людина породжує висловлення, тобто використовує мову. Іншими словами, генеративна граMATика моделює мовно-розумовий апарат людини в аспекті засвоєння мови, тобто здатності людини будувати граматично правильні речення.

Лінгвістична теорія Хомського стала основою для створення різних комп'ютерних систем, що працюють із природною мовою. Але спочатку, коли мова програмування була недосконалою, комп'ютер не міг розрізняти семантичні тонкощі природної мови (наприклад: синонімія, антонімія), тому він породжував речення, які були правильні з погляду синтаксису, але змістовно аномальними з погляду семантики. Так, Юрій Апресян наводить один із прикладів такого речення: «Коли інженер Смол змашений, вода у дзвониках підігрівається». Отже, послідовники Хомського почали розуміти, що головним має бути не синтаксис, а зміст, тобто семантика речення. Тому новим напрямком у сучасній структурній лінгвістиці наприкінці 1960-х та початку 1970-х рр. виник новий напрям – генеративна семантика.

Генеративна семантика є подальшим розвитком теорії породжувальних граматик, що характеризується побудовою моделі мови за схемою «від значення до тексту» і «від тексту до значення». Підхід до моделювання мови полягає у тому, що граMATика породжує основний зміст речень і потім перетворює цей зміст у реальні речення. До представників генеративної семантики належать Джон Росс, Пол Постал, Джеймс МакКоулі, Джордж Лакофф, Юрій Апресян, Ігор Мельчук, Олександр Жолковський та інші. На відміну від генеративного синтаксису, синтаксичні структури визначаються на основі змісту (семантики). За таким підходом змісти породжуються граMATикою як глибинні структури, а потім такі структури набувають вигляду звичайних речень (поверхневі структури).

Розглянемо поняття глибинної та поверхневої структур. **Глибинна структура** – це теоретична конструкція, що прагне пов'язати низку речень, спільних за своїм змістом, але при цьому з різними синтаксичними конструкціями. Глибинну структуру можна подати у вигляді дерева безпосередніх складників або у вигляді формального запису у квадратних дужках (див. рис. 4.3). Структура дерева формується у такий спосіб: кореню дерева відповідає початковий символ S ; а вузлам дерева – кожний символ послідовності, на який замінюється попередній символ (початковий символ S на першому кроці), і до якого проводиться дуга з кореня.

Поверхневою структурою називають *реалізацію глибинної структури у вигляді речення з певною синтаксичною будовою*. Наведемо такий приклад поверхневих структур:

- (1) Бобри будують греблю. (3) Будівництво греблі бобрами.
 (2) Гребля будується бобрами. (4) Будівництво греблі здійснюється бобрами.

Як бачимо, наведені вище речення мають різні синтаксичні конструкції, але однаковий зміст, який задається глибинною структурою, поданою на рисунку нижче. На основі концепції глибинних та поверхневих структур Юрій Апресян, Ігор Мельчук, Олександр Жолковський побудували теорію «Смисл $\leftrightarrow$ Текст», про яку детально висвітлено в темі 12.

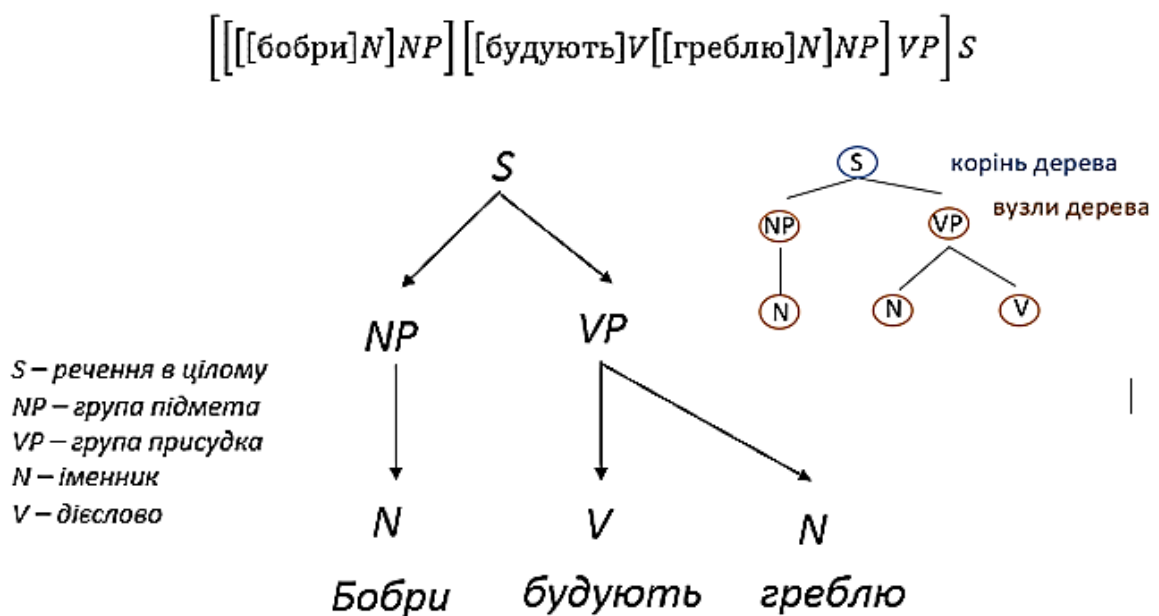


Рис. 4.3. Глибинна синтаксична структура.

Виходячи з вищесказаного, можна подати таке визначення структурної лінгвістики: **структурна лінгвістика** – важливий розділ мовознавчої науки, предметом якої є репрезентація системи мови та її елементів у вигляді певних структур. Отже, завданнями структурної лінгвістики є: (1) опис мовного матеріалу у вигляді структур; (2) моделювання мовних структур (формальних та змістових аспектів мови).

Контрольні запитання:

1. У чому полягають особливості вивчення і репрезентації мовної структури в класичній та сучасній структурній лінгвістиці?
2. Які можна виявити обмеження у використанні методів класичної та структурної лінгвістики?
3. Які переваги дає метод моделювання у традиційній та комп'ютерній лінгвістиці?

Практичні завдання

1. Виконайте компонентний аналіз слів, інтегральною семою яких є «предмет меблі для сидіння», а саме визначте диференційні семи, за якими ці слова відрізняються одне від одного:

stool, chair, armchair, sofa, rocking chair, high chair, love seat, lawn chair.

Для пошуку визначень можна використовувати будь-який тлумачний словник, наприклад: <https://www.oxfordlearnersdictionaries.com/>

Приклад розкладу значення слова *stool* на окремі семи на основі визначення:

Stool: “a seat with legs but with nothing to support your back or arms”.

Фрагмент визначення	Сема №	Сема
with legs	C1	Наявність ніжок
with nothing to support your back	C2	Відсутність спинки
but with nothing to support your arms	C3	Відсутність підлокітників

Після цього заповніть таблицю та впишіть до неї усі семи, які були виокремлені для всіх слів. Якщо те чи інше слово має відповідну сему, позначте відповідну клітинку знаком «+», якщо ні – знаком «–»:

Слово	C1	C2	C3	C4	...	...	...	...	...	...

Поясніть, чому деякі слова можуть мати спільні диференційні семи? Надайте стислий аналіз отриманої таблиці.

2. Подані речення, які начебто мають однакову синтаксичну будову:

- (1) Рахунок складається бухгалтером.
- (2) Людина зачісується гребінцем.
- (3) Микола повернувся старим.

Визначте, чи відповідають подані речення одній глибинній структурі? Як можна перетворити структуру речення, для того щоб це визначити?

3. Як можна перефразувати, тобто синтаксично перетворити, наведені словосполучення, не змінюючи при цьому їхній зміст? Поясніть ці перетворення.

- (1) Вити по-вовчому.
- (2) Порада лікаря.
- (3) Командирський наказ.

4. Скільки речень (поверхневих структур) можна побудувати з наведеного вихідного речення (глибинної структури):

High voltage caused fatal failure.

Спробуйте зобразити вихідне речення (глибинну структуру) у вигляді дерева, як на рис. 4.3.

4.2. Структура і система: співвідношення понять

Система і структура є основоположними поняттями, як у класичній, так і сучасній структурній лінгвістиці. Першими⁴, хто поклав в основу опису мови та мовних елементів розуміння системних відношень є Бодуен де Куртене та

⁴ Принцип системного підходу ще обстоював Вільгельм фон Гумбольдт: у мові немає нічого одиничного, кожний факт мови треба розглядати в його зв'язках.

Фердинанд де Соссюр. Останній стверджував, що **мова** – це система, елементи якої утворюють ціле, а значущість одного елемента виникає тільки від одночасної наявності інших. У період раннього структуралізму **система** та **структура** практично не розрізняються, як наприклад у працях Бодуена де Куртене та Хеймана Штейнталя⁵. Якщо проаналізувати історію структурної лінгвістики, то ми можемо побачити, що в різні періоди (див. рис. 4.1.) спостерігалися різні бачення мови як системи і структури. У підрозділі 3.1 було подано одне із визначень цих понять. Але досі залишається проблема їхнього розмежування, тобто як вони співвідносяться одне з одним. Цей розділ не має на меті детальний аналіз усіх наявних дефініцій, виокремлюючи їхні переваги та недоліки. Проте необхідно зауважити, що в різних джерелах аналізовані поняття:

а) не розмежовуються, тому для їх позначення вживають один із термінів, або використовують обидва як синоніми;

б) розмежовуються, проте для їхнього позначення використовуються обидва терміни у двох однакових значеннях;

в) послідовно розмежовуються, але те, що у одного автора називається структурою, у іншого номінується системою.

Не зважаючи на розбіжності й непослідовності їхнього трактування, можна зробити певні узагальнення, які допоможуть чітко визначити та розрізнити аналізовані поняття. У період раннього структуралізму (1920-40-і рр.), який ми називаємо класичною структурною лінгвістикою, мовознавці ставили за головну мету інвентаризувати мовний матеріал, тобто виокремити одиниці мови та репрезентувати їх у вигляді класифікаційних схем, іншими словами виявити системну організацію мови. Проте такі дослідження по-різному здійснювалися у тих чи інших школах структуралізму. Так, у Празькому лінгвістичному гуртку до аналізу брали конкретні мовні одиниці (наприклад, фонеми) та розподіляли їх по класах на основі диференційних або функційних

⁵ Поняття **структури** вперше уведено в лінгвістику представниками Празької лінгвістичної школи.

ознак⁶. На відміну від пражських, копенгагенські структуралісти прагнули створити загальну теорію мови під назвою «Глосематика», використовуючи несубстанціональний («іманентний») підхід до визначення усіх основних одиниць та відношень між ними. Звідси випливає, що система мови має як субстантивні, так і несубстантивні властивості. Що означає опозиція субстантивність / несубстантивність мовної системи? Очевидно, що мова йде про **субстанцію**, матеріальну оболонку, тобто мовний інвентар, з яких складається мова та з яких будуються ті чи інші лінгвістичні об'єкти.

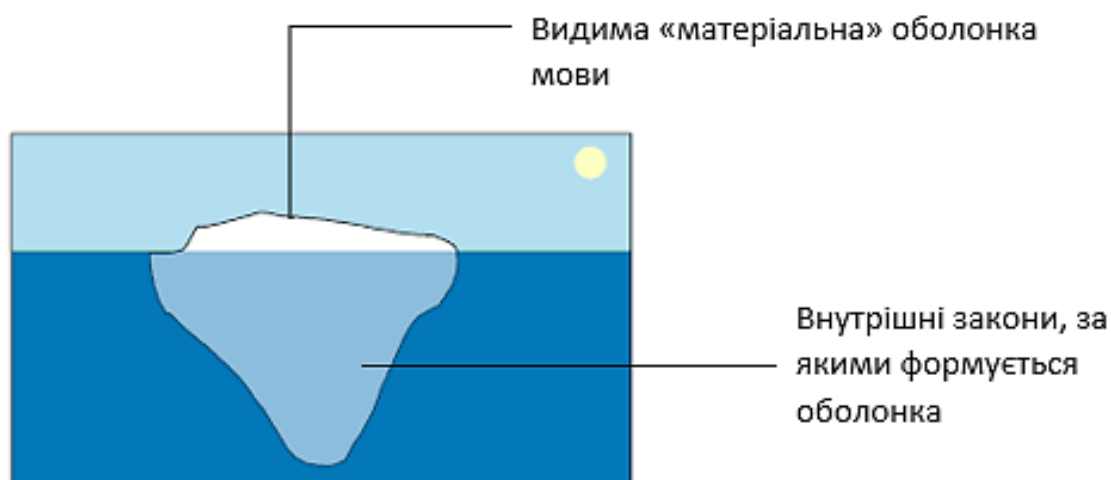


Рис. 4.4. Схематичне зображення системної організації мови.

До нематеріальної частини («не субстанції») належать внутрішні закони, правила, завдяки яким система мови формується та функціонує. У своїй теорії «Глосематиці» Луї Єльмслев передбачає наявність «неспостережуваної» частини мовної системи, яка лежить в основі будь-якої конкретній мові і є однаковою для всіх конкретних мов. Несубстантивна частина подається як мережа функцій або залежностей, що об'єднують мовні елементи (див. рис. 4.4). Розглянемо тепер таке питання: як корелює поняття **структури** з поняттям **системи**? Структура мови – це будова мови у вигляді: а) ієрархії елементів, від нижчого до вищого, якщо ми беремо певний рівень або б) у вигляді лінійної послідовності елементів, з яких формується інший елемент більш високого рівня (з фонем – морфемами, з морфем – лексемами і т. д.). Тобто **структура** – це

⁶ Американських дескриптивістів цікавила не стільки система мови, скільки синтагматичні відношення, на яких було засновано дистрибутивний аналіз.

спосіб організації мовних елементів у межах системи за рахунок різного типу відношень. **Систему** при цьому уявляють як класифікаційну схему, у якій показано одиниці та відношення між ними.

У період домінування генеративізму у 1950-70-х роках, коли почала активно розвиватися кібернетика та з'являтися генеративні граматики, **система мови** вже сприймається як *механізм, що утворює будь-які складні лінгвістичні об'єкти за допомогою певного набору граматичних правил*. Зокрема, така система моделює здібність носія мови утворювати граматично правильні речення. З огляду на це, Ролан Барт у статті «Структуралізм як діяльність» розкриває призначення сучасної структурної лінгвістики: «структуралізм є моделювальною діяльністю. Він ставить за мету відтворення об'єкта таким чином, щоб у подібній реконструкції виявилися правила функціонування цього об'єкта»⁷. У цьому контексті моделюється вміння людини володіти мовою. Наступним виникає питання: як співвідноситься поняття **структури** з таким розумінням **системи**? У складі механізму, що називається породжувальною граматиною наявний шаблон, або модель, що задає в абстрактному вигляді конструкцію породжуваних речень, показує складники речення, встановлює їхній порядок слідування, а також відображає взаємозв'язки між ними. Саме цей шаблон (модель) вважається **структурою**. Тобто, **структура** є елементом **системи**. Під час породження речення, система «накладає» на шаблон відповідний мовний матеріал, у результаті чого абстрактна конструкція перетворюється на конкретне речення (див. рис. 4.3).

За допомогою відповідних шаблонів можна змоделювати інші явища мови, такі як словотвір, словозміна, утворення словосполучень тощо. Виходячи з цього, Генадій Мельников вважає як **структуру**, так і **субстанцію** іманентними ознаками поняття системи. Така інтерпретація відображає властивості мови як засобу передавання інформації. До таких належать передусім субстантивні властивості мови, оскільки мова має реальні «матеріальні» втілення, зокрема в інформаційно-комунікаційних системах. Проте Володимир

⁷ R. Bartes. Structuralist activity. URL: <https://fliphtml5.com/tiioo/tlsz/basic>

Широков зазначає, що цього недостатньо, адже в означенні поняття системи активну роль відіграє ще один елемент, а саме – **суб'єкт**, здатний ідентифікувати, інтерпретувати структурно-субстанційні властивості та елементи системи і здійснювати з ними операції за певними правилами та алгоритмами. Суб'єктом може бути людина, спільнота, комп'ютерна система, нейромережа тощо. Отже, будь-яка система символічною репрезентується тріадою «структура – субстанція – суб'єкт»:

$$C = C + C + C,$$

де ліва частина C означає поняття системи, а права частина $C + C + C$ демонструє взаємодію твірних компонентів цього поняття (власне: структура, суб'єкт та субстанція). Системна тріада проявляє себе на всіх рівнях мови. Наприклад, на рівні словозміни:

- **структуру** системи репрезентує формальна модель, яка показує розклад мовної одиниці на незмінну (основу) та змінну (флексію) частини слова;
- **субстанція** визначається набором мовних елементів для побудови змінної частини відповідної відмінюваної одиниці;
- **суб'єкт** застосовує правила і алгоритми, за допомогою яких здійснюється побудова словозмінної парадигми за заданою структурою з використанням субстанційних елементів.

Виходячи з вищесказаного, можна подати таке визначення **системи мови**: *абстрактний механізм, що породжує лінгвістичні об'єкти будь-якого ступеня складності, маючи в основі три важливих компоненти: структуру, суб'єкт та субстанцію. Структура – головний елемент системи, який в абстрактному вигляді задає будову породжуваного лінгвістичного об'єкта. Така інтерпретація понять є основою для розроблювання різних комп'ютерних систем опрацювання природної мови.*

Контрольні запитання:

1. Як ви розумієте словосполучення: *синтаксична структура / синтаксична система, морфологічна структура / морфологічна система, семантична структура / семантична система?*

2. Яким чином репрезентується мова в сучасних інформаційних системах? Чому саме таким чином? Поясніть свою думку.

3. Якщо уявляти систему мови у вигляді тріади «структура – субстанція – суб'єкт», схарактеризуйте, що являє собою кожний з її елементів, коли говоримо, наприклад, про словотвірний, лексичний чи синтаксичний рівні?

Практичні завдання

1. Подано українські іменники, що належать, відповідно, першій, другій, третій та четвертій відмінам:

бідолаха, в'яз, ведмежа, вовча, глибочінь, глуш, голуб, далеч, дуб, зайча зброяр, здобич, зелень, кача, княжа, коровай, красень, мова, надія, небожа, палац, паморозь, північ, подорож, поріг, рік, рілля, розбишака, розкіш, сирота, собача, староста, сув'язь, суша, хмара, шпача.

Подайте у вигляді тріади «структура – субстанція – суб'єкт» систему розподілення українських іменників за відмінами. Для цього потрібно:

- провідмінювати усі іменники;
- визначити формальний показник (або показники), за якими відбувається цей розподіл;
- указати, що являє собою кожний елемент системної тріади та яку інформацію він містить;
- розписати покрокову процедуру розподілу іменників за відповідними відмінами.

2. У словотвірній системі сучасної української мови утворення фемінітивів (бандуристка, доповідачка, лікарка) відбувається за певними правилами:

автор, вершник, виборець, депутат, дизайнер, директор, диякон, доцент, засновник, керівник, лідер, очільник, патрон, плавець, поет, речник, рятувальник, учень, фахівець.

Для виконання цього завдання можна використовувати довідковий ресурс:
<https://kppk.com.ua/pedrada/Lozova.pdf>

Схарактеризуйте нижче елементи цієї системи:

Структура:

Субстанція:

Суб'єкт:

4.3. Граматика і словник як дві підсистеми мови

У попередньому розділі було зазначено, що система мови проявляє структурні та субстанційні властивості. Це означає, що вона має у своєму складі два компоненти – граматику і словник, які взаємодіють за посередництвом суб'єкта під час побудови тих чи інших лінгвістичних об'єктів. Тепер виникає питання, що являють собою словник і граматика? У мовознавчій літературі можна знайти чимало визначень цих понять. Зокрема, можна зустріти такі визначення поняття **словник**:

- книга, в якій в алфавітному чи тематичному порядку подано слова якоїсь мови (з тлумаченням, перекладом іншою мовою і т. ін.);
- упорядкований в алфавітному чи тематичному порядку лексикографічно опрацьованих перелік заголовних слів;
- довідкове зібрання у вигляді книги (тепер і в електронному варіанті) слів, а також інших мовних одиниць з інформацією про їхнє значення, вживання, будову, походження тощо;
- рукописне, друковане або електронне видання, до якого звертаються наразі потреби підвищити або перевірити рівень володіння лексикою.

Пропоновані визначення акцентують увагу на тому, що словник має на меті описати змістові (лексичні) характеристики мовних одиниць, тоді як насправді він містить також інформацію про їхні граматичні властивості, якщо взяти будь-який сучасний тлумачний словник іноземної мови (див. наприклад: <https://www.oxfordlearnersdictionaries.com/>). При цьому спосіб репрезентації граматичної інформації може бути різним: від простих словникових ремарок до ілюстративних прикладів, що показують, яка має бути будова речення, коли

слово використовується в тому чи іншому значенні. Це говорить про те, що словник, хоч і має свій предмет опису системи мови, тісно пов'язаний або інтегрований (за термінологію Юрія Апресяна) у граматику.

Виходячи з вищенаведеного, словник, з одного боку містить змістові властивості певної одиниці, а з іншого – наводить інформацію, за якою можна побудувати з цієї одиниці той чи інший лінгвістичний об'єкт (словотвірну або словозмінну парадигму, речення). Перейдемо тепер до розгляду наступного аспекту опису системи мови – **граматики**. Граматика здебільшого розглядається як наука (дисципліна), що вивчає будову мовних елементів та їх сполучення у слова, фрази та речення:

- лінгвістична наука, що вивчає будову мови, тобто будову і форми слова, речення і словосполучення;

- розділ мовознавства, що вивчає закономірності будови, змін та утворення форм слів певної мови (морфологія), а також словосполучень і речень (синтаксис);

- вся незвукова і нелексична організація мови, що виявляється в її морфологічних категоріях, одиницях і формах, тобто структурна основа мови, її устрій, лад;

- розділ мовознавства, що вивчає закономірності семантики, будови, творення, функціонування словоформ, лексем (слів), словосполучень, речень.

Вироблення принципів узгодження граматичного та лексичного описів набуло особливої актуальності у зв'язку з розвитком граматичних теорій, зокрема породжувальної граматики, пов'язаних із вивченням синтаксису, а також систем машинного перекладу. Хоча вже існують теорії, що присвячені проблемі зв'язку лексики з граматиною, але особливої актуальності вона набула завдяки активним дослідженням в галузі породжувальної граматики. Питання взаємодії між синтаксисом та лексикою активно обговорювалося в аспекті породжувальної граматики, засновником якої є Ноам Хомський. Першою його працею, що дала початок цьому напрямку в лінгвістиці, є монографія «Syntax Structures» (1957 р.), де мову розглянуто як механізм породження нескінченної

множини речень за допомогою скінченного набору граматичних засобів. Цей механізм було названо породжувальною граматиною, будова якої включає термінальний і нетермінальний словники (див. тему 10). Формальна граMATика відображає інтелектуальний механізм побудови граматично правильних речень відповідно до заданої моделі, використовуючи правила та інвентар одиниць. Термінальний словник містить інвентар одиниць, власне кажучи, мовний матеріал, що «накладається» на структуру речення. Хоча словник не містить семантичної інформації про слова, що використовуються для побудови речення, сам факт наявності словника свідчить про його необхідність для функціонування граматики.

Ноам Хомський вперше зробив спробу ввести семантичний компонент у граматику у своїй монографії «Aspects of the Theory of Syntax» (1965 р.). Цей компонент учений розуміє як сукупність правил семантичної інтерпретації, що приписують значення глибинним структурам. Ідеї взаємозв'язку синтаксису і семантики отримали подальший розвиток у роботі Ігоря Мельчука, Юрія Апресяна, Олександра Жолковського та Лідії Іорданської, присвяченій розробленню моделі «Смисл ↔ Текст» (див. підрозділ 12.3). Така модель – це складно організована сукупність правил, чисто механічне застосування яких має в ідеалі забезпечувати: 1) перехід від певного тексту досліджуваною мовою до формального опису змісту цього тексту, тобто до його семантичної репрезентації; 2) перехід від змісту, тобто від семантичної репрезентації до тексту, який є носієм цього змісту; якщо запропонований зміст можна передати більш ніж одним способом, то для нього мають будуватися всі синонімічні тексти (як це було показано в підрозділі 4.1). Важливим елементом моделі є Тлумачно-комбінаторий словник. Показовим для нього є те, що він містить таку граматичну інформацію:

- морфологічні характеристики мовної одиниці;
- модель керування;
- обмеження до моделі керування;
- ілюстративні приклади до моделей керування та їх обмежень.

Подальший розвиток теорія «Смисл ↔ Текст» отримала у працях Юрія Апресяна з інтегрального опису мови та системної лексикографії. Учений вважає, що словник і граматики мають бути узгоджені між собою як за типом розміщуваної інформації, так і за формою її подання. Тобто на рівні словника прописують лінгвістичні властивості, звернення до яких потребує правил граматики, а на рівні граматики побудова правил відбувається з урахуванням інформації про описувану лексему. Тлумачний словник як компонент інтегрального опису має включати два типи інформації – **класифікаційну** та **операційну**. Перша *стосується властивостей лексем, на основі яких вони стають об'єктами правил* (інформація про частину мови, синтаксичні характеристики). Друга – *це інформація про самі правила, тобто яким чином певний об'єкт має бути перетворений в інший об'єкт за наявності тієї чи іншої сукупності умов*.

Словник і граматики є двома важливими елементами системи мови: у словнику описано лексичні одиниці, а у граматиці сформульовано правила їх зміни (морфологія) і правила побудови складніших одиниць із простіших (словотвір, синтаксис). Для того, щоб лінгвістичний опис мови загалом був ефективним, необхідно досягти максимальної узгодженості словникового та граматичного описів. Отже, ці поняття можна інтерпретувати таким чином: **граматика** – *підсистема мови, що містить сукупність правил для породження цілісних об'єктів із мовних одиниць, враховуючи інформацію про їхні лексико-семантичні властивості зі словника*; **словник** – *підсистема мови, що містить інформацію про семантичні та інші властивості, що впливають на утворення граматиною лінгвістичного об'єкта*.

Контрольні запитання:

1. Чим пов'язано включення до сучасних словників граматичної інформації (наприклад, словотвір, словозміна, сполучуваність слів тощо)?
2. Яка інформація лексичного характеру має враховуватися граматиною для побудови, наприклад, словозмінної парадигми?

3. Чи можна стверджувати, що граматика і словник є автономними підсистемами мови?

4. Наведіть приклади зі знайомої вам мови, коли граматичне значення виражається лексичними засобами?

Практичні завдання

1. Нижче подано низку деяких визначень для дієслова «йти» та іменника «хліб» зі словника української мови (sum.in.ua). Оберіть для них параметри словозміни. Визначте кількісний склад їхньої парадигми залежно від лексичного значення.

2. Розпишіть парадигму словозміни відповідно до кожного значення.

ІТИ (ЙТИ), іду (йду), ідеш (йдеш); мин. ч. ішов (йшов), ішла (йшла), ішло (йшло), ішли (йшли); наказ. сп. іди (йди); недок.

1. Ступаючи ногами, пересуватися, рухатися, змінюючи місце в просторі (про людину або тварин).

2. Рухатися в якомусь напрямі (про види транспорту та інші засоби пересування).

3. Виходити, вирушати в певному напрямі, до кого-небудь. <...>

9. *за кого і без додатка.* Брати шлюб (про жінку); виходити заміж. <...>

15. Наставати, наближатися (про певний відрізок часу, період життя, розвитку, про якусь подію).

ХЛІБ

1. Харчовий продукт, що випікається з борошна. *Хліб усьому голова.*

// *перен.* Основний життєво необхідний харчовий продукт. *Картопля – половина хліба.*

// *Виріб з борошна певної форми (паляниця, буханець, коровай тощо).* *На столі лежав ясний да високий хліб.*

// *Тісто, приготовлене для випікання таких виробів.* *Стала хліб місити, піч затопила, в хаті убрала, – так мені любо!*

2. *тільки одн.* Зерно, з якого виготовляють борошно. Швидше й швидше летять ешелони, хліб, і вугіль, і нафту везуть...

3. *перев. мн. хліба, хліби, ів. Зернові культури. Хліби зійшли, Радіють люде, бога просять...*

// перев. одн. Необмолочені зернові. Коли б я хліба не возив, То що б стеріг тоді кудлатий.

// перев. одн. Урожай зернових. Моя мати між двома хлібами вмерла: одного не дочекала, а другий уже минувся.

4. *перев. одн., перен., розм. Засоби до існування; заробіток. Коли людина має цікаву роботу, ...коли вона не має побоювань про завтрашній хліб, фізично здорова...*

5. *перен., розм. Харчі, їжа. Чий хліб їси, під того й дудочку скачи!*

ТЕМА 5 СТРУКТУРНА І ТАКСОНОМІЧНА ЛІНГВІСТИКА

План заняття:

1. Синхронія та діяхронія. Статика і динаміка
2. Специфіка структурної і таксономічної лінгвістики
3. Системна тріада: структура, суб'єкт і субстанція

У цьому розділі здійснено порівняльний аналіз двох наукових дисциплін: таксономічної і власне структурної лінгвістики. Вони обидві є важливими для розуміння принципів формування та функціонування мови як системи. Таксономічна лінгвістика надає класичний погляд на систему мови як усталену, упорядковану сукупність мовних одиниць, яку ми можемо уявити у вигляді ієрархії елементів (статика). А структурна лінгвістика репрезентує її як механізм породження нових об'єктів у мові (динаміка). Не зважаючи на те, що ці дисципліни розв'язують різні завдання, вони пов'язані між собою. Перша упорядковує елементи мови відповідно до їхніх характеристик або властивостей. А друга моделює механізм утворення лінгвістичних об'єктів, використовуючи вихідні елементи з потрібними для цього характеристиками. Здобутки таксономічної та структурної лінгвістики використовують під час розроблення комп'ютерних систем, що опрацьовують природну мову. Як таксономічна, так і структурна лінгвістика оперує поняттями синхронії та діяхронії, а також їх співвідношенням із поняттями статичності та динаміки.

5.1. Синхронія та діяхронія. Статика і динаміка

Мова, як і будь-яка інша система, постійно перебуває в розвитку, тобто еволюціонує. Це супроводжується появою або зникненням тих чи інших одиниць та зміною характеру взаємозв'язків між ними. На різних рівнях мови (див. підрозділ 3.1) процес змін може відбуватися з різним ступенем інтенсивності. Наприклад, зміни на лексичному рівні є більш частими, порівняно зі змінами на фонологічному або морфемному рівнях. Звідси випливає, що мові як системі притаманна така властивість, як час. Те, що наявне в мові зараз, є результатом змін, що відбулися в ній у минулому. Ми можемо розглядати її такою, як вона є на певному етапі або у її історичній

перспективі, наприклад: від витоків до сьогодення. У першому випадку мова йде про синхронію, а в другому – про діяхронію. Отже, **синхронія** – *умовне виокремлення певного історичного етапу в розвитку мови, який береться як об'єкт лінгвістичного дослідження*. Синхронне вивчення передбачає аналіз мовних явищ в одному якомусь часовому проміжку: на сучасному етапі або в певну історичну добу, наприклад, у XIV або XVII ст., але без пояснення того, які зміни в попередні періоди розвитку мови привели до сучасного стану чи стану мови певної історичної доби. **Діяхронія** – *часовий зріз мови, при якому об'єктом лінгвістичного аналізу стає історичний розвиток мови*. Це означає, що при діяхронічному вивченні передбачається простежити весь шлях, який пройшов певний структурний елемент мови (звук, слово, речення).

Розмежування синхронії і діяхронії виникло у контексті історичних студій. Як відомо, у XIX столітті об'єктом лінгвістики вважали стародавні мови та пошуки прамови. Порівняльно-історична парадигма мовознавства попередніх століть, що перетворила дослідження мови в самостійну науку, мала за мету відновлення прамови. Можливість аналізувати мовні факти різних ступенів давності надав порівняльно-історичний метод. Теоретичний фундамент різниці між синхронією і діяхронією заклав Вільгельм фон Гумбольдт, хоча самі терміни він не використовував у своїх працях. Але на XIX – початку XX століть провідні лінгвісти Пилип Фортунатов та Іван Бодуен де Куртене, продовжуючи розвивати порівняльно-історичний підхід до вивчення мови, науково обґрунтували, що життя мови має як діяхронічну, так і синхронічну складову.

Починаючи з 1910-х рр. у лінгвістичній науці почала простежуватися нова тенденція, а саме – поворот до розгляду мови як безпосередньої даності (її синхронічної складової), не торкаючись питань про генезис окремих її елементів. Подібний підхід став реакцією проти пануючих у мовознавстві тієї епохи методів молодограматичної школи й отримав назву «лінгвістичний модернізм» (див. підрозділ 1.3). Ідеї синхронічного та діяхронічного аспектів продовжували розвиватися у парадигмі нової течії – структуралізму. У межах цього напрямку Микола Трубецькой, Роман Якобсон, Лев Щерба та Іван

Бодуен де Куртене розділяли лінгвістику на статичну (описову) і динамічну (історичну).

Але найбільш детально розглянув це питання та ввів у науковий обіг терміни **синхронія** і **діахронія** сам засновник структурної лінгвістики Фердинанд де Соссюр. Згідно з його твердженнями, синхронія передбачає вивчення системи мови з точки зору зв'язків між її складниками в один період часу. Відповідно, діахронічний аспект – це не історія окремих звуків і форм, а історія розвитку системних відношень. Основні історичні процеси в мові інтерпретуються відповідно до законів розвитку системи і визначаються прагненням до уніфікації мовних одиниць. Протиставлення і одночасно взаємозв'язок між синхронією та діахронією учений пояснює тим, що у кожний момент мовна діяльність передбачає й усталену систему, й еволюцію; в будь-яку хвилину мова є і живою діяльністю, і продуктом минулого. Для розкриття зазначених понять він пропонує схему (див. рис. 5.1), за якою розмежовуються мова і мовлення. Вони тісно пов'язані між собою та взаємно обумовлюють одне одного: мова необхідна, щоб мовлення було зрозумілим, а мовлення, у свою чергу, необхідне для того, щоб встановилася мова; історично факт мовлення завжди передує появленню мови. Отже, мова вивчається та розуміється як система, що існує не тільки в сьогоденні, але й у минулому. Синхронія та діахронія – це ще одна стала дихотомія у лінгвістиці, поряд із «мова і мовлення». Використовуючи цей подвійний принцип класифікації, Фердинанд де Соссюр вважає, що все діахронічне у мові стає таким лише через мовлення. Таким чином, мовлення є джерелом усіх змін: будь-які інновації з'являються спочатку в мовленні окремих осіб, а потім перетворюються на факт мови тоді, коли приймаються мовним колективом.



Рис. 5.1. Аспекти вивчення мови за Фердинандом де Сосюром.

Досліджуючи мову як систему, Фердинанд де Соссюр зазвичай віддавав перевагу синхронічному аспекту, хоча багато своїх робіт він присвятив діяхронії. Він вважав, що системні відношення можливі лише тоді, коли система мови вже усталена, і це дає змогу чітко виявити її складники та їхні взаємозв'язки⁸. Виходячи з розмежування синхронії та діяхронії, дослідник пропонує поділити лінгвістику на синхронічну (статичну) та діяхронічну (динамічну). Предметом **синхронічної лінгвістики** мають бути логічні та психологічні відношення, що пов'язують співіснуючі елементи у спільну систему, вивчаючи їх так, як вони сприймаються однією й тією самою колективною свідомістю. **Діахронічна** лінгвістика, навпаки, має вивчати відношення, що пов'язують елементи, які послідовно змінюють один одного в часі і не утворюють у своїй сукупності систему.

До другої половини XX століття синхронія ототожнювалася лише зі статикою, а діяхронія – лише з динамікою у мові. Але, на думку Себастьяна Шаумяна, синхронії притаманна як статика, так і динаміка. Синхронічні **статика** і **динаміка** є двома ярусами мови: до статичності належить мережа таксономічних відношень, а до динаміки – мережа глибинних відношень, пов'язаних із законами породження того чи іншого лінгвістичного об'єкта із найпростіших первинних елементів мови (рис. 5.2). До таксономічних учений зараховує відношення між елементами мови, що доступні дослідженню шляхом класифікаційних методів.



Рис. 5.2. Аспекти вивчення мови за Себастьяном Шаумяном.

⁸ За Іваном Бодуеном де Куртене, основна вимога діалектики – вивчати явища одночасно у їх зв'язку і у їх розвитку. Розрив синхронії та діяхронії, проголошений Фердинандом де Соссюром, двічі порушує цю вимогу, бо його синхронічне вивчення мови розглядає явища у їх зв'язку, але поза розвитком, а діяхронічне вивчення розглядає явища у їх розвитку, але поза зв'язком.

За наведеною вище схемою учений також вважає доцільним розрізняти поняття структури та квазіструктури. **Структурою** мови він називає *мережу лінгвістичних відношень, що належать до динаміки синхронії*, а **квазіструктурою** – *мережу лінгвістичних відношень, що належать до статички синхронії*. Префікс *квазі-* у терміні **квазіструктура** використано за аналогією до деяких математичних термінів, таких як «квазіпорядок», «квазігрупа» тощо. У таких випадках префікс означає послаблення вимог до поняття. Так, якщо поняття структури обов'язково передбачає наявність мережи саме глибинних відношень у мові, то квазіструктура не вимагає глибинності й розглядається як проста мережа відношень у мові (таксономія). Зупинимось детально на глибинності як характеристики структури. Це пов'язано з напрацюваннями Ноама Хомського, де він використовує термін **глибинна структура** (див. визначення у підрозділі 4.1). Як приклад, розглянемо речення «John hit the ball». Зазвичай ми його сприймаємо як послідовність слів у відповідному графічному або фонетичному (звуковому) оформленні. Умовно кажучи, це те, що ми бачимо на поверхні (рис. 5.3). А що власне приховано під поверхнею або за цією послідовністю слів? За нею лежить певна абстрактна конструкція, яка:

- 1) задає загальну будову речення;
- 2) будову окремих складників речення;
- 3) ієрархію складників у межах загальної будови речення.

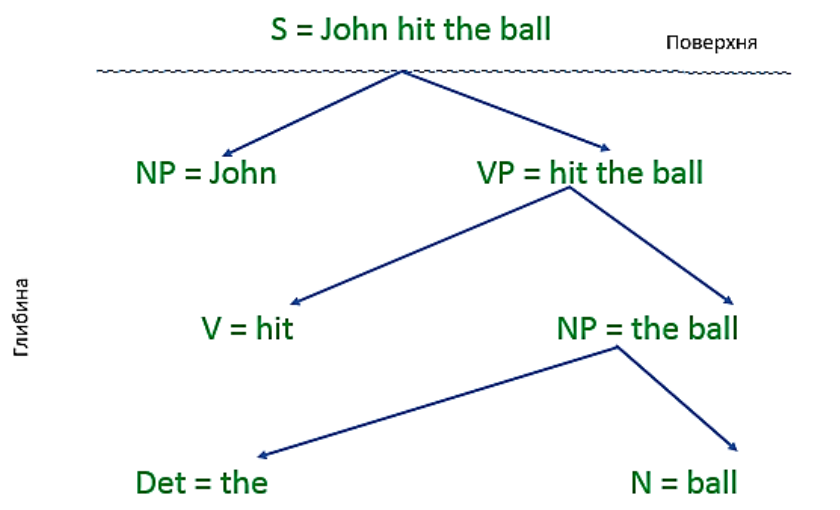


Рис. 5.3. Схематичне зображення глибинної структури.

Отже зазначений термін учений запропонував позначати засіб репрезентації в абстрактному вигляді будови речення⁹. Така будова дає можливість відобразити змістову близькість низки речень, що мають ті самі лексичні засоби та відрізняються граматичними характеристиками. З появою генеративної семантики замість нього стали використовувати термін **семантична структура**, у якому відображається одночасно і синтаксична, і змістова будова речення. Глибинність семантичної структури може використовуватися для об'єднання речень з однаковим змістом, але з відмінним лексичним складом і синтаксичною будовою (див. рис. 4.3). Варто зазначити, що термін **глибинна структура** застосовний не лише до синтаксису, а й також до морфології – словотвору та словозміні. Як синонім можна використовувати **формальна модель**.

Контрольні запитання:

1. Схарактеризуйте різницю в інтерпретації понять синхронії та діяхронії у доструктурній та структурній лінгвістиці.
2. Які приклади синхронічних та діяхронічних досліджень мови ви можете навести?
3. На якій підставі Себастьян Шаумян виокремлював таксономічну та структурну лінгвістику?
4. Чи є коректним твердження, що синхронія є лише статикою, а діяхронія – це лише динаміка? Поясніть свою думку.

Практичні завдання

Нижче наведені твердження, що стосуються мови або мовлення. Визначте відповідність кожного твердження до певного поняття та впишіть їх у таблицю:

абстрактне поняття, відношення підпорядкованості елементів «від нижчого до вищого рівня», притаманна динаміка, відношення визначають послідовність елементів, притаманна статика, містить набір правил побудови

⁹ З таким самим успіхом можна також використовувати словосполучення **абстрактна структура** замість цього терміна.

лінгвістичних об'єктів, реалізація лінгвістичних явищ, конкретне явище, можна репрезентувати як породжувальний механізм, можна зобразити структуру як дерево безпосередніх складників.

Мова	Мовлення

5.2. Специфіка структурної і таксономічної лінгвістики

У зв'язку з уведенням і протиставленням понять структури і квазіструктури Себастьян Шаумян вважає необхідним переглянути визначення предмета структурної лінгвістики. На його думку, **структурна лінгвістика** є наукою про динамічний аспект синхронії мови. У свою чергу, структурну лінгвістику він протиставляє **таксономічній лінгвістиці** – науці про статичний аспект синхронії мови. Як синоніми терміна «таксономічна лінгвістика» учений вважає доречним вживати терміни «статична лінгвістика» або «дескриптивна лінгвістика» (не плутати за американським дескриптивізмом, про який висвітлено у підрозділі 2.3). Зіставлення структурної лінгвістики та таксономічної лінгвістики дає змогу з'ясувати наочно специфіку обох наукових дисциплін. Динамічний аспект синхронії мови стосується процесу розробки формальних граматик, тобто гіпотетичного породжувального пристрою. Що ж до статичного аспекту, то вона охоплює питання побудови таксономічних, тобто класифікаційних схем. Отже, структурна лінгвістика пов'язана з генеративними граматиками, а таксономічна лінгвістика має справу з класифікаційними системами.

Таксономія передбачає розробку класифікації, яка відображає ієрархічну організацію мовних одиниць. Це відображається у вигляді ієрархії таксономічних рангів, пов'язаних відношенням включення від нижчого рангу до вищого. Так, одиниця фонологічного рівня складає одиницю морфемного

рівня. Кожний ранг (мовний рівень) складають відповідні класи об'єктів (таксони), які об'єднують мовні одиниці за спільними ознаками (див. рис. 3.1). Наприклад, фонеми об'єднуються у різні класи за участю голосу та способом їх творення. На побудові класифікаційних схем немає необхідності особливо зупинятися, оскільки питання класифікації частково розглядалися в темах 2 і 3. Будь-який механізм генерації поєднує три компоненти:

1. Набір вихідних елементів, з яких породжуються складні лінгвістичні об'єкти.

2. Набір операцій, які застосовуються до вихідних елементів для породження складних лінгвістичних об'єктів.

3. Набір структурних характеристик, приписуваних породжуваному лінгвістичному об'єкту.

Механізм генерації можна порівняти з грою «конструктор». У ній також наявні три компоненти: 1) набір вихідних деталей, із яких конструюються певні предмети; 2) інструкція, що вказує, за допомогою яких саме операцій можна отримувати з вихідних деталей конструйований об'єкт; 3) схеми, тобто структурні характеристики конструйованих предметів. Щодо понять **системи** і **структури**, у таксономічній лінгвістиці вони мають визначення, що подане в підрозділі 3.1, а у структурній – як у підрозділі 4.2, де мова йде про генеративні граматики. Також є і різниця у методах репрезентації понять. Таксономічна лінгвістика використовує для цього описовий метод, а структурна – метод моделювання.

Не менш важливим питанням Себастьян Шаумян вважає роль фонології у структурній і таксономічній лінгвістиці. По відношенню до породжувальної граматики, фонологія має розглядатися як допоміжна дисципліна, завданням якої є розроблення методів кодування глобальних знаків, якими оперує породжувальна граматика. Для граматики є важливими знаки лінгвістичних одиниць незалежно від їхнього фонологічного членування. Але, оскільки в природних мовах знаки членуються на фонологічні компоненти, структурна лінгвістика, для щоб бути адекватною теорією природних мов, має займатися

також фонологічними компонентами знаків. На фонологічному рівні, як і на граматичному, перед таксономічною та структурною лінгвістикою постають різні завдання. Для таксономічної завдання полягає в тому, щоб подати опис фонологічної системи самої по собі, тоді як структурна лінгвістика має розглядати фонологічну систему як кодування одиниць породжувальної граматики. У механізмах генерації маємо на вході інформацію про кінцеву кількість одиниць (морфем) та інформацію про правила утворення з цих одиниць нових лінгвістичних об'єктів (слів і речень). За кінцеве число кроків механізм генерації дає на виході граматично правильні слова та речення та забезпечує їх відповідними структурними характеристиками. На відміну від породжувальної граматики, **розпізнавальна граматика** (див. підрозділ 9.2) отримує на вході інформацію про кінцеву кількість допустимих у мові типів речень. Таким чином, за кінцеве число кроків розпізнавальна граматика дає на виході інформацію про спосіб породження кожного речення з первинних елементарних компонентів.

Оскільки у другій половині XX ст. лінгвістика стає підґрунтям для розроблення комп'ютерних систем опрацювання текстової інформації, окремо також слід розглянути взаємозв'язок структурної лінгвістики та кібернетики – науки про те, як складні керівні системи отримують, зберігають і опрацьовують інформацію. Якщо розглядати структурну лінгвістику у системі сучасних наук, її можна трактувати як лінгвістичну галузь кібернетики. Щоб обґрунтувати цю тезу, варто зупинитися на предметі кібернетики. Існує три основні типи машин: 1) які перетворюють один вид енергії на інший (наприклад, парові машини, газові турбіни); 2) які змінюють властивості та положення об'єктів праці (наприклад, текстильні, металообробні, транспортні машини); 3) які служать для перетворення одного виду інформації на іншу (це машини, які можуть бути використаними для опрацювання результатів наукових та технічних досліджень, автоматизації процесів пошуку інформації в науковій літературі, автоматизації перекладу з одних мов іншими тощо).

Саме останній тип машин входить до об'єкта досліджень кібернетики. Слід зазначити, що кібернетика має справу з такими машинами не в їх конкретному вигляді, а як із абстрактними системами високого ступеня узагальненості. Як зазначає академік Андрій Колмогоров, «кібернетика вивчає системи будь-якої природи, здатних сприймати, зберігати, опрацьовувати інформацію та використовувати її для управління і регулювання». Виходячи із загального уявлення про абстрактні кібернетичні машини, можна розглядати систему знаків будь-якої мови як різновид кібернетичної машини, що переробляє один вид лінгвістичної інформації в іншу. Звідси ясно, що структурна лінгвістика має вважатися лінгвістичною галуззю кібернетики.

Наукова теорія Ноама Хомського про формальні (породжувальну та розпізнавальну) граматики задала новий тренд у сучасній структурній лінгвістиці, який у науковій літературі називають генеративізмом. Характерним для нього є активне розроблення моделей породження лінгвістичних об'єктів, таких як слів, словосполучень, речень, висловлювань, текстів. У цьому контексті слід зазначити, що предметом моделювання у лінгвістиці є *не самі одиниці конкретного мовного рівня, а інтелектуальні процеси (механізми), що пов'язані з використанням мови*. Так, у теорії формальних граматик Ноам Хомський прагне змодельовати інтелектуальний механізм, що відповідає за володіння мовою. При цьому саме володіння мовою дослідник розуміє як здатність людини будувати граматично правильні речення. Згодом виникла потреба моделювати процес породження не тільки суто граматично правильних речень, а й таких висловлювань, що мають зміст. Подальші напрацювання в галузі лінгвістичного моделювання привели до появи теорії «Смисл ↔ Текст» Юрія Апресяна, Ігоря Мельчука, Олександра Жолковського та Лідії Іорданської (див. підрозділ 12.3).

Обидві теорії акцентують увагу на синтаксисі. Але проблеми моделювання мають охоплювати інші явища у системі мови, такі як словотвір та словозміна. З огляду на це особливої уваги заслуговує теорія системи мови Володимира

Широкова, яку він репрезентує у вигляді системної тріади, яка висвітлена у наступному розділі.

Контрольні запитання:

1. У чому полягає системний підхід у лінгвістиці?
2. Схарактеризуйте особливості опису системи мови у таксономічній та структурній лінгвістиці.
3. У яких галузях використовують напрацювання таксономічної та структурної лінгвістики?
4. Чим зумовлене інше бачення системи мови в структурній лінгвістиці, якщо порівнювати з таксономічною лінгвістикою?

Практичні завдання

1. Заповніть таблицю «Таксономічна / структурна лінгвістика»

Таксономічна лінгвістика	Структурна лінгвістика
Предмет	
Метод опису мови	
Інтерпретація поняття системи	
Інтерпретація поняття структури	

2. Розподіліть подані нижче слова на парадигматичні типи, вважаючи що два слова входять у той самий парадигматичний тип, якщо їх словозміна відбувається за однаковими граматичними категоріями або параметрами (приклади граматичних категорій – відмінок, число, час тощо):

верблюд, який, хто, фарба, лисий, мужва, хтось, п'ятий, п'ять, синій, три, чотири, депо, кольє, таксі, штани, ягода.

Для кожного типу вкажіть, за якими граматичними параметрами змінюються слова, що входять до нього:

Тип	Параметри	Слова

Розпишіть словозміну за параметрами для одного слова з кожного виокремленого парадигматичного типу.

З'ясуйте, чи за всіма параметрами змінюються слова у визначених парадигматичних типах.

5.3. Системна тріада: структура, суб'єкт і субстанція

Як було зазначено в підрозділі 4.3, систему мови можна репрезентувати символічною тріадою «структура – субстанція – суб'єкт» ($C = C + C + C$), де C у лівій частині означає поняття системи, а права частина рівняння демонструє взаємодію твірних компонентів цього поняття (власне: **Структура**, **Суб'єкт** та **Субстанція**). Під час свого функціонування система проявляє структурні та субстантивні властивості, як відзначає Геннадій Мельніков¹⁰. Структурні характеризують скільки вихідних елементів застосовуються для утворення того чи іншого лінгвістичного об'єкта, яким має бути порядок слідування цих елементів, якими відношеннями вони об'єднуються в межах конструйованого об'єкта. **Структура** подається як певна абстрактна модель, на яку накладається мовний матеріал (**субстанція**), і в результаті маємо конкретний лінгвістичний об'єкт у вигляді словоформи, словосполучення, речення тощо. При цьому Володимир Широков наголошує, що задання структури та вибір відповідних засобів матеріального її оформлення відбувається третім елементом тріади, **суб'єктом** за відповідними правилами, алгоритмами.

Системна тріада проявляє себе на всіх мовних рівнях. Продемонструємо як вона функціонує на рівні словозміни. Спочатку необхідно дати визначення, що таке словозміна і зобразити її у вигляді моделі. Отже, **словозміна** – це *сукупність граматичних форм тієї самої лексеми, що утворюються шляхом додавання певного набору закінчень (квазіфлексій) до її незмінюваної частини (квазіоснови¹¹)*. Виходячи з цього, словоформу як елемент парадигми можна подати у такому вигляді:

$$x = c(x) * f(x),$$

¹⁰ V. Shyrokov. Language. Information. System. K.: Академперіодика, 2021. С. 75

¹¹ Префікс *квазі-* означає, що терміни «основа» і «флексія» не завжди відповідають їхньому визначенню в традиційній лінгвістиці.

де x – словоформа, $c(x)$ – квазіоснова словоформи x , $f(x)$ – квазізакінчення словоформи x .

Так, аналізуючи реалізацію дієвідміни, можна побачити, що квазіфлексії можуть охоплювати все слово (у випадку суплетивних форм¹²), частину кореня або збігатися із закінченням. У випадку суплетивних форм маємо нульову квазіоснову. Нехай для будь-якої одиниці x існує відповідна квазіоснова $c(x)$ та набір відповідних квазіфлексій $f(x)$. Наведемо у таблиці 5.1 відповідні елементи трьома мовами.

Таблиця 5.1. Квазіоснови і квазіфлексії в різних мовах.

Особа / число	Українська		Англійська		Іспанська	
	$c(x)$	$f(x)$	$c(x)$	$f(x)$	$c(x)$	$f(x)$
1 особа однини	ма	ю	ha	ve	t	engo
2 особа однини	ма	єш	ha	ve	t	ienes
3 особа однини	ма	є	ha	s	t	iene
1 особа множини	ма	ємо	ha	ve	t	enemos
2 особа множини	ма	єте	ha	ve	t	enéis
3 особа множини	ма	ють	ha	ve	t	ienen

Введемо додаткові позначення: $[x]$ – повна парадигма слова x ; $[f(x)_1 \dots f(x)_n]$ – набір квазіфлексій, що входять до складу парадигми $[x]$. Тоді модель словозміни набуває такого вигляду:

$$[x] = c(x) * [f(x)_1 \dots f(x)_n]$$

Отже, структура системи репрезентована розкладом мовної одиниці на константну та змінну частини слова; субстанція – набором елементів (квазіфлексій) для побудови змінної частини відповідної відмінюваної одиниці; суб'єкт – правилами та алгоритмами, за допомогою яких здійснюється побудова словозмінної парадигми з набору структурних та субстанційних елементів. Розпишемо елементи системної тріади, яка моделює процес утворення словозмінної парадигми:

- структура – формальна модель словозміни: $[x] = c(x) * [f(x)_1 \dots f(x)_n]$;
- субстанція – матеріальне оформлення $c(x)$ та $f(x)$;
- суб'єкт – алгоритм утворення парадигми $[x]$ з $c(x)$ та $[f(x)_1 \dots f(x)_n]$;

¹² **Суплетивізм** – утворення граматичних форм одного й того ж слова від різних основ або коренів.

Говорячи про словозміну, необхідно враховувати, що квазіоснова $c(x)$, так само як і квазіфлексія $f(x)$ можуть набувати таких значень:

– $c(x) = \emptyset$, якщо слово має суплетивну форму, як наприклад, англійське дієслово *to be* (бути) у другій та третій особі однини теперішнього часу дійсного способу: *am, is, are; was i were* у Past Simple (див. таблицю 5.2);

Таблиця 5.2. Квазіоснови і квазіфлексії дієслова *to be*.

Особа / число	$c(x)$	$f(x)$
1 особа однини	$\emptyset$	<i>am</i>
2 особа однини	$\emptyset$	<i>are</i>
3 особа однини	$\emptyset$	<i>is</i>
1 особа множини	$\emptyset$	<i>are</i>
2 особа множини	$\emptyset$	<i>are</i>
3 особа множини	$\emptyset$	<i>are</i>

– $f(x) = \emptyset$, якщо компонент x є незмінюваним. Порівняйте, наприклад, дієслово *must* в усіх особових формах простого теперішнього часу (див. таблицю 5.3);

Таблиця 5.3. Квазіоснови і квазіфлексії дієслова *must*.

Особа / число	$c(x)$	$f(x)$
1 особа однини	<i>must</i>	$\emptyset$
2 особа однини	<i>must</i>	$\emptyset$
3 особа однини	<i>must</i>	$\emptyset$
1 особа множини	<i>must</i>	$\emptyset$
2 особа множини	<i>must</i>	$\emptyset$
3 особа множини	<i>must</i>	$\emptyset$

– $c(x) = x - f(x)$, якщо x має змінювану квазіфлексію, як, наприклад, у дієслова *to steal* в особових формах минулого часу (***steal*** – ***stole***) та пасивного дієприкметника (***steal*** – ***stolen***), як це показано у таблиці 5.4.

Таблиця 5.4. Квазіоснови і квазіфлексії дієслова *steal*.

Особа / число	$c(x)$	$f(x)$	$c(x)$	$f(x)$
1 особа однини	<i>st</i>	<i>eal</i>	<i>st</i>	<i>ole</i>
2 особа однини	<i>st</i>	<i>eal</i>	<i>st</i>	<i>ole</i>
3 особа однини	<i>st</i>	<i>eals</i>	<i>st</i>	<i>ole</i>
1 особа множини	<i>st</i>	<i>eal</i>	<i>st</i>	<i>ole</i>
2 особа множини	<i>st</i>	<i>eal</i>	<i>st</i>	<i>ole</i>
3 особа множини	<i>st</i>	<i>eal</i>	<i>st</i>	<i>ole</i>

Така система словозміни була реалізована у вигляді низки електронних словників, розроблених співробітниками Мовно-інформаційного фонду НАН України для української та турецької мов (ulif.org.ua), а також співробітниками й здобувачами Міжвідомчого центру прикладної лінгвістики на базі кафедри інтелектуальних комп'ютерних систем НТУ «ХПІ». Наприклад, електронний словник словозміни англійських дієслів, (рис. 5.3), який подає парадигму як для окремого дієслова, так і для дієслів певного словозмінного класу.

	Active voice						Passive voice					
	singular			plural			singular			plural		
	1st person	2nd person	3rd person	1st person	2nd person	3rd person	1st person	2nd person	3rd person	1st person	2nd person	3rd person
Fut...	will finish	will finish	will finish	will finish	will finish	will finish	will be finis...	will be finis...	will be finis...	will be finis...	will be finis...	will be finis...
Fut...	will be finis...	will be finis...	will be finis...	will be finis...	will be finis...	will be finis...						
Fut...	will have fin...	will have fin...	will have fin...	will have fin...	will have fin...	will have fin...	will have be...	will have be...	will have be...	will have be...	will have be...	will have be...
Fut...	will have be...	will have be...	will have be...	will have be...	will have be...	will have be...						
▼ Subjecti...												
Simple	finished	finished	finished	finished	finished	finished	were finished	were finished	were finished	were finished	were finished	were finished
Progr...	were finishing	were finishing	were finishing	were finishing	were finishing	were finishing						
Perfect	had finished	had finished	had finished	had finished	had finished	had finished	had been fi...	had been fi...	had been fi...	had been fi...	had been fi...	had been fi...
Perfe...	had been fi...	had been fi...	had been fi...	had been fi...	had been fi...	had been fi...						
Imperati...		finish		finish	finish			be finished		be finished	be finished	

Рис. 5.3. Електронний словник словозміни англійських дієслів.

Проте сама реалізація словозміни за поданою структурою залежатиме певною мірою від субстантивних характеристик конкретної мови. Перш за все, маються на увазі параметри словозміни, які визначають відповідний набір квазіфлексій та місце слова у цій системі словозміни. Так, якщо дієслова змінюються за особою, часом та способом, то у випадку з іменниками параметри будуть іншими. На сьогодні теорії Ноама Хомського, Володимира Широкова та інших є теоретичною основою для розроблення систем автоматизованого перекладу, питально-відповідальних систем (чат-боти) та взагалі у всіх систем, що використовують під час опрацювання природних мов.

Контрольні запитання:

1. Порівняйте погляди структуралістів (Фердинанда де Соссюра) та сучасних лінгвістів (Себастьяна Шаумяна, Володимира Широкова) на системно-структурну організацію мови.

2. За якими параметрами відбувається словозміна? Від чого залежить реалізація словозміни тим чи іншим словом?

3. Схарактеризуйте динаміку поглядів на структурну лінгвістику від її зародження до сьогодення.

4. У якому вигляді відбилися синхронічні та діяхронічні погляди у теорії системи мови Володимира Широкова?

Практичні завдання

1. Подайте у вигляді таблиці власні приклади словозмінної парадигми (будь-якою мовою) відповідно до теорії системи мови Володимира Широкова, так щоб були дотримані такі умови:

$$- c(x) = \emptyset;$$

Особа / число	$c(x)$	$f(x)$
1 особа однини	$\emptyset$	
2 особа однини	$\emptyset$	
3 особа однини	$\emptyset$	
1 особа множини	$\emptyset$	
2 особа множини	$\emptyset$	
3 особа множини	$\emptyset$	

$$- f(x) = \emptyset;$$

Особа / число	$c(x)$	$f(x)$
1 особа однини		$\emptyset$
2 особа однини		$\emptyset$
3 особа однини		$\emptyset$
1 особа множини		$\emptyset$
2 особа множини		$\emptyset$
3 особа множини		$\emptyset$

$$- c(x) = x - f(x);$$

Особа / число	$c(x)$	$f(x)$
1 особа однини		
2 особа однини		

3 особа однини		
1 особа множини		
2 особа множини		
3 особа множини		

2. Подано займенники і займенникові прислівники угорської мови разом із перекладами українською:

amely – який, hogy – як?, néhány – декілька, valamikor – коли-небудь, néhol – де-інде, aki – хто.

Перекладіть з угорської:

valamely, hol, ahogy, némely, valaki, hány

Визначте алгоритм утворення займенників та займенникових прислівників. Керуючись визначенням алгоритмом, утворіть від поданих українських відповідні угорські займенники та займенникові слова:

скільки-небудь, де, як-небудь, коли?, де-небудь, інколи.

Побудуйте формальну модель, яка відображає морфологічну структуру займенників та займенникових прислівників. Схарактеризуйте складники системної тріади, що репрезентує словотвір: структура, суб'єкт і субстанція.

3.* Нижче подано фрази угорською мовою:

Az asztalon a térkép van.	На столі мапа.
Az asztalokban a térképek vannak.	У столах мапи.
A füzetnél az újság van.	Біля зошиту газета.
Az újságokon a füzetek vannak.	На газетах зошити.
Az ablakoknál a pad van.	Біля вікон лава.
A székeken a kasok vannak.	На стільцях кошики.

Побудуйте механізм породження речень угорською мовою. Для цього необхідно:

- Визначити граматичні правила побудови речень.
- З'ясувати засоби вираження категорій відмінка, числа.
- Побудувати синтаксичну структуру речення угорською мовою.
- Визначити складники механізму (у вигляді тріади).

ТЕМА 6 СТРУКТУРНА ЛІНГВІСТИКА

СЕРЕД ІНШИХ ГАЛУЗЕЙ МОВОЗНАВСТВА

План заняття:

1. Структурна лінгвістика і порівняльне мовознавство
2. Структурна лінгвістика і навчання мові
3. Структурна лінгвістика і процес перекладу

На сучасному етапі розвитку структурної лінгвістики виникла необхідність у новому формулюванні як її завдань, так і її предмету, що зумовлено зміною наукової парадигми. Так, предметом стає динамічний аспект синхронії мови: тобто науковці зосереджують увагу на інтелектуальних процесах, пов'язаних із володінням мовою. Під володінням мається на увазі здатність утворювати будь-який лінгвістичний об'єкт із вихідних елементів. Побудова такого об'єкта відбувається за певними шаблонами або моделями, які у формальному вигляді задають його загальну структуру. Можна сказати, що сучасна структурна лінгвістика вивчає та описує механізми породження мовних явищ на структурах, репрезентованих для наочності у вигляді моделей. Завдяки цьому, її методика використовується також в інших мовознавчих галузях, які також вимагають структурного принципу опису мови. Ознайомимося з деякими з них.

6.1. Структурна лінгвістика і порівняльне мовознавство

Порівняльне мовознавство – галузь лінгвістики, що присвячена зіставленню різних мов з метою виявлення спільних рис на основі їхньої генетичної спорідненості або спільних рис у будові та принципах функціонування безвідносно до їхньої спорідненості. Виходячи з цього, розрізняють, відповідно, порівняльно-історичне і типологічне мовознавство. Застосування поряд зі структурними методами інших прийомів і методів дослідження, у тому числі методів порівняльного мовознавства, дало можливість удосконалити методику строгого опису системи мови. Розглянемо роль структурної лінгвістики у цих двох напрямках.

Структурна лінгвістика належить до синхронічних дисциплін. Але її методи застосовні також до вивчення історії мови шляхом так званої

реконструкції. Такий тип реконструкції розширює систему мови на кілька систем, одна з яких розглядається як найдавніша (рис. 6.1), а інші – як результат перетворень найдавнішої системи із застосуванням до неї відповідних операцій. Методи структурної лінгвістики починають застосовувати в 1920-30-х рр. у своїх роботах з порівняльно-історичного мовознавства Роман Якобсон, Андре Мартіне, Єжи Курилович та інші. На відміну від молодограматиків (див. підрозділ 1.3), котрі також здійснювали порівняльно-історичні студії, об'єктом уваги структуралістів були не окремі мовні факти, а система мови в цілому, тобто які історичні зміни сприяли появі системи мови у сучасному вигляді.

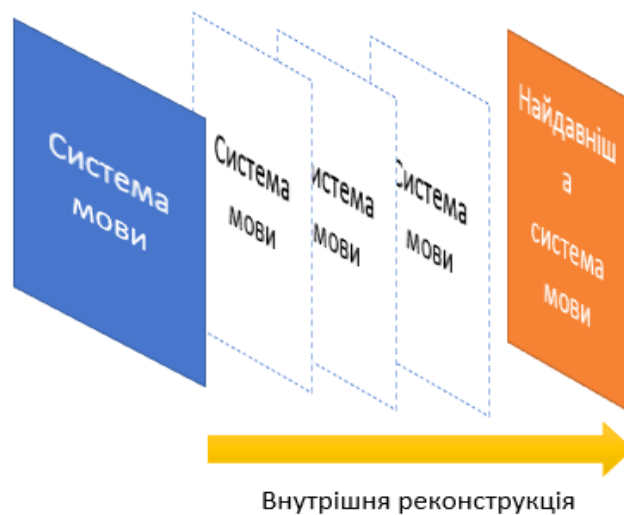


Рис. 6.1. Розширення системи мови від сучасної до найдавнішої.

Реконструкцією називають *комплекс прийомів та процедур відновлення незасвідчених мовних станів, форм, явищ шляхом історичного порівняння відповідних одиниць окремої мови, групи та сім'ї мов*. Предметом реконструкції можуть бути будь-які елементи мови – фонemi, морфemi, просодія, синтаксис, парадигми та семантичні поля, діалектні членування та мовні стани в цілому. Методика реконструкції має дати відповіді на два питання:

1. Як отримати інформацію про прамову в цілому або про окремі рівні (фонологічний, словотвірний, лексичний, синтаксичний) на основі даних про існуючі мови-нащадки?

2. Яким є ступінь достовірності одержуваних при цьому результатів та як підвищити цю достовірність?

Кожна окрема зміна не є хаотичною, а охоплює всю систему мови, тобто ці зміни є регулярними. Дослідник має виявити риси прамови, які збереглися в мовах-нащадках, відокремити їх від пізніших трансформацій та надати їм лінгвістичну інтерпретацію. Говорячи про реконструкцію прамови, насамперед мається на увазі відтворення її словникового складу. Її основу, як правило, складають слова, наявні в переважній більшості мов світу, що здебільшого позначають назви частин тіла, деяких природних об'єктів. На основі реконструкції цих слів можна скласти певну систему фонетичних відповідників, що дає змогу реконструювати мову, якою говорили предки того чи іншого народу. Крім фонетичної та лексичної систем, об'єктами реконструкції також є граматична система – морфологія та синтаксис.

Дуже важливо реконструювати основні парадигматичні системи граматики прамови, до яких, наприклад, належить система іменного відмінювання. Вона складається з певної кількості відмінків, яку можна реконструювати на підставі відомих сучасних мов. Реконструювавши два компоненти будь-якої відмінкової форми (це, власне, саме звучання форми та значення форми), можна тим самим перейти поступово до відтворення усієї системи. До виникнення структурного напрямку в мовознавстві цей метод використовували науковці та дослідники до різних мов з початку XIX сторіччя (Франц Бопп, Расмус-Крістіан Раск, Якоб Грім, Август Шлейхер, Олександр Востоков, Олександр Потебня та інші). Так, Франц Бопп досліджував порівняльно-історичним методом дієвідмінювання основних дієслів у санскриті, грецькій, латинській та готській мовах, зіставляючи основи та флексії. З огляду на це, Расмус Крістіан Раск зауважує, що граматичні відповідники є важливішими за лексичних, оскільки ніколи не буває запозичення словозміни і, зокрема, флексій. Слов'янські мови досліджували цим методом Олександр Востоков і Олександр Потебня. На матеріалах української мови Олександр Потебня у своїй праці «Із записок

руської граматики» подає порівняльно-історичне дослідження граматичної будови усієї системи східнослов'янських мов у їх зв'язках з іншими мовами.

Розберемо приклад із сучасної наукової розвідки Ольги Крижанівської¹³, присвяченій системі українського відмінювання іменників першої відміни, витоки якої спостерігаються у системі праслов'янської мови. Ми розглянемо лише форму однини. Для початку наведемо сучасну парадигму для слів *риба*, *пісня*, *буква*, *дочка*:

Н.	риба	пісня	буква	дочка
Р.	риби	пісні	букви	дочки
Д.	рибі	пісні	букві	дочці
З.	рибу	пісню	букву	дочку
О.	рибою	піснею	буквою	дочкою
М.	на / у рибі	на / у пісні	на / у букві	на / у дочці
Кл.	рибо	пісне	букво	дочко

Як зазначає дослідниця, на вибір відмінкових флексій іменників сучасної першої відміни впливає кінцевий приголосний основи (твердий нешиплячий, м'який чи твердий шиплячий), а також належність іменника до категорії істот чи неістот (це проявляється у формах множини). За її спостереженнями, в історичному плані система відмінкових закінчень першої відміни склалася шляхом уніфікації відмінкових закінчень колишніх іменників переважно жіночого та частково чоловічого роду, що належали до різних праслов'янських типів відмінювання¹⁴:

- перший тип: усі іменники *-*а-*, *-*ја-* основ: *рука*, *риба*, *земля*, *воєвода*;
- другий тип: більшість іменників з давньою основою на *-*ǫ-*: *тиква*, *церква*, *хоругва*, *буква*, *кроква*;
- третій тип: деякі іменники жіночого роду на *-*ї-* основи, які набули закінчення *-а(-я)*: *долоня*, *миша*, *пісня*.
- четвертий тип: іменники на приголосний *-*г-*: *дочка*.

¹³ Повний текст статті доступний за посиланням: <https://core.ac.uk/download/pdf/53036482.pdf>

¹⁴ Наведено транскрипцію латиною відповідно до робіт учених, присвячених дослідженню праслов'янської мови.

Відповідно, у праслов'янській мові існувало чотири типи іменного відмінювання для наведених вище груп іменників. Наведемо у таблиці повну парадигму на прикладі деяких іменників.

Відм.	I тип		II тип	III тип	IV тип	Сучасні закінчення I відміни
	-*а-	-*ја-	-*ǫ-	-*ĩ-	-*r-	
Н.	рыба	земля	буки	пѣснь	дѣчери	-а, -я
Р.	рыбы	землѣ	букѣвѣ	пѣсни	дѣчере	-и, -і
Д.	рыбѣ	земли	букѣви	пѣсни	дѣчери	-і
З.	рыбу	землю	букѣвь	пѣснь	дѣчерь	-у, -ю
О.	рыбою	земелею	букѣвь(и)ю	пѣснь(и)ю	дѣчерь(и)ю	-ою, -ею
М.	рыбѣ	земли	букѣвѣ	пѣсни	дѣчери(-є)	-і
Кл.	рыбо	землє	буки	пѣсни	дѣчерь	-о, -е, -ю

Отже, для визначення давньої парадигми для сучасного українського слова, необхідно з'ясувати, яку він мав основу в праслов'янській мові (для цього можна використовувати, наприклад, етимологічні словники):

а) іменники чоловічого або жіночого роду з основою на -а, -я (типу *рука, нога, риба, земля*; праслов'янською: *рука, нога, рыба, земля*) мають парадигму I типу;

б) іменники жіночого роду, що мають в основі -кв-, -гв- (*буква, тиква хоругва, церква*) та деякі інші (*любов, свекруха*), у праслов'янській мові мають закінчення на -кы, -гы (*буки, тыки, хоругы, цѣркы*) або -ы (*любы, свекры*) відмінюються за II типом;

в) іменники жіночого роду типу *долоня, миша, пісня*, що мали давню флексію -ь (*долонь, мышь, пѣснь*), відмінюються за III типом;

г) український іменник *дочка* у праслов'янській мові відмінюється за IV типом (спільнослов'янською *дѣчери*).

До парадигматичних систем у мові також зараховують систему іменних класів, систему займенників, які теж підлягають реконструкції. Не зважаючи на відмінності між парадигматичними системами прамови та сучасної мови, загальні закономірності все-таки існують. У результаті реконструюється автентична система, верифікована на базі мовного матеріалу. Саме в цьому полягає основне завдання реконструкції граматики прамови.

Найскладнішою є реконструкція синтаксису, яка є окремим розділом, що вивчає питання побудови речень. Складність полягає в тому, що невідомо, як саме могли будуватися речення у прамові, тобто невідомою є структура речення. Одним із можливих шляхів відтворення синтаксису прамови можуть бути, наприклад, стійкі фрази ідіоматичного характеру. Для багатьох давніх індоєвропейських мов відтворено одні й ті самі стійкі словосполучення, які практично однаково будуються кількома мовами, але цього недостатньо для того, щоб припустити, як будувалися фрази в індоєвропейській прамові. Синтез структурної лінгвістики та порівняльно-історичного мовознавства надає ширші можливості розуміння, розшифрування та переклад давніх текстів, а також сприяє більш точному трактуванню текстів та розв'язанню питань шифрування.

Перейдемо до розгляду зв'язків структурної лінгвістики з типологічним мовознавством. Структурна лінгвістика – окремий розділ лінгвістичної науки, який, досліджуючи загальні принципи лінгвістичної структури, будує різного роду породжувальні граматики як абстрактні лінгвістичні механізми. **Типологія мов** – це дисципліна, що має своїм завданням систематизацію структурних типів конкретних мов світу. Розв'язання такого завдання можливе за допомогою породжувальних граматик у якості певних мов-еталонів, які є предметом структурної лінгвістики та водночас одиницями виміру для типологічного порівняння мов. Прийнявши ту чи іншу систему породжувальної граматики¹⁵ можна отримати абстрактні мовні конструкції, що будуть універсальними для двох або більше порівнюваних мов. Використовуючи у мові-еталоні ті чи інші правила перетворення, цілком можливо отримати умовну будову конкретної мови, яка може охоплювати як морфологію, так і синтаксис. Варто наголосити, що мова-еталон має трактуватися як гіпотетичний інваріант мов світу.

¹⁵ Термін породжувальна граMATика у цьому контексті можна використовувати в інтерпретації теорії Ноама Хомського або теорії системної тріади Володимира Широкова, або будь-яку іншу математичну модель, яка відтворює граматичну будову, характерну для тієї чи іншої мови.

Наведемо простий приклад використання породжувальної граматики як мови-еталону для моделювання синтаксичних особливостей мов номінативного та ергативного типу, тобто особливостей побудови речень у цих мовах. **Номінативними** називають мови такого типу, в яких будова речення спрямована на розрізнення об'єкта та суб'єкта дії. Так, у реченні «Кіт спіймав мишу» *кіт* є суб'єктом дії, а *миша* – об'єктом дії. У флективно-синтетичних мовах опозиція відображається за допомогою відмінкових закінчень, а в аналітичних – порядком слів (як в англійській «The cat caught a mouse»). До номінативних належать усі мови індоєвропейської, тюркської та монгольської сімей, а також афразійської та уральської макросімей. В **ергативних мовах** опозиція формується шляхом протиставлення більш активних, менш активних та самотійних дій. У таких мовах аналізоване речення матиме вигляд на кшталт «Коту миша піймалася», для них характерним є поширене використання пасивного стану. До ергативних належать іберійсько-кавказькі мови, баскська мова, папуаські, австралійські, чукотсько-камчатські та ескімосько-алеутські мови.

Побудуємо мову-еталон, на основі якої можна отримати структуру речення для мов як номінативного (1), так і ергативного типів (2). Розберемо синтаксичну структуру речення в обох мовах:

(1) Кіт спіймав мишу;

(2) Katuak sagua harrapatu zuen (досл. баскськ.: Коту миша піймалася).

Можна побачити, що речення (1) і (2) мають однакову будову, тобто можна виокремити в них групу підмета і групу присудка, тобто $S \rightarrow NP + VP$, де:

(1) $NP \rightarrow$ Кіт, $VP \rightarrow$ спіймав мишу;

(2) $NP \rightarrow$ Katuak, $VP \rightarrow$ sagua harrapatu zuen.

В обох реченнях групу підмета складає іменник, що можна відобразити у вигляді такого правила: $NP \rightarrow N$. Щодо групи присудка, то у мові (1) вона репрезентована дієсловом та іменником $VP \rightarrow V + N$; у мові (2) група присудка включає іменник та групу присудка $VP \rightarrow N + VP$:

(1) $VP \rightarrow V + N$, де $V \rightarrow$ спіймав, $N \rightarrow$ мишу;

(2) $VP \rightarrow N + VP$, де $N \rightarrow \text{sagua}$, $VP \rightarrow \text{harrapatu zuen}$.

Зупинимося детально на групі присудка $VP \rightarrow \text{harrapatu zuen}$. Її складниками є дієслово та допоміжне дієслово (позначає давність завершення дії, як *had* в англійській мові). Формально аналізована група присудка набуває такого вигляду:

$VP \rightarrow V + Vaux$, де $V \rightarrow \text{harrapatu}$, а $Vaux \rightarrow \text{zuen}$.

Проте це може бути особливістю лише баскської мови. В інших ергативних мовах групу присудка може складати лише дієслово, тобто:

$VP \rightarrow V$

Виходячи з вищесказаного, побудована мова-еталон для мов (1) і (2) має такий перелік трансформаційних правил:

- | | |
|----------------------------|------------------------------|
| 1. $S \rightarrow NP + VP$ | 4. $VP \rightarrow N + VP$ |
| 2. $NP \rightarrow N$ | 5. $VP \rightarrow V + Vaux$ |
| 3. $VP \rightarrow V + N$ | 6. $VP \rightarrow V$ |

Отже, для побудови речень у мові (1) використовуються правила 1, 2 і 3, а у мові (2) застосовуються правила 1, 2, 4 і 5. Процес побудови речень являє собою послідовну заміну одних елементів іншими і матиме такий вигляд:

(1) $S \rightarrow NP + VP \rightarrow N + VP \rightarrow N + V + N$.

(2) $S \rightarrow NP + VP \rightarrow N + VP \rightarrow N + N + VP$.

Так, речення мови (1) будується у три кроки. На першому застосовуємо правило 1 та отримуємо загальну структуру речення ($S \rightarrow NP + VP$). На другому NP замінюємо на N і отримуємо $N + VP$. На третьому VP замінюємо на $V + N$ і, як результат, маємо синтаксичну структуру $N + V + N$. Аналогічним чином виглядає процес побудови речень мови (2). Таким чином в абстрактному вигляді будується структура речень як еталон для номінативної (1) та ергативної мов (2).

Звісно, будь-яка мова-еталон може вважатися не більш як гіпотеза про інваріантну основу мов світу, тому в інтересах типологічного дослідження необхідно використовувати по можливості найрізноманітніші гіпотези для того, щоб вибрати серед них найефективніші. Отже, структурна лінгвістика й типологія мов – це різні лінгвістичні дисципліни, але між ними існує тісний

зв'язок. Структурна лінгвістика, надаючи у розпорядження типології мов породжувальні граматиками як мови-еталони, слугує теоретичною базою типології мов. Типологія мов, використовуючи мови-еталони як одиниці для систематичного типологічного порівняння конкретних мов світу, слугує свого роду полем експериментальної перевірки ефективності породжувальних граматик, створюваних структурною лінгвістикою. Можна стверджувати, що типологія мов є однією з найважливіших галузей експериментальної перевірки результатів структурної лінгвістики. Синтез структурної лінгвістики та типології мов дає змогу розробляти системи синтаксичного та морфологічного аналізу, питально-відповідальні та інші системи аналізу та синтезу природних мов.

Контрольні запитання:

1. Які складнощі можуть виникнути під час реконструкції системи тієї чи іншої мови?
2. Чим зумовлена необхідність застосування методів структурної лінгвістики у порівняльно-історичних розвідках?
3. У чому полягає перевага використання породжувальних граматик як мови-еталону?
4. Які можуть існувати обмеження у використанні мов-еталонів (у вигляді породжувальних граматик) під час типологічних досліджень?

Практичні завдання

1. На основі матеріалу наукової статті¹⁶ відтворіть історичну систему відмінювання українських іменників, які зараз належать до другої відміни.

Друга відміна сучасної української мови включає переважно іменники чол. р. з нульовим закінченням (*гай, явір, плуг, друг*), із закінченням *-о* (тато, дядько, Павло), одиничні на *-е* (*вітрище*) та субстантиви сер. роду із флексіями *-о* (*озеро, вікно, відро*), *-е* (*серце, море, поле*) і *-я* (*колосся, обличчя, між гір'я*).

¹⁶<http://dspace.cuspu.edu.ua/jspui/bitstream/123456789/1784/1/Історичні%20коментарі%20відмінкових%20закінчень%20іменників%20української%20мови%20%28друга%20відміна%29.pdf>

Історично вона склалася шляхом об'єднання іменників зі старими основами на:

1) *-o-, *-jo-; 2) *-ŷ-; 3) *-ĩ-; 4) *-n-; 5) *-s-.

Розподіліть знайдені іменники по п'яти групах, що репрезентують відповідну основу, та подайте до них відповідний набір флексій:

Стара основа	Слова	Набір флексій

Опишіть історичну систему відмінювання іменників, використовуючи модель словозміни (див. підрозділ 5.3) та теорію системи мови Володимира Широкова, визначаючи зміст кожного складника системної тріади:

Структура:

Суб'єкт:

Субстанція:

2. В угорській мові, так само як і в українській, можна утворювати від одних слів інші різними способами: а) додаванням до кореня різних суфіксів або закінчень; б) з'єднанням декількох коренів у складне слово тощо. Нижче подано слова угорською мовою та їх переклад українською:

háború	війна	körről	про коло, про кільце
háborúról	про війну	bátor	сміливий
bűn	вина	bátorságról	про сміливість
bűntelen	безвинний	bátortalan	несміливий
bűnről	про вину	föld	земля, поле
út	дорога, шлях	földtelen	безземельний
útról	про дорогу, про шлях	sötét	темний
kör	коло, кільце	sötétség	темнота

Побудуйте модель утворення слів, яку можна застосувати як для угорської, так і української мов. Для цього потрібно:

– Проаналізувати морфологічну структуру угорських слів та порівняйте її зі структурою в українській мові.

– *Визначити правила використання тих чи інших суфіксів в угорській мові (суб'єкт).*

– *Визначити компоненти, які мають бути присутніми у морфологічній моделі (структура).*

– *Схарактеризувати реалізацію компонентів моделі в угорській та українській мовах (субстанція).*

6.2. Структурна лінгвістика і навчання мові

Прикладна лінгвістика – розділ мовознавчої науки, що вивчає проблеми застосування мови у галузях людської діяльності. Серед її галузей дослідження – навчання рідній та іноземній мовам, що передбачає процес оволодіння навичками усного і письмового мовлення, тобто вміння використовувати мову у різних комунікативних ситуаціях. Своєю чергою, вміння володіти мовою означає здатність будувати висловлювання з певним змістом, використовуючи відповідні граматичні та лексичні засоби. Тобто від вивчення та аналізу мови прикладна лінгвістика переходить до її синтезу, однією із ефективних методик якого може стати створення мов-еталонів, що, у свою чергу, є предметом вже структурної лінгвістики. За допомогою мови-еталону можна задати, зокрема, перехід від певної семантичної категорії до системи засобів вираження, наприклад: для категорії відмінка можна запропонувати таку послідовність «семантичний відмінок – граматичне оформлення – відмінкова система конкретної мови».

Спочатку з'ясуємо, що таке семантичний відмінок, і як він співвідноситься з граматичним відмінком. Сам термін «семантичний відмінок» узято з відмінкової граматики, основоположником якої є Чарльз Філмор. Сутність її полягає в тому, що план змісту речення репрезентується набором семантично-синтаксичних функцій (або ролей), склад якого задається лексичним значенням предикату (тобто дієслова). Кожному дієслову відповідає певний набір відмінків, і кожному з них відповідає певний учасник дії (ситуації) або **партиципant**. Наприклад, англійською та українською дієслово *to move* – *рухати* передбачає наявність щонайменше трьох партиципantів, які

виражаються іменниками у різних відмінках. В українській мові це може бути виражено називним, давальним та знахідним відмінками: *ХТО рухає ЩО і КУДИ*; а в англійській мові – структурою: *SOMEBODY moves SOMETHING SOMWHERE*. Цілком можливо розширити цей набір за рахунок необов’язкових учасників *КОЛИ, ЯКИМ ЧИНОМ і ЧИМ (WHEN, HOW, WITH SMTH / BY SMB)*. У семантиці обов’язкові учасники також називають **актантами**, а необов’язкові – **сирконстантами**. Отже, **семантичним відмінком** називають *роль предметного імені (або учасника) у структурі речення*.

Наведемо у таблиці приклади семантичних відмінків у тих чи інших реченнях.

Таблиця 6.1. Приклади семантичних відмінків у реченнях.

Речення	Кількість семантичних відмінків	Семантичні відмінки
Дитина спить	1	Суб’єкт стану
Кішка п’є молоко	2	1. Суб’єкт дії
		2. Об’єкт дії
Батько подарив синові іграшку	3	1. Суб’єкт дії
		2. Адресат
		3. Об’єкт дії
Хлопчик пише ручкою лист матері	4	1. Суб’єкт дії
		2. Інструмент
		3. Об’єкт дії
		4. Адресат

У своїй роботі «The case for case» Чарльз Філмор виокремлює такий набір семантичних відмінків:

- (А) – агентив: суб’єкт дії;
- (Д) – датив: особа, до якої стосується дія;
- (Т) – інструмент: сила або предмет, за допомогою якого здійснюється дія;
- (Г) – фактив: результат дії (особа, предмет тощо);
- (Л) – локатив: місце або простір, де здійснюється дія;
- (О) – об’єктив: власне об’єкт дії.

У теоретичній семантиці кількість семантичних відмінків розширилася до ста для відображення більшої кількості можливих семантичних відношень.

Наведемо приклад одного з розширень, що запропонував Чарльз Філмор у своїй граматиці:

- (1) експерієнсер – партиципant, на внутрішній стан якого впливає дія;
- (2) джерело – партиципant, що є джерелом, звідки відбувається дія;
- (3) реципієнт – партиципant, який набуває щось у результаті здійснення дії.
- (4) напрямок або мета – партиципant, що є місцем спрямування дії;
- (5) стимул – партиципant, що викликає почуття або емоції;
- (6) час – часовий проміжок, у який здійснюється дія;
- (7) природня сила – партиципant, що несвідомо виконує дію;
- (8) образ дії – тобто яким чином виконується дія;
- (9) мета – для чого здійснюється дія;
- (10) причина – чому відбулася дія.

Тепер покажемо, як можна здійснити перехід від семантичного відмінка до граматичних засобів його вираження на прикладі флективно-синтетичної (української) та аналітичної (англійської) мов. Треба звернути увагу, що семантичні відмінки визначають синтаксичну структуру речення:

Хлопчик пише ручкою лист(а) матері у кімнаті батька ввечері.

A boy writes a letter to his mother with a pen in his father's room in the evening.

Дієслово *писати* (англ. *to write*) у семантично-синтаксичній структурі речення передбачає трьох обов'язкових партиципantів (актантів) або семантичних відмінків: агентив (*ХТО пише*), об'єктив (*ЩО пише*) і реципієнт (*КОМУ пише*). У цьому прикладі наявні додаткові партиципанти (сирконстанти): інструмент (*ЧИМ пише*), локатив (*ДЕ пише*), посесив (*у ЧИЙ кімнаті пише*) і час (*КОЛИ пише*). Семантичні відмінки в аналізованих мовах позначають прийменниками, а також відповідними флексіями (як в українській) або порядком слів (як в англійській). Англійські іменники, що виконують семантико-синтаксичну роль в реченні, супроводжуються детермінантом, тобто артиклем (*a, the*) або вказівним займенником (*his*).

Таблиця 6.2. Семантичні відмінки та їхні ознаки в реченні.

Семантичний відмінок	Значення сем. відмінка		Формальна ознака	
	українська	англійська	українська	англійська
Агентив (А)	хлопчик	a boy	має Ø закінчення, стоїть перед дієсловом	стоїть перед дієсловом
Об'єктив (О)	лист(а)	a letter	має Ø закінчення, стоїть після дієслова	стоїть після дієслова
Реципієнт (R)	матері	his mother	має закінчення «-erі»	стоїть після об'єктива, має прийменник <i>to</i>
Інструмент (І)	ручкою	a pen	має закінчення «-ою»	має прийменник <i>with</i>
Локатив (L)	в кімнаті	room	має прийменник + закінчення «-і»	має прийменник <i>in</i>
Час (Т)	ввечері	the evening	має прийменник закінчення «-erі»	має прийменник <i>in</i>
*Посесив (Р)	батька	his fater's	закінчення «-а»	має закінч-я «'s»

*Посесив є окремим присвійним відмінком в англійській мові, але в українській мові він лише відображає додаткову семантичну характеристику будь-якого партиципанта.

Побудуємо семантико-синтаксичну структуру для двох аналізованих речень, яка включає семантичні відмінки, наведені в таблиці 6.2, та розпишемо їхні відповідники в українській та англійській мовах:

$$S \rightarrow A + V + O + R + I + L[P] + T$$

Уведемо позначення такого вигляду $N_i (UKR)$ та $N_i (ENG)$ на позначення українських та англійських граматичних форм, що відповідають елементам семантико-синтаксичної структури. Індекс i позначає номер того чи іншого

відмінка відповідно до загальноновживаної системи відмінювання. Так, для української мови застосовні такі правила підстановки:

$A \rightarrow N_1 (UKR)$ – називний відмінок;

$O \rightarrow N_4 (UKR)$ – знахідний відмінок;

$R \rightarrow N_3 (UKR)$ – давальний відмінок;

$I \rightarrow N_5 (UKR)$ – орудний відмінок;

$L \rightarrow PR (UKR) + N_6 (UKR)$ – місцевий відмінок;

$P \rightarrow N_2 (UKR)$ – родовий відмінок;

$T \rightarrow N_6 (UKR)$ – місцевий відмінок.

Нехай $N_1 (UKR) = \{x\}$ – множина усіх українських словоформ, які співвідносяться з називним відмінком. Своєю чергою, будь-яка словоформа x розкладається на основу $c(x)$ та флексію $f(x)$, тобто:

$$x = c(x) * f(x), \text{ де } f(x) \in F^i$$

Символ F^i позначає множину флексій, які належать до відмінка i . Наприклад, F^1 позначає множину флексій називного відмінка:

$$F^1 = \{f_1^1(x), f_2^1(x), f_3^1(x), f_4^1(x), f_5^1(x)\},$$

де $f_1^1(x) = \text{«-а»}$, $f_2^1(x) = \text{«-я»}$, $f_3^1(x) = \text{«-о»}$, $f_4^1(x) = \emptyset$, $f_5^1(x) = \text{«-е»}$ ¹⁷.

Аналогічно можна отримати флексії для інших відмінків української мови – родового (F^2), давального (F^3), знахідного (F^4), орудного (F^5), місцевого (F^6) та кличного (F^7). Але потрібно зважати, що той самий відмінок може використовуватися для відображення інших семантико-синтаксичних відношень. Наприклад, флексії знахідного відмінка можуть виражати значення як предмета або властивості (*мати волю*, *мати яблуко*, *мати дім*), так і міри (*працювати рік*, *годину*, *тиждень*). Варто зазначити, що один семантичний відмінок може передаватися різними відмінковими формами. Так, наприклад, відмінок агенс (або агентив) в українській мові може відповідати називному (*дитина спить*), родовому (*сміх дитини*), давальному (*дитині спиться*) та орудному (*малюнок намальовано дитиною*) відмінкам. Але в той самий час морфологічний відмінок не завжди дорівнює семантичному відмінку. Називний

¹⁷ Цей перелік не є повним, оскільки не включає відмінкові флексії для форм множини.

може називати ситуацію взагалі (*ранкова тиша*), мати значення ознаки (*Сашко – молодець!*), значення кількості (*гроші – кури не клюють*) тощо. Наведемо таблицю відповідності між семантичними та граматичними відмінками в українській мові:

Таблиця 6.3. Семантичні та граматичні відмінки в українській мові.

Семантичний відмінок	Граматичний відмінок	Приклад
Агентив	Називний	Дитин-а спить
	Родовий	Сміх дитин-и
	Давальний	Дитин-і сниться
Об'єктив	Знахідний	Бачу дім-Ø, бачу батьк-а
	Родовий	Не вистачає сміливост-і
	Місцевий	Говорити про батьк-а
Реципієнт	Давальний	пишу [лист] матер-і
Інструмент	Орудний	Пишу ручк-ою, працюю рук-ами
Локатив	Місцевий	Лежати на стол-і
	Родовий	Стояти коло хат-и
	Орудний	Літати над мор-ем, йти ліс-ом
Час	Родовий	Робити протягом годин-и
	Знахідний	Упоратися за годин-у
	Орудний	Працювати годин-ами
	Місцевий	Працювати з ранк-у

Останнім етапом є угруповання слів української мови по відмінах, залежно від набору флексій для їх відмінювання. Перейдемо до розгляду семантично-синтаксичної структури англійського речення. Як було зазначено вище, синтаксична роль іменника в англійській мові визначається використанням прийменників та (або) порядком слів у реченні. Виходячи з таблиці 6.2, семантичні відмінки виражаються у такий спосіб:

$$A \rightarrow N_A (ENG) \rightarrow DET + N;$$

$$O \rightarrow N_O (ENG) \rightarrow DET + N;$$

$$R \rightarrow N_R (ENG) \rightarrow PR + DET + N \rightarrow to + DET + N;$$

$$I \rightarrow N_I (ENG) \rightarrow PR + DET + N \rightarrow with + DET + N;$$

$$L \rightarrow N_L (ENG) \rightarrow PR + DET + N \rightarrow in + DET + N;$$

$$P \rightarrow N_P (ENG) \rightarrow PR + DET + N \rightarrow DET + N.$$

$$T \rightarrow N_T(ENG) \rightarrow PR + DET + N$$

На відміну від української, в англійській мові семантичні відмінки передаються синтаксичними (аналітичними) засобами, тобто за допомогою прийменників, а їх ідентифікація визначається їхнім місцем у реченні. Таким чином отримуємо модель системи відмінювання для певної мови-еталону, яка в абстрактному та узагальненому вигляді описує перехід від семантичних відмінків до граматичного засобу їх оформлення, а далі – до системи відмінювання у відповідній мові-еталоні. Така модель може бути основою для посібників та комп'ютерних систем навчання рідній або іноземній мовам, а також для створення автоматичних систем перекладу.

Контрольні запитання:

1. Схарактеризуйте роль мов-еталонів у теоретичному та практичному мовознавстві.
2. Які сучасні мови зберегли систему відмінків, що були наявні у латинській мові?

Практичні завдання

1. Визначте усі можливі семантичні відмінки (партиципанти), яких потребують наведені нижче англійські дієслова. Як виявлені відмінки можна передати засобами української мови?

To bring, to buy, to create, to detect, to destroy, to enjoy, to feel, to love, to paint, to receive, to take.

Примітка: у відмінковій граматиці Чарльз Філмор виокремлює такі семантичні відмінки: агент, експерієнсер, інструмент, джерело, напрямок, локатив, час, стимул, об'єктив, природня сила, реципієнт, спосіб дії, мета, причина. Стислу довідку про семантичні відмінки можна знайти за посиланням: <https://cuckoo.js.ila.titech.ac.jp/~yamagen/ling/ling-fillmore-single2021.pdf>

Під час виконання цього завдання можна також використовувати ресурс, що надає приклади вживання дієслів – <https://skell.sketchengine.eu/#home?lang=en>

Дієслово	Семантичні відмінки імен

Якими граматичними засобами відображаються виявлені семантичні відмінки:

Семантичний відмінок	Граматичні засоби

2. Визначте, якими граматичними засобами відображаються семантичні відмінки в іспанському (а), німецькому (б) або французькому (в) варіантах українського речення: Хлопчик пише ручкою лист(а) матері у кімнаті батька ввечері.

(а) Un niño escribe una carta a su madre con un bolígrafo en la habitación de su padre.

(б) Ein Junge schreibt abends im Zimmer seines Vaters mit einem Stift einen Brief an seine Mutter.

(в) Le soir un garçon écrit une lettre à sa mère avec un stylo dans la chambre de son père.

Заповніть таблицю:

Семантичний відмінок	Формальна ознака	
	українська	іспанська / німецька / французька
Агентив (А)		
Об'єктив (О)		
Реципієнт (R)		
Інструмент (І)		
Локатив (L)		
Час (Т)		
*Посесив (Р)		

3. Якими граматичними відмінками відображаються подані нижче семантичні відмінки у знайомій вам мові?

Семантичний відмінок	Граматичний відмінок	Приклад
Експерієнсер		

Джерело		
Реципієнт		
Напрямок або мета		
Стимул		
Образ дії		
Мета		
Причина		

6.3. Структурна лінгвістика і процес перекладу

Розглянемо роль структурної лінгвістики у процесі перекладу. На сьогодні існує велика кількість моделей опису процесу перекладу, але найбільш відомими є семантична, ситуативна, трансформаційна, психолінгвістична, комунікативна та інформативна. Як приклад, розглянемо трансформаційну модель, яка має в основі теорію породжувальних граматик Ноама Хомського (див. теми 10 і 11). Сутність цієї граматики полягає у спробі сформулювати правила синтаксичних перетворень, за допомогою яких можна отримати багатоміття висловлювань з використанням ядерних, тобто простих синтаксичних структур. Отримані з ядерних структур висловлювання називаються трансформами. Ядерні структури і трансформи мають спільний елемент – план змісту. За рахунок правил перетворень та обмеженого інвентаря елементів (звуків, слів), носій мови за теорією Ноама Хомського здатний по-різному оформлювати цей план змісту і сприймати усе розмаїття зворотів рідною мовою. За трансформаційною моделлю весь процес перекладу відбувається в три етапи:

- 1) етап аналізу – розкладання синтаксичної структури вихідною мовою на мінімальні ядерні структури;
- 2) етап переходу від ядерних структур до загальної інваріантної структури, яка відображає план змісту висловлювання вихідною мовою;
- 3) етап трансформації загальної структури у синтаксичну структуру висловлювання мовою перекладу.

Продемонструємо роботу трансформаційної моделі перекладу на прикладі такого речення: *However, with the definition of Covid deaths narrowed and mass*

testing no longer compulsory, government data is no longer reflective of the true scale of the outbreak. На першому етапі розкладемо вихідне речення на окремі ядерні структури:

- 1) *the definition of Covid deaths has been narrowed;*
- 2) *mass testing is no longer compulsory;*
- 3) *government data does not reflect any longer the true scale of the outbreak;*
- 4) *government data about the true scale of the outbreak is no longer reflective;*
- 5) *government data about the scope of the outbreak does no longer reflect the truth.*

Після того, як речення розкладено на більш прості синтаксичні структури, другим етапом є утворення загальної структури, що відображає план змісту речення мовою оригіналу. Використавши певні трансформації, можна отримати такі структури:

(1) *However, government data does not reflect any longer the true scale of the outbreak since the definition of Covid deaths has been narrowed and mass testing is no longer compulsory;*

(2) *However, government data about the true scale of the outbreak is no longer reflective, since the definition of Covid deaths has been narrowed and mass testing is no longer compulsory;*

(3) *However, government data about the true scope of the outbreak is no longer reflective, because of the Covid deaths definition narrowed and mass testing no longer compulsory;*

(4) *However, government data about the scope of the outbreak does no longer reflect the truth, because of the Covid deaths definition narrowed and mass testing no longer compulsory.*

Отже, шляхом відповідних перетворень і заміщень (1) – (4) можна отримати загальну синтаксичну структуру, яка передається засобами мови перекладу: *Однак урядові дані про масштаби спалаху вже не відповідають дійсності через те, що показники смертності від Ковіду були занижені, а масове тестування більше не є обов'язковим. Як недолік трансформаційної*

моделі перекладу можна відзначити відсутність пояснення того, як здійснюється вибір кінцевої структури або одиниці на другому та третьому етапах. Тому, для повного відтворення процесу перекладу слід використовувати також інші моделі. Однак така модель демонструє у спрощеному вигляді процес аналізу синтаксичної структури вихідного речення перекладачем, визначення змісту на основі цього аналізу та побудови речення мовою перекладу з використанням відповідних синтаксичних конструкцій.

Контрольні запитання:

1. Чим відрізняється процес породження з погляду теорії генеративних граматик від процесу породження у сфері перекладу?
2. Які сильні та слабкі сторони має трансформаційна модель процесу перекладу?

Практичні завдання

1. Опишіть процес перекладу поданих нижче речень з використанням трансформаційної моделі, як це було показано вище:

(1) Nothing can replace the ability to connect, observe, and create with peers that comes from being physically together, nor the opportunity to grow professionally by learning from leaders and mentors.

(2) Despite the actions of the protesters, in the hours before the chaos, security had appeared tight, with the roads closed for about a block around the parliament area and armed police pairs guarding every entrance into the area.

2. Як можна граматично трансформувати наведені вище речення таким чином, щоб їхній зміст залишався без змін? Наведіть приклади.

МОДУЛЬ 3 ЛІНГВІСТИЧНІ МОДЕЛІ

ТЕМА 7 МОДЕЛІ В ЛІНГВІСТИЧНИХ ДОСЛІДЖЕННЯХ

План заняття:

1. Поняття *моделювання* і *моделі*
2. Гіпотетико-дедуктивний метод
3. Моделювання в лінгвістиці

Основним методом структурної лінгвістики є метод моделювання. Його застосовують для відображення інтелектуальних процесів, пов'язаних із використанням мови, а також репрезентації різних мовних явищ, таких як словотвір, словозміна, синтаксис. Такі моделі мають як теоретичне – створення нових лінгвістичних теорій, – так і практичне значення, тобто на їх основі розробляють комп'ютерні системи опрацювання природномовної інформації.

7.1. Поняття моделювання і моделі

У минулих темах ми неодноразово використовували термін «модель». Він пов'язаний з іншим – «моделювання», що позначає метод наукового пізнання, який полягає у створенні та дослідженні моделей. Якщо точніше, моделювання – це зображення об'єкта у вигляді моделі для отримання інформації про цей об'єкт шляхом проведення над нею експериментів. Будь-яке моделювання полягає в тому, що для отримання нових знань про реальний об'єкт (його найважливіші властивості, характеристики, роботу) його необхідно замінити об'єктом-моделлю або просто моделлю. Існує навіть окрема теорія – теорія моделювання, яка висвітлює питання заміщення реальних об'єктів на моделі та дослідження властивостей об'єктів на прикладі цих моделей. Отже, **моделювання** можна визначити як *метод опосередкованого пізнання, за якого досліджуваний об'єкт-оригінал перебуває в певній відповідності з іншим об'єктом-моделлю, причому модель здатна в тому чи іншому відношенні замінювати оригінал на деяких стадіях пізнавального процесу.*

Що означає словосполучення «опосередковане пізнання», яке було використано у дефініції? Це означає, що ми пізнаємо об'єкт, явище, процес або дійсність не безпосередньо, а через інший предмет. Одним із прикладів

опосередкованого пізнання є дослідження процесу перекладу у вигляді мисленнєвих операцій над мовними одиницями речення або тексту. У цьому випадку безпосереднє пізнання неможливе, тобто ми не можемо напряду бачити, які нейрони активуються та як вони працюють під час перекладу. Тому єдиним можливим рішенням стає використання певного об'єкту, що замінює так званий «інтелектуальний механізм». І таким об'єктом є модель, що відображає послідовність дій, за допомогою яких можна розв'язати певне перекладацьке завдання за заданих умов процесу перекладу.

Моделювання може бути різним: так, за галузями людської діяльності виокремлюють математичне, біологічне, хімічне, соціологічне, економічне, лінгвістичне та інші. У свою чергу, за способом реалізації розрізняють матеріальне та ідеальне (рис. 7.1).



Рис. 7.1. Види моделювання.

Матеріальним зазвичай називають моделювання, під час якого реальному об'єкту протиставляється його збільшена або зменшена копія, яка дає змогу виконувати дослідження в лабораторних умовах з подальшим перерахуванням характеристик процесів або явищ з моделі на об'єкт на основі теорії подібності. Прикладами такого моделювання є побудова макетів будинків в архітектурі, моделей літальних апаратів у авіабудуванні тощо. **Ідеальне** моделювання засновано не на матеріальній аналогії об'єкта й моделі, а на аналогії ідеальної, уявної. Мисленнєвий експеримент вперше реалізував Галілео Галілей, завдяки чому учений прийшов до ідеї інерційного руху тіла. **Знакове моделювання** – вид моделювання, що використовує у вигляді моделі певні знакові перетворення: схеми, графіки, малюнки. **Математичне моделювання** – це

моделювання, під час якого дослідження об'єкта відбувається за допомогою моделі, побудованої мовою математики (наприклад, дослідження законів механіки Ньютоном за допомогою математичних формул).

Перейдемо тепер до терміна **модель**, для якого існує чимало визначень, оскільки використовується в різних сферах. Навіть у лінгвістиці зазначений термін не матиме єдиного визначення (пор. модель речення, модель словотвору, модель словозміни, модель породжувальної граматики, модель процесу перекладу, модель «Смисл – Текст» тощо). Якщо проаналізувати більшість визначень, то можна знайти різні риси поняття, позначеного терміном «модель» (див. Таблицю 7.1), серед яких можна виокремити такі: 1) слово, через яке визначається поняття; 2) призначення, тобто з якою метою створено модель; 3) додаткові характеристики, тобто як можуть проявляти себе моделі, залежно від свого призначення.

Таблиця 7.1. Характеристика поняття моделі.

Галузь	Визначальне слово	Призначення	Додаткові характеристики
Наука і техніка	Спрощений об'єкт	Дослідження властивостей реальних об'єктів	Містить важливі властивості реального об'єкта
Інформатика	Система	Отримання інформації про іншу систему	Спосіб уявлення реального процесу, пристрою або концепції
Математика	Система	Прогнозування поведінки реального об'єкта	Являє собою той чи іншого ступінь ідеалізації об'єкта
Фізика	Технічний пристрій	Дослідження характеристик або властивостей	Імітує певні характеристики або властивості матеріальних об'єктів або фізичних явищ
Лінгвістика (словотвірна модель)	Шаблон, взірць	Утворення похідних слів	Схема побудови слів певної частини мови

Якщо зробити узагальнення, то можна виокремити спільні для наведених галузей (у тому числі для структурної лінгвістики) риси поняття **модель**:

1) репрезентує певний реальний об'єкт: предмет, механізм, систему, явища або процеси;

2) імітує реальний об'єкт: його форму, будову, принципи функціонування тощо;

3) має прогностичний характер: прогнозує можливу поведінку об'єкта у майбутньому, результати функціонування або можливі майбутні характеристики реального об'єкта.

Кожен об'єкт має велику кількість різних властивостей, тому при створенні його моделі беруть до уваги лише найсуттєвіші. Наприклад, модель літака повинна мати геометричну подібність до оригіналу, модель атома – правильно відображати фізичні взаємодії, архітектурний макет міста – ландшафт тощо. Ознака або величина, які характеризують будь-яку властивість об'єкта і можуть набувати різних значень, називаються параметрами моделі. Модель відображає в заздалегідь обумовленому вигляді будову та властивості реального об'єкта, явища або процесу, які є суттєвими з точки зору дослідників для проведення того чи іншого експерименту. Підсумовуючи вищесказане, можна подати таке визначення: **модель** – *матеріальний або уявний замінник реального об'єкта, що містить у собі його найсуттєвіші характеристики або властивості, на основі якого отримують нові знання про цей реальний об'єкт і формують прогнози щодо його майбутньої поведінки.*

Розглянемо питання класифікації моделей. Існує чимало критеріїв, за якими можна розподілити моделі по різних типах. Так, наприклад, їх можна класифікувати за галуззю використання (навчальні, експериментальні, науково-технічні, ігрові, імітаційні), за роллю часового чинника (статичні та динамічні), за наявністю впливу на систему (детерміновані й стохастичні) та за способом репрезентації (матеріальні та інформаційні). Найширшою є остання класифікація – за способом репрезентації (див. рис. 7.2).

Матеріальні моделі відтворюють геометричні та фізичні властивості оригіналу та завжди мають реальне втілення, тобто реалізують матеріальний підхід до вивчення об'єкта, явища або процесу. Вони так називаються тому, що

їх можна побачити, почути, доторкнутися до них. Прикладами таких моделей є: 1) дитячі іграшки, за допомогою яких дитина пізнає світ; 2) шкільні посібники, а також фізичні та хімічні досліди (так, під час хімічних дослідів моделюється реакція кисню з воднем, тобто утворення небезпечної «гримучої суміші» з простих природних елементів); 3) мапи для вивчення історії або географії, схеми сонячної системи тощо.

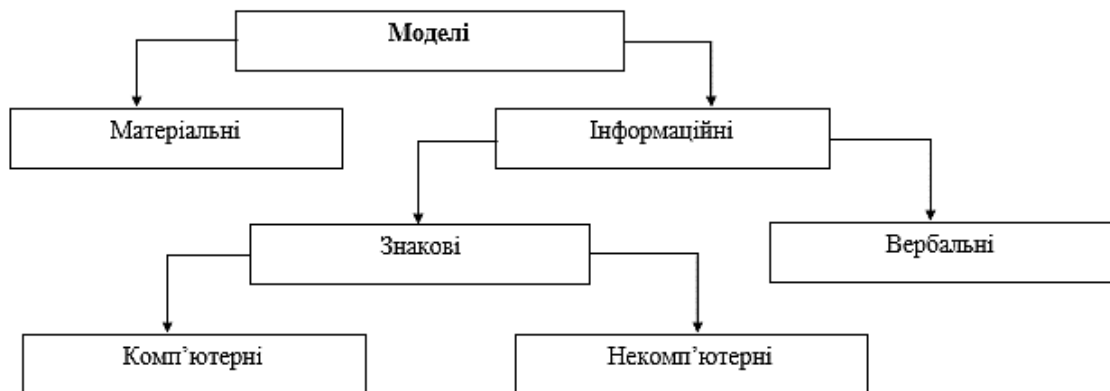


Рис. 7.2. Типи моделей за призначенням.

Інформаційні моделі являють собою сукупність інформації щодо стану та властивості об'єкта, процесу або явища. Такі моделі називаються так, тому що їх неможна побачити або доторкнутися до них, і вони будуються лише на основі інформації, що характеризує об'єкт або процес. Форма, спосіб і засоби подання такої інформації можуть бути різними. До інформаційних належать знакові та вербальні моделі. **Вербальна** – модель, що репрезентується у мисленнєвій або розмовній формі за допомогою словесного опису у вигляді тексту. Інформаційні моделі формуються в результаті роздумів та умовиводів і можуть залишатися мисленнєвими або вираженими словесно. **Знаковою моделлю** називають модель, що репрезентована за допомогою спеціальних знаків, тобто засобами будь-якої формальної мови. Прикладами таких моделей можуть бути малюнки, тексти, схеми, графіки тощо. Залежно від засобу реалізації, знакові моделі поділяються на комп'ютерні та некомп'ютерні. **Комп'ютерна модель** реалізується засобами програмного середовища.

За формою репрезентації, інформаційні моделі в цілому поділяються на:

(а) геометричні – графічні форми або об'ємні конструкції;

- (б) словесні – усний або письмовий опис з ілюстраціями;
- (в) математичні – математичні формули, що відображають зв'язок різних параметрів об'єкта або процесу;
- (г) структурні – схеми, графіки, таблиці тощо;
- ґ) логічні – моделі, що репрезентують різні варіанти вибору дій на основі умовиводів і аналізу умов;
- д) спеціальні – ноти, хімічні формули тощо.

Контрольні запитання:

1. Яке теоретичне та практичне значення мають моделі у лінгвістичній галузі?
2. Які властивості мови, на ваш погляд, відображає модель породжувальної граматики та який реальний об'єкт вона заміщує?
3. У вигляді чого можна реалізувати модель системи мови у системній тріаді «структура – суб'єкт – субстанція»?

Практичні завдання

1. Подано знакові моделі разом із їхньою інтерпретацією українською мовою:

$A \text{ } 1 \text{ } B - A \text{ з'їло } B.$

$C : (A \text{ } 1 \text{ } B) - C \text{ зробило так, щоб } A \text{ з'їло } B.$

Нехай тепер 2 означає «побачити», а 3 – «мати». Побудуйте аналогічні моделі для передання змісту наведених нижче фраз, використовуючи символи B (ворона), L (лисиця) і C (сир) на позначення відповідних іменників:

- | | |
|--|--------------------------------|
| 1. Був у ворони сир. | 4. Ворона показала сир лисиці. |
| 2. Помітила сир лисиця. | 5. Лисиця відібрала сир. |
| 3. Лисиця постаралася попастися на очі вороні. | 6. Ворона побачила лисицю. |

Проінтерпретуйте зашифрований зміст, що містять подані нижче знакові моделі:

1. $L : (L \text{ } 3 \text{ } C)$

3. $L \text{ } 1 \text{ } C$

2. $L : (B \text{ } 2 \text{ } C)$

4. $B : (L \text{ } 1 \text{ } C)$

2. Подано такі речення, що містить ті чи інші логічні твердження:

(1) Усі малі діти нерозумні.

(2) Усі, хто приборкує крокодилів, заслуговують на повагу.

(3) Усі нерозумні люди не заслуговують на повагу.

Нехай символами позначено: (C) – «малі діти», (S) – «розумні люди», (T) – «ті, хто приборкує крокодилів», (R) – «ті, хто заслуговують на повагу». Уведемо символ « $\bar{}$ », що заперечує твердження: $\underline{C}$ – «не малі діти», $\underline{S}$ – «нерозумні люди», $\underline{T}$ – «ті, хто не приборкує крокодилів» і $\underline{R}$ – «ті, хто не заслуговує на повагу». Символ « $\subseteq$ » позначає відношення включення, тоді, наприклад, запис $C \subseteq \underline{S}$ відповідає реченню (1).

Проінтерпретуйте зміст таких формальних записів:

Запис	Інтерпретація
$S \subseteq \underline{C}$	
$\underline{R} \subseteq \underline{T}$	
$R \subseteq \underline{S}$	

Визначте зміст групи формальних записів:

$(C \subseteq \underline{S}, \underline{S} \subseteq \underline{R}) \Rightarrow C \subseteq \underline{R} \quad (T \subseteq R, R \subseteq S) \Rightarrow T \subseteq S \quad (\underline{S} \subseteq \underline{R}, \underline{R} \subseteq \underline{T}) \Rightarrow \underline{S} \subseteq \underline{T} \quad (R \subseteq S, S \subseteq \underline{C}) \Rightarrow R \subseteq \underline{C} \} \Rightarrow C \subseteq \underline{T}; T \subseteq \underline{C}.$

Запис	Інтерпретація

7.2. Гіпотетико-дедуктивний метод

У назві цього методу наявні два ключових терміни «гіпотеза» і «дедукція», що визначають його сутність. Будь-яке дослідження не починається з побудови моделей та проведення відповідних експериментів. Будь-яка модель є відображенням певної теорії, у тому числі лінгвістичної, яка була апробована шляхом багатьох спостережень, численних експериментів та перевірок. Щодо цього Лев Щерба зазначив, що лінгвістичне дослідження має включати збір фактичного мовного матеріалу, побудову гіпотези й перевірку її на нових фактах. Отже, для того, щоб побудувати ту чи іншу модель, необхідно виконати

попередній етап – сформулювати гіпотезу на основі спостережень за тими чи іншими явищами. Наступним етапом стає апробація шляхом експериментів, яка має на меті перевірити істинність гіпотези не лише на спостережуваних, а й на інших додаткових фактах. Результатом апробації гіпотези є нова теорія або повне чи часткове спрощення сформульованої гіпотези. Заключним етапом стає формування нової наукової теорії. Вищесказане відображає саму сутність гіпотетико-дедуктивного методу (рис. 7.3). Наведемо визначення: **гіпотетико-дедуктивний метод** – це особливий метод науково-дослідної роботи, що полягає у виробленні припущень (гіпотез) та їх перевірка емпіричним шляхом, тобто шляхом спостережень, перевірок і експериментів. Іншими словами це теоретичний метод наукового пізнання, що ґрунтується на дедуктивному виведенні висновків із гіпотези (або системи гіпотез) та їх емпіричній (експериментальній) перевірці.



Рис. 7.3. Гіпотетико-дедуктивний метод.

Реалізація гіпотетико-дедуктивного методу складається з таких етапів: 1) збір фактичного матеріалу; 2) узагальнення індуктивним шляхом, тобто формування теорії у вигляді гіпотези; 3) дедукція, тобто виведення нових ідей, що передбачає перевірку попередньо сформованої гіпотези. Розглянемо детально перелічені етапи.

Етап 1: збір фактичного матеріалу. Встановивши критерії об'єкта, що підлягає дослідженню, дослідник збирає фактичний матеріал. Перш за все, збір та аналіз матеріалу дослідження має відповідати меті та завданню дослідження. Фактичний матеріал має відображати сутність аналізованої проблеми та підтверджувати сформульовані гіпотези. Наприклад, основними джерелами лінгвістичного матеріалу є:

- *лексикографічні джерела*: словники різного типу, у тому числі загальнономовні та спеціальні;
- *нормативна документація*: галузеві стандарти, енциклопедії, довідники та підручники;
- *аутентичні тексти* (корпуси текстів різного типу), у яких наочним чином відображається реалізація одиницями мови своїх характеристик (структурних, семантичних).

Зібраний фактичний матеріал узагальнюється індуктивним шляхом. **Індуктивним** називається *висновок, у якому думка розвивається від знання одиничного до знання загального*. Саме так формуються гіпотези, з яких складається теорія. Теорією називатимемо деяку систему ідей, що дає наукове пояснення явищу чи класу явищ. Далі певна гіпотеза уточнюється багатьма шляхами, наприклад, зіставленням з відомими концепціями і виявленням суперечностей. Різницю між дедукцією та індукцією наочно наведено на рисунку 7.4 (гіпотеза у цьому прикладі може мати вигляд: «передбачається, що всі люди, як спостережувані, так і неспостережувані, мають одну голову»).

Етап 2: формулювання гіпотези. На основі зібраного матеріалу вчений може зробити попередні висновки або припущення у вигляді **гіпотези** тобто *твердження, що передбачає доказ, на відміну від аксіом та постулатів*. Іншими словами, вона є науковим припущенням, істинність якого невизначено. Формулюючи гіпотезу, дослідник будує припущення про те, яким чином має намір досягти поставленої мети. До гіпотез можна зарахувати окремі наукові твердження (припущення про етимологію конкретного слова) або цілі складні науково-теоретичні комплекси, такі як гіпотеза Сепіра-Уорфа.



Рис. 7.4. Різниця між дедукцією та індукцією.

Наведемо деякі приклади лінгвістичних гіпотез. Так, спостерігаючи за тим, як людина у дитячому віці за доволі стислий час опановує мову, Ноам Хомський зробив припущення, що в мозку людині наявний «лінгвістичний механізм», який власне відповідає за опанування мови і який відсутній у інших живих істот – у тварин і птахів. Такий механізм учений назвав породжувальною граматикою, якій він присвятив низку робіт. Іншим прикладом є гіпотеза Сепіра – Уорфа про те, що структура мови впливає на світосприйняття і бачення його носіїв, а також на їхні когнітивні (пізнавальні) процеси.

Гіпотеза вважається науковою, якщо вона, відповідно до наукового методу:

- пояснює факти, що охоплюються цією гіпотезою;
- не є логічно суперечливою;
- принципово спростовується, тобто потенційно може бути перевірена критичним експериментом;
- не суперечить раніше встановленим законам.

Окремим питанням є, яким саме чином потрібно формувати гіпотези. Щодо цього не існує певних правил, проте гіпотеза має бути сформульована таким чином, щоб її зміст (сутність) була зрозумілою.

Етап 3: дедукція, тобто виведення нових ідей, тверджень, нових теорій.

На цьому заключному етапі використовують **дедукцію**, що означає *спосіб міркування, за яким нове положення виводиться суто логічним шляхом від загальних положень до часткових висновків* (на противагу індукції). Основу будь-якої дедукції складає окремий умовивід, який впливає із загальних передумов. Метою дедукції є підтвердження факту істинності висунутої гіпотези та формування нових ідей, тверджень або теорій.

Наведемо простий приклад формулювання гіпотези під час використання дедуктивного методу. У підрозділі 6.1 мова йшла про мови-еталони для опису типологічних властивостей мов. Розглянемо побудову речення в ергативних мовах. Так, на початковому етапі у нас є речення «Кіт піймав мишу» баскською (1), грузинською (2) та азербайджанською мовами (3):

(1) *Katuak* *sagua* *harrapatu* *zuen*.

(2) კატუაკი საგა ჰარპათუ ჯუენი [k'at'am tagvi daich'ira].

(3) *Pişik* *siçanı* *tutdu*.

Експериментальним шляхом (за допомогою системи Google Translate) було встановлено, що першим йде іменник на позначення агенса (*KIT*), другим – іменник на позначення пацієнса (*МИША*), а третім – група присудка на позначення дії, що реалізує агенс (*СПИЙМАВ*). Групу присудка може складати дієслово або граматична конструкція з допоміжним дієсловом (у першому реченні *zuen* є допоміжним дієсловом) Перший іменник на позначення агенса стоїть в ергативному відмінку (який вказує на джерело спрямованості дії), а другий – в основному відмінку (відповідає знахідному відмінку в українській мові). У перших двох реченнях ергатив позначається певним закінченням (*-ak* в баскській або *-am* в грузинській), а в третьому – лише місцем в реченні. Отже, гіпотезу щодо структури речення в ергативних мовах ми можемо подати у вигляді моделі:

$$S \rightarrow N_x + N + VP$$

де, N_x – іменник в ергативному відмінку, x – наявність морфологічного показника ергативного відмінка, N – іменник в абсолютному відмінку, VP – група присудка, яка може набувати такого вигляду $VP \rightarrow V + V_{aux}$ або $VP \rightarrow V$.

Далі гіпотеза перевіряється на фактах інших мов ергативного типу. Цілком припустимо, що запропонована модель може зазнавати виправлень під час цієї перевірки. Якщо емпіричні перевірки дають негативний результат, то це свідчить про некоректність вихідної гіпотези або внесення відповідних змін до процедур перевірок та засобів логічного формулювання висновків.

Контрольні запитання:

1. Схарактеризуйте своїми словами сутність гіпотетико-дедуктивного методу.
2. Яку роль відіграє гіпотетико-дедуктивний метод у структурній лінгвістиці?
3. Як відбувається формулювання гіпотези і що саме для цього потрібно під час проведення наукових досліджень?

Практичні завдання

1. Німецьке слово **Die Losung** має жіночий рід. В українській мові слово **лозунг** сходить до цього німецького слова. Сформулюйте гіпотезу, яка пояснює чому слово лозунг стало словом чоловічого роду. Для цього необхідно визначити:

- ознаки жіночого роду в німецькій мові;
- ознаки, за якими слова зараховують до граматичного роду в українській мові;
- які слова іноземного (необов'язково німецькою) походження в українській мові є виключенням.

2. У наведеному нижче списку все слова підкоряються певному правилу, за яким в однокладових словах у родовому відмінку наголос може падати або на основу, або на закінчення:

Блат, болт, гриб, кіт, м'яч, полон, піп, піт, руль, світ, спирт, хміль, джміль, шторм, шум, щит.

Сформулюйте це правило у вигляді гіпотези. Для цього рекомендується розподілити слова на два класи залежно від наголосу (світ – світу; щит – щита́). Знайдіть те спільне, що об'єднує семантичні групи в межах кожного класу і протиставляє самі ці класи один одному.

На які з цих ознак може спиратися правило: довжина слова, характер кінцевої приголосної, характер кореневої голосної, наявність тих чи інших суфіксів або префіксів, походження слова (питоме чи запозичене), наявність того чи іншого стилістичного забарвлення, лексичне значення, лексико-граматичні особливості.

3. У поданих нижче відіменникових прикметниках наголос падає на:

- основу: фондовий, карликовий, маковий;*
- суфікс: парниковий;*
- закінчення: боковий, голосовий, жировий.*

Для цього вам потрібно провідмінювати іменники, від яких походить той чи інший іменник, та визначити куди падає у них наголос під час відмінювання.

Сформулюйте правило у вигляді гіпотези та перевірте, чи працює воно на інших словах: <https://uk.makeword.info/end/овий>

7.3. Моделювання в лінгвістиці

Як вже було зазначено у підрозділі 7.1, моделювання є методом наукового пізнання, який передбачає вивчення певних явищ, процесів і станів за допомогою моделей. Моделі створюються на ґрунті гіпотетико-дедуктивного методу і застосовуються як заміник «оригіналу». У цьому зв'язку, виникає питання, що саме слід моделювати в лінгвістиці? З огляду на це Володимир Широков зазначає, що об'єктом моделювання мають бути психофізичні стани і процеси, які відбуваються у мовно-розумовому апараті людини¹⁸. Іншими словами, усі інтелектуальні процеси, що пов'язані з використанням мови, мають братися за об'єкт моделювання. Так, наприклад: Ноам Хомський змодельовав процес породження граматично коректних речень відповідно до заданих правил; Ігор Мельчук розробив модель, яка відображає процес

¹⁸ В. А. Широков. Комп'ютерна лексикографія. Київ: «Наукова думка», 2011. С. 65.

побудови текстів із заданим змістом; Володимир Широков запропонував теоретичний апарат для моделювання психофізичних процесів у низці своїх робіт¹⁹.

Чому саме ці процеси потрібно подавати у вигляді моделі? Це пояснюється тим, що вони не повною мірою доступні для безпосереднього спостереження. Результати цих процесів можна спостерігати в усному та писемному мовленні. Володимир Широков також зазначає, що мова не виникає у голосових зв'язках, оскільки її походження є значно глибшим. Неможливість спостереження за самими процесами пояснюється ефектом «чорної скриньки», під яким розуміють прихований механізм, принцип роботи якого невідомий. Заздалегідь невідомо, як він виглядає і яким чином він породжує об'єкти (наприклад: речення або словозмінну парадигму). Про нього можна судити за результатами його діяльності, тобто за характером або властивостями об'єктів, що породжує цей механізм. Для дослідника важливими є ті аспекти функціонування механізму, які пов'язані з породженням спостережуваної сукупності об'єктів. Інші особливості функціонування механізму (його матеріал, складники) є несуттєвими.

Аналізуючи сукупність об'єктів, породжену механізмом, цілком можливо створити гіпотетичний опис цього механізму. Для того, щоб перевірити опис, необхідно побудувати на його основі модель механізму. Це буде модель, а не копія механізму, оскільки не буде враховано багато конкретних властивостей, а в деяких відношеннях модель зовсім не буде схожою на сам механізм. Але якщо ця модель під час свого функціонування буде породжувати точно такі самі об'єкти, що й досліджуваний механізм, то можна вважати, що розроблена модель є адекватною, тобто відповідною реальному механізму. Отже, для перевірки гіпотетичного опису прихованого від спостереження механізму потрібно, щоб відповідна цьому опису модель породжувала точно такі об'єкти, що й досліджуваний механізм. Створення моделі ґрунтується на гіпотезі про

¹⁹ Див.: Лінгвістично-інформаційні студії: праці Українського мовно-інформаційного фонду НАН України: у 5 т. / В. А. Широков та ін. Т. 1: Наукова парадигма та основні мовно-інформаційні структури. Київ. Український мовно-інформаційний фонд НАН України. 2018.

можливу будову оригіналу і часто являє собою функціональний аналог оригіналу, що дає змогу переносити отримані знання з моделі на оригінал. Критерієм адекватності моделі є експеримент.

Поняття лінгвістичної моделі активно використовується в комп'ютерній лінгвістиці, тому що без побудови лінгвістичних моделей неможливо розв'язати багато практичних завдань створення і використання лінгвістичних ресурсів (наприклад, машинних перекладачів, електронних словників, навчальних програм) та завдання автоматичного опрацювання мови. Наприклад, модель системи мови, у межах якої створюється словозміна для будь-якої мовної одиниці, можна подати у вигляді математичного запису такого вигляду:

$$V(PAR) \equiv \{\Lambda(PAR), P(PAR), H^{PAR}, H^{LEM}\}$$

Теоретичним фундаментом побудови наведеної моделі є інформаційна теорія лексикографічних систем провідного українського академіка Володимира Широкова. Термін **лексикографічна система** запропоновано і введено ученим на позначення способу формальної репрезентації будь-якої інформаційної системи, принцип функціонування якої схожий на природну мову. Цей принцип функціонування полягає у породженні одного чи кількох мовних явищ, таких як словотвір, словозміна, побудова речень (синтаксис) тощо. У наведеній моделі символами позначено:

- 1) $V(PAR)$ – систему мови, що містить множину всіх слів;
- 2) $\Lambda(PAR)$ – множину слів (x_0) у вихідній формі (в українській мові, наприклад, для іменників такою є форма називного відмінка однини, для прикметників – форма називного відмінка чоловічого роду однини, для дієслів – форма інфінітива);
- 3) $P(PAR)$ – множина усіх граматичних форм ($[x]$), які набуває слово в процесі словозміни;
- 4) H^{PAR} – оператор відмінювання, що створює словозмінну парадигму для будь-якого слова ($[x]$);

5) H^{LEM} – оператор лематизації, що перетворює будь-яку граматичну форму слова у вихідну.

Наведену вище модель системи словозміни можна також репрезентувати у вигляді схеми (рис. 7.5).

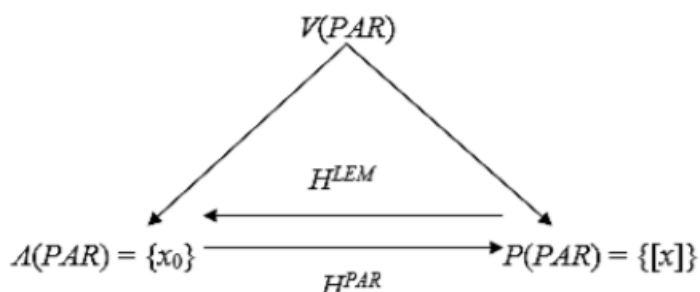


Рис. 7.5. Модель-схема системи словозміни.

Дію оператора відмінювання можна також описати у вигляді моделі, яка відображає власне процес утворення парадигми:

$H^{PAR}x = [x]; [x] = \{x^1, x^2, \dots, x^N\}, x^i = c(x) \star f(x^i)$
де x – певне слово, H^{PAR} оператор відмінювання, $[x] = \{x^1, x^2, \dots, x^N\}$ – повна словозмінна парадигма слова x ; $x^i, i = 1, 2, \dots, N$, – компоненти парадигми; $c(x)$ – основа слова x ; $f(x^i)$ – флексія словоформи x^i .

Одним із прикладів реалізації поданої моделі $V(PAR)$ є комп'ютерна система автоматичної побудови англійських дієслівних парадигм, що була розроблена співробітниками й здобувачами Міжвідомчого центру прикладної лінгвістики на базі кафедри інтелектуальних комп'ютерних систем НТУ «ХП». Через інтерфейс (рис. 7.6) відбувається робота системи за таким алгоритмом:

- Користувач вводить будь-яке речення, що містить дієслова.
- Система виявляє у ньому дієслова та додає їх у перший стовпчик.
- До кожного знайденого дієслова у другий стовпчик додаються початкові форми дієслів (H^{LEM});
- Для знайдених у тексті дієслів система будує повну парадигму (H^{PAR}) та подає її у відповідних стовпчиках.
- На заключному етапі система подає граматичну характеристику типу «present tense, 3rd person, singular», «gerund/present participle», «base form» тощо.

Розроблена програма імітує пошук учнем дієслів та дієвідмінювання ним знайдених дієслів. Вона слугує свого роду тренажером під час самостійних занять із практики англійської мови.

Автоматическое построение английской парадигмы

Введите текст:

Tom has been working for the company for a month, but he still doesn't know how to use the printer. The manager asked Mike to teach him.

Очистить Найти

Найденные глаголы	Начальная форма глагола	1 и 2 форма наст. вр.	3 форма наст. вр.	окончание -ing	прошед. время окончание -ed или 2-3 форма неправ. глагола	Характеристика найденного глагола
0 has	have	have	has	having	had, had	present tense, 3rd person, singular
1 been	be	am, are	is	being	was/were, been	past participle
2 working	work	work	works	working	worked	gerund/present participle
3 does	do	do	does	doing	did, done	present tense, 3rd person, singular
4 know	know	know	knows	knowing	knew, known	base form, present tense
5 use	use	use	uses	using	used	base form, present tense
6 asked	ask	ask	asks	asking	asked	present tense
7 teach	teach	teach	teaches	teaching	taught	present tense

Рис. 7.6. Система автоматичної побудови англійських дієслівних парадигм.

Проблема створення лінгвістичних моделей тісно пов'язана із зародженням комп'ютерної лінгвістики. Першою подією, яка дала поштовх до її появи, став Джорджтаунський експеримент²⁰, проведений 7 січня 1954 р. Його метою була демонстрація можливостей комп'ютера здійснювати переклад з однієї мови іншою. На сьогодні існує чимало перекладацьких систем: Systran, Prompt, Lingee, Deeple, Bing Translator, Microsoft Translator, а також Google Translate, який зараз є найпопулярнішим. Згодом почали з'являтися інші типи систем, що працюють з природною мовою: питально-відповідальні – першу таку систему – ELIZA – розробив Джозеф Вейценбаум у 1966 р. Система імітувала діалог із психотерапевтом, реалізуючи техніку активного слухання. Пізніше з'явилася більш досконала система – A.L.I.C.E. (скор. від Artificial Linguistic Internet Computer Entity; досл. «Штучна лінгвістична інтернет-комп'ютерна спільнота»), що використовує техніку евристичного порівняння фраз користувача з аналогічними, що наявні у її базі даних. На основі таких систем виникли сучасні чат-боти на кшталт Сірі та інші, що використовуються

²⁰ https://uk.wikipedia.org/wiki/Джорджтаунський_експеримент_у_співпраці_з_IBM

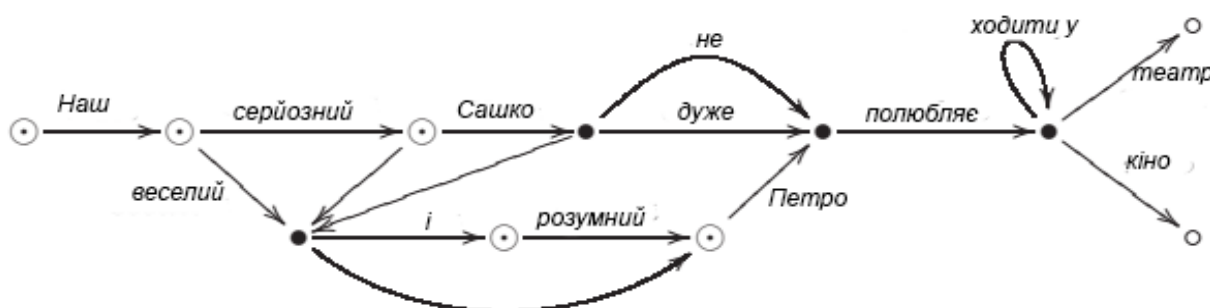
на сайтах комерційних та фінансових організацій та на сайтах з розважальною тематикою. Зараз чат-боти також доступні в різних месенджерах та соціальних мережах. Системи анотування і реферування мають на меті виокремити з наданих текстів найважливішу інформацію та на її основі породити новий текст, що відображатиме короткий та узагальнюючий зміст первинних текстів²¹. Одним напрямком комп'ютерної лінгвістики є комп'ютерна лексикографія, що розв'язує питання укладання електронних словників як для користувачів, так і для систем опрацювання текстової інформації, а також розроблення систем роботи зі словниками у цифровому середовищі (віртуальні лексикографічні лабораторії, системи підготовки словників до видання). Створення комп'ютерних лексикографічних ресурсів також вимагає побудови відповідних моделей, які у формальному вигляді задають репрезентацію лінгвістичної інформації та спосіб її опрацювання.

Контрольні запитання:

1. Що є предметом моделювання в лінгвістиці?
2. Схарактеризуйте поняття «чорної скриньки» у моделюванні.
3. За яких умов модель відповідає реальному об'єкту?
4. Схарактеризуйте наукове і практичне значення моделей, розроблюваних у лінгвістичній галузі.

Практичні завдання

1. Подано модель породжувальної граматики у вигляді схеми, за якою можна отримати чисельну кількість речень:



²¹ Один із прикладів системи автоматичного реферування та анотування текстів, створеної в межах дипломного проекту: <https://openarchive.nure.ua/server/api/core/bitstreams/f5c3d917-9c97-48e6-937c-3158ba9e7d30/content>

Побудова речення починається з тих точок, які обведені кружком. З кожної точки можна йти по одній зі стрілок, виписуючи при цьому слово, яке написано на ній (якщо воно є). Слово в реченні вважається останнім, коли досягаємо точки, з якої не виходить жодної стрілки. Породжувані речення можуть бути як правильними, так і неправильними з погляду граматики.

1а. Побудуйте усі можливі речення за цією моделлю.

1б. Визначте, які правильні та неправильні речення можна отримати.

1в. Змініть схему так, щоб вона породжувала лише правильні речення.

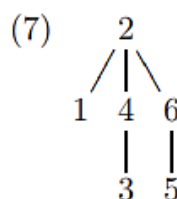
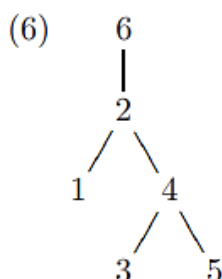
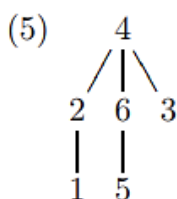
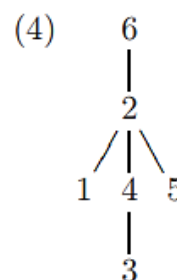
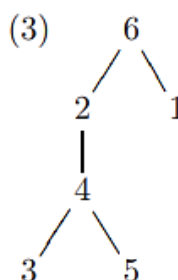
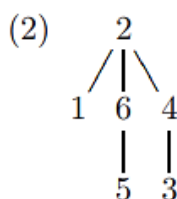
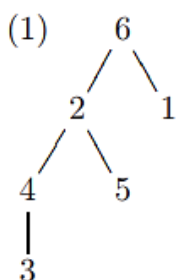
2. Нижче подано речення та відповідні їм **графи**²², що відображають синтаксичну структуру. Встановіть відповідність наведених речень цим графам, а також укажіть закономірності, що забезпечують відповідність:

1. Дорогий подарунок моєму другові з Гонконгу затримався.

2. Я побачив у міському парку маленьку дівчинку.

3. Сильна людина легко підніме цю валізу.

4. Він опублікує це оголошення в нашій газеті.



Заповніть таблицю відповідності між реченнями і графами:

Номер речення	Номер графа

²² Графом називають сукупність об'єктів із зв'язками між ними.

ТЕМА 8 ПОБУДОВА ЛІНГВІСТИЧНИХ МОДЕЛЕЙ

План заняття:

1. Лінгвістична модель та вимоги до неї
2. Лінгвістична модель і «чорна скринька»
3. Моделі у наукових дослідженнях

Створення лінгвістичних моделей не є самоціллю, за їх допомогою розв'язуються певні завдання. Одні моделі мають науково-теоретичний характер, тобто підтверджують гіпотези, створюють нові теорії, а інші – практичний характер, наприклад створення систем автоматизованого перекладу, експертних систем для аналізу усного та писемного мовлення, тональності текстів тощо. Крім цього, методи моделювання використовують у психолінгвістиці, статистичних дослідженнях текстових корпусів, криптографії для дешифрування текстів тощо.

8.1. Лінгвістична модель та вимоги до неї

Моделлю у лінгвістиці називають *штучно створюваний реальний або уявний об'єкт, феномен або процес, що відтворює своєю поведінкою (зазвичай у спрощеному вигляді) поведінку оригіналу в лінгвістичних цілях*. Головна мета моделювання в лінгвістиці – це моделювання здатності людини використовувати мову. Поняття лінгвістичної моделі з'явилося в структурній лінгвістиці (Карл Бюлер, Зеліг Гарріс, Чарльз Гоккет), але увійшло в науковий обіг у 1960-70-ті рр. у зв'язку з виникненням математичної лінгвістики та проникненням у мовознавство ідей і методів кібернетики (Валентина Перебийніс, Володимир Широков, Людмила Алексієнко, Ніна Клименко, Ноам Хомський, Юрій Апресян, Ігор Мельчук, Ісаак Рєвзін, Олена Падучева та інші). Як зазначає Людмила Алексієнко, модель – поняття загальнонаукове, що використовується в математичних, природничих, соціальних і гуманітарних науках. У зв'язку з такою поширеністю воно набуло різних значень. Неоднозначно його розуміють навіть у межах лінгвістики, і найчастіше ним позначають такі типи моделей:

Тип 1: *взірці*, за якими утворюють різні одиниці мови (слова, речення, словосполучення), а також групи мовних одиниць, такі як словотвірна або словозмінна парадигма;

Тип 2: *символи* або *схеми* для опису мовних об'єктів, наприклад схема моделі безпосередніх складників у синтаксисі;

Тип 3: *формалізована теорія структури мови*, до якої належать формальні граматики.

Якщо говорити про структурну лінгвістику, то модель – це теорія структури мови. А якщо говорити більш конкретно – механізм функціонування мови. Оскільки, на думку Людмили Алексієнко, важливою складовою моделі є її символи та правила оперування ними, можна стверджувати, що **модель** у структурній лінгвістиці – *це штучна мова (метамова), яка є гіпотетичним аналогом структури природної мови*. Підсумовуючи вищесказане, **лінгвістична модель** – *це певний абстрактний механізм, зображуваний формальними засобами, який за своєю будовою й принципом свого функціонування імітує поведінку реального об'єкта, який використовує природню мову*.

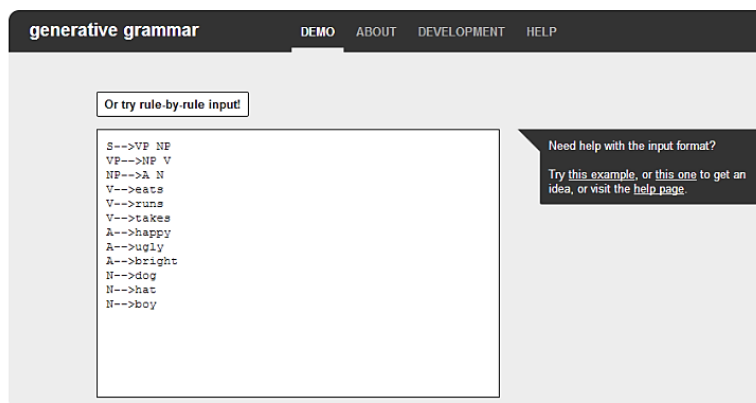
Предметом вивчення сучасної структурної лінгвістики є переважно третій тип лінгвістичних моделей, побудова яких здійснюється відповідно до вимог, основні з яких сформулював Юрій Апресян у своїй праці, присвяченій ідеям і методам сучасної структурної лінгвістики²³:

Вимога 1. *Моделюванню підлягають лише ті явища, суттєві властивості яких вичерпуються їхніми структурними (функціональними) характеристиками та ніяк не пов'язані з їхньою фізичною природою. Моделлю об'єкта слід вважати будь-який функціонально схожий на нього заміник, що відображає лише функціональні властивості цього об'єкта. При цьому матеріал заміника майже завжди відрізняється від матеріалу, з якого побудовано об'єкт (пор., наприклад, нейронну мережу людини і комп'ютерну нейронну мережу). Якщо говорити про моделювання інтелектуального механізму породження*

²³ http://science.lp.edu.ua/sites/default/files/Papers/19_151.pdf

речень, який за Ноамом Хомським називається породжувальною граматикою, то його матеріальним втіленням може стати комп'ютерна програма.

GENERATIVE GRAMMAR



Platform: **Browser** Languages: **JavaScript, CSS3, HTML5** Project Length: **5 months**
Team (Programmers): **Melody Lee, Rose Marie Tan**

Рис. 8.1. Реалізація грамматики Хомського у вигляді комп'ютерної програми.

Вимога 2. *Модель завжди є деякою ідеалізацією об'єкта.* Це означає, що модель має відображати ідеальну поведінку об'єкта, тобто таку, якої не може бути в реальності. Таким чином, модель створює певний ідеальний образ об'єкта, який найчастіше важко досягти, але який зберігає основі якості цього об'єкта з певною точністю. Одним із прикладів ідеалізації, який покладено в основу створення моделі мовного механізму, є те, що він може породжувати нескінченну кількість речень без обмежень до їх довжини (тобто вона може бути більшою за ту довжину, яка була задана раніше). Варто зазначити, що реальна кількість написаних або вимовлених речень будь-якою мовою є доволі великою, але скінченною. Тим не менше, якщо поставити завдання пояснити здатність мовця будувати речення, які раніше він не вимовляв або не писав, то в цьому випадку потрібно розглянути не реально спостережувану ситуацію, а саме цю ідеальну ситуацію.

Вимога 3. *Модель оперує поняттями про ідеальні, а не реальні об'єкти – певні конструкти, які не виводяться безпосередньо і однозначно з даних, а побудовані «вільно» на основі сформульованих гіпотез.* В ідеалі лінгвістичні

конструкти²⁴ будуються безвідносно до фонетичної або семантичної субстанції тих мовних явищ, для вивчення яких вони створені. Так, модель системи мови, що було розглянуто у підрозділі 7.3, показує механізм утворення парадигми, незалежно від субстантивних властивостей конкретних мов і не спирається на традиційні лінгвістичні концепції. Покажемо це на прикладі мовного феномена словозміни в українській, англійській та іспанській мовах. Так, дієслова *мати* (укр.), *to have* (англ.), *avoir* (франц.), *haben* (нім.) як показано нижче у таблиці 8.1, по-різному дієвідмінюються в цих мовах.

Таблиця 8.1. Реалізація словозміни в різних мовах.

Особа / число	Українська	Англійська	Французька	Німецька
1 особа однини	маю	have	ai	habe
2 особа однини	маєш	have	as	hast
3 особа однини	має	has	a	hat
1 особа множини	маємо	have	avons	haben
2 особа множини	маєте	have	avez	habt
3 особа множини	мають	have	ont	haben

Якщо відкинути субстантивну частину слів, то можна побачити, що в аналізованих мовах структуру дієслова складають незмінювана частина (яка залишається незмінюваною під час дієвідмінювання) та змінювана частина, що приєднується до першої. У таблиці змінювану частину позначено жирним шрифтом. Тобто в цих мовах словозміна полягає у поєднанні змінюваних компонентів до незмінюваної частини слова. Відмова від субстантивної частини означає відмову від емпіричної бази, тобто знань про правила утворення парадигми у цих мовах, про відмінності в структурі слова та понять (корінь, суфікс та закінчення), прийнятих у традиційній лінгвістиці та граматиці. Замість цього вводяться поняття квазіоснови та квазіфлексії як елементи інваріантної структури феномена – словозміни. Зверніть увагу, що у французькій мові квазіоснови не мають матеріального вираження.

Вимога 4. *Будь-яка модель повинна бути формальною, тобто в ній мають бути однозначно та у явному вигляді задані вихідні об'єкти та правила*

²⁴ Конструкт – ідеальний об'єкт, вироблений свідомістю (синонім поняття модель).

поводження з ними (правила утворення або виокремлення нових об'єктів і тверджень). В ідеалі будь-яка формальна модель є математичною системою, тому в певному сенсі поняття формальності рівнозначно поняттю математичної точності або однозначності. Формальна модель зв'язується з даними за допомогою тієї чи іншої інтерпретації. Дати інтерпретацію моделі означає вказати правила підстановки об'єктів мови замість об'єктів (символів) моделі. Так, наприклад, модель:

$$H^{PAR}x = [x]; [x] = \{x^1, x^2, \dots, x^N\}, x^i = c(x) \star f(x^i)$$

передбачає правило, за яким квазіфлексії послідовно додаються до квазіоснови, в результаті чого отримуємо словозмінну парадигму. На прикладі українського дієслова *мати* формально це можна зобразити таким чином:

$$c(x) = \text{ма}; f(x^0) = \text{ти}; f(x^1) = \text{ю}; f(x^2) = \text{єш}; f(x^3) = \text{є}; f(x^4) = \text{ємо}; \\ f(x^5) = \text{єте}; f(x^6) = \text{ють};$$

$$i = 0; x^0 = c(x) \star f(x^0) = \text{ма} \star f(x^0) = \text{мати};$$

$$i = 1; x^1 = c(x) \star f(x^1) = \text{ма} \star f(x^1) = \text{маю};$$

$$i = 2; x^2 = c(x) \star f(x^2) = \text{ма} \star f(x^2) = \text{маєш};$$

$$i = 3; x^3 = c(x) \star f(x^3) = \text{ма} \star f(x^3) = \text{має};$$

$$i = 4; x^4 = c(x) \star f(x^4) = \text{ма} \star f(x^4) = \text{маємо};$$

$$i = 5; x^5 = c(x) \star f(x^5) = \text{ма} \star f(x^5) = \text{маєте};$$

$$i = 6; x^6 = c(x) \star f(x^6) = \text{ма} \star f(x^6) = \text{мають}.$$

Вимога 5. *Будь-яка побудована модель повинна мати властивість пояснювальної сили.* Модель володіє цією властивістю, якщо вона: 1) пояснює факти або дані проведеного експерименту, які не можна пояснити за допомогою раніше розробленої теорії; 2) передбачає не відому раніше, але принципово можливу поведінку об'єкта, яка пізніше підтверджується даними спостереження або новими експериментами. Так, наприклад, Ноам Хомський використав метод моделювання прихованого мовного механізму для того, щоб пояснити здатність людини у дитячому віці опанувати мову за досить короткий проміжок часу і тим самим спростував домінуючу на той час теорію

біхевіористів, за якою дитина вчиться правильно говорити за допомогою стимулів та реакцій із зовнішнього середовища.

Контрольні запитання:

1. На основі поданої дефініції схарактеризуйте ознаки поняття лінгвістичної моделі.
2. Які властивості мови є предметом лінгвістичного моделювання?
3. З якою метою створюється та використовується ідеальний образ модельованого об'єкта?
4. Чому для репрезентації лінгвістичних моделей використовують мову математики?

Практичні завдання

1. Подано слова грузинською мовою разом із їхніми відповідниками українською мовою:

mravalsauḱunovani	багатовіковий
ṭomi	том
ṭiraḟi	тираж
didṭiraḟiani	багатотиражний
sauḱune	вік, сторіччя
didmnišvnelovani	багатозначний
saxe	вид
mravalsaxovani	багатоманіття
ozaxi	родина
mravalṭomiani	багатотомний
ataṣi	тисяча
tvali	поверх

Побудуйте формальну модель, яка утворює грузинські прикметники з префіксами, відповідними українському **багато-**. Для цього необхідно:

- визначити грузинські відповідники до українського префіксу **багато-**;
- з'ясувати, в яких випадках потрібно вживати той чи інших грузинський префікс (можливо, це пов'язано зі семантикою прикметника, наприклад: багатовіковий – що має / налічує багато століть);
- виокремити в структурі грузинських слів суфікси, що є формальними показниками прикметників;
- визначити випадки, коли використовується той чи інший суфікс;

– на основі отриманої інформації прописати алгоритм утворення грузинських дієслів;

– побудувати формальну модель утворення грузинських прикметників та визначити її складники.

2. Побудуйте грузинські прикметники за моделлю, утвореною у завданні 1: багатородинний, багатотисячний, багатоповерховий, тисячний та родинний.

З'ясуйте, що означають грузинські слова *didi* та *mravali*?

8.2. Лінгвістична модель і «чорна скринька»

Впровадження конструктивних методів у лінгвістиці спостерігається на початку 1960-х рр., що пов'язано з прагненням застосовувати методи точних наук, зокрема математики, до природної мови. Одне з основних положень математичного конструктивізму наголошує на можливості та необхідності створення теоретичних конструктів, які стають об'єктами теорії тоді, коли їх можна побудувати або змодельовати. При цьому конструктивним визнається не реальний об'єкт, а його теоретичний аналог, або модель. Наприклад, замість реальних висловлювань можна використовувати їхні теоретичні аналоги, а саме структурні схеми або символічні записи у вигляді математичних формул.

У конструктивних методах структурної лінгвістики того періоду об'єктом вивчення було речення, репрезентоване у вигляді теоретичної моделі. Цими методами досліджували не лише синтаксис речення, а також ті явища семантики і лексики, які безпосередньо були пов'язані з реченням. Семантика, так само як і фонологія, виконує по відношенню до синтаксису інтерпретуючі функції. Починаючи з датського структуралізму, конструктивні методи дедалі ширше використовуються для побудови теоретичної моделі мови в цілому. Ще однією важливою особливістю конструктивних методів є бачення мови як динамічної системи, у якій закони побудови речення (висловлювання) є універсальними, тоді як національні, історично змінювані особливості мови розглядаються як форма реалізації цих універсальних закономірностей.

Модель є відображенням властивостей будь-якого об'єкта під час його вивчення. Під час моделювання властивості певного об'єкта переносяться на

модель таким чином, щоб властивості моделі та властивості цього об'єкта були аналогічними. Моделювання застосовують тоді, коли досліджуваний об'єкт недоступний безпосередньому спостереженню. У цих випадках він зазвичай уподібнюється так званій «чорній скриньці», про яку відомо лише, що саме вона отримує «на вході» і який кінцевий результат вона видає «на виході». Завдання полягає в тому, щоб дізнатися зміст «чорної скриньки», тобто прихований від дослідника механізм, який перетворює початкові матеріали у кінцеві продукти. Оскільки неможливо розібрати «чорну скриньку», не порушивши її функціонування, залишається єдиний спосіб пізнання такого об'єкта: на основі зіставлення початкових і кінцевих даних побудувати модель цього об'єкта, тобто висунути гіпотезу про його можливу структуру та реалізувати її, наприклад, у вигляді комп'ютерної програми, яка здатна опрацьовувати деякий матеріал так само, як це робить механізм у «чорній скриньці».

Будь-яка річ, предмет, або явище завжди спочатку виступає як «чорна скринька». Сама назва образно підкреслює повну відсутність відомостей про внутрішній зміст «скриньки»: у моделі задають, фіксують, перераховують тільки вхідні та вихідні зв'язки системи із середовищем. Умовне графічне зображення «чорної скриньки» показано на рисунку 8.2. Це поняття виникло в електротехніці, а потім його почали використовувати в багатьох інших галузях, наприклад, у фізіології вищої нервової діяльності та атомній фізиці, бо ми не можемо спостерігати «безпосередньо» ні діяльності мозку, ні процесів, що протікають в атомі.

Принцип «чорної скриньки» у багатьох випадках є не лише корисним, а й єдино застосовним. Це взагалі стосується таких досліджень, у результаті проведення яких потрібно отримати дані про систему у звичайній для неї обстановці, де слід спеціально дбати про те, щоб вимірювання якомога менше впливали на саму систему. Інша причина того, що доводиться обмежуватися «чорною скринькою», – дійсна відсутність даних про внутрішню будову системи. Наприклад, ми не знаємо, як саме «влаштований» електрон, але

знаємо, як він взаємодіє з електричним, магнітним та гравітаційним полями. Це і є опис електрона на рівні моделі «чорної скриньки».



Рис. 8.2. Графічне зображення чорної скриньки.

У лінгвістичних дослідженнях також широко використовується принцип «чорної скриньки»: єдиною реальністю, з якою лінгвіст безпосередньо має справу, є текст, а механізми мови, які лежать в основі мовленнєвої діяльності людини, не піддаються прямому спостереженню. Тому й в лінгвістиці одним з основних засобів пізнання є побудова моделей. Будь-яке лінгвістичне дослідження відображає, описує і пояснює явища, що існують об'єктивно. Іншими словами, будується модель деякого явища, причому базові принципи її побудови можуть бути різними. Вже з другої половини XX століття лінгвісти під час моделювання широко використовують принцип «чорної скриньки», за яким створювана модель не претендує на точне відображення внутрішньої природи явища, вона лише має бути ідентичною своєму об'єкту в його зовнішніх проявах (невідомо, що відбувається всередині «чорної скриньки», маючи лише дані про те, що відбувається на її вході та виході). У лінгвістиці такою «чорною скринькою» є мовно-розумові процеси, які не піддаються прямому дослідженню.

У теорії формальних граматики Ноам Хомський пропонує розглянути так званий «гіпотетичний механізм засвоєння мови» – «чорна скринька», на вхід якої подається речення, а на виході отримується синтаксична конструкція. З огляду на це «чорна скринька» відображає мовну компетенцію слухача

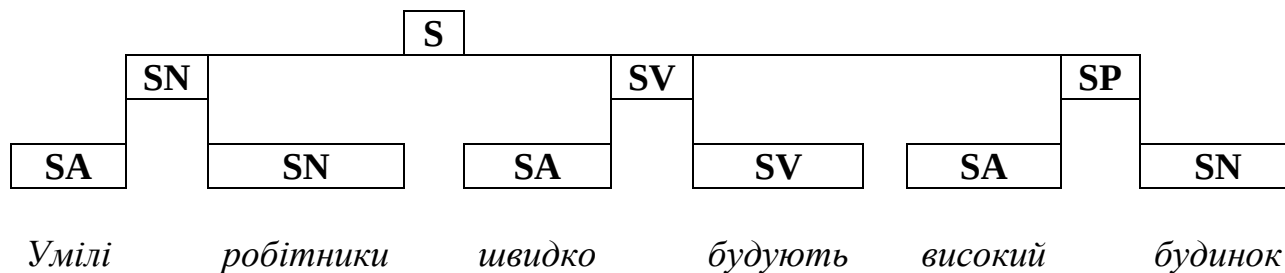
аналізувати мову. Проте залишається невідомим, як саме здійснюється таке перетворення та як саме цей механізм «визначає» граматичну форму речення. Якщо вдасться виявити «граматичний аналізатор», наявний у мовно-розумовому апараті людини, відтворити принципи його роботи, то стає цілком можливим зрозуміти, як працює мислення, формуючи ідеї у кінцеві речення. Під час сприйняття усного або писемного мовлення, носій мови здійснює певний аналіз вхідних даних і за його підсумками отримує граматичний профіль висловлювань. «Зчитуючи» граматичну структуру речень, носій переходить до аналізу змісту. За синтаксичною структурою, яка відображає відношення між елементами речення, встановлюється зміст висловлювання. Так, наприклад, висловлення зі структурою типу *SN-PV-SP* описує багато випадків, коли той чи інший суб'єкт або стан виконують певну дію по відношенню до одного або групи об'єктів. Символами позначено:

1) *SN (subject noun)* – підметовий іменник, тобто особа, місце, річ або ідея, що виконує дію в реченні; його також можна назвати дійовою особою в реченні;

2) *PV (predicate verb)* – дієслово-присудок на позначення дії, що здійснюється по відношенню до об'єкта або групи об'єктів;

3) *SP (subject phrase)* – фраза або речення, що слідує за дієсловом-присудком, допомагаючи розкрити його значення.

Наведемо приклад речення, що відповідає структурі *SN-PV-SP*: *Умілі робітники швидко будують високий будинок*. Структуру цієї фрази можна розгорнути таким чином:



Подане речення містить кілька ідей. Синтаксичне дерево, побудоване за правилами трансформації, дає змогу побачити, що «робітники будують будинок» – це основна ідея. Повертаючись до гіпотетичного аналізатора Ноама

Хомського і до того, як працює мовна здібність, слід зазначити, що люди не лише не усвідомлюють всієї її роботи, а й можуть і не знати, як працює граматика, як породжуються правильні речення, вони просто користуються результатом – розуміють або розмовляють.

Лінгвістичне моделювання і пов'язаний з ним принцип «чорної скриньки» також використовується у перекладацьких дослідженнях, коли потрібно будувати теоретичні моделі процесу перекладу. Оскільки він здійснюється за допомогою низки розумових операцій у голові перекладача, його неможливо спостерігати напряду, і такий процес являє собою «чорну скриньку». Лінгвістичну модель перекладацького процесу можна репрезентувати у вигляді низки послідовних перетворень тексту оригіналу в текст перекладу, за допомогою яких досягається кінцевий результат. Будь-яка модель перекладу має гіпотетичний характер, оскільки немає прямих доказів, що перекладач діє саме так, як це впливає з моделі. Але збіг результату перекладу, прогнозований за моделлю, показує, що вона має певну пояснювальну силу. У сучасному перекладознавстві існує низка моделей, які надають можливість здійснювати переклад різними способами. Один із них, з використанням трансформаційної граматики, наведено у підрозділі 6.3.

Контрольні запитання:

1. Схарактеризуйте зв'язок понять «лінгвістична модель» та «чорна скринька». У чому полягає теоретичне значення принципу «чорної скриньки» в лінгвістиці?
2. Чому модель має бути формальною? У чому полягає формалізація досліджуваного предмета або явища?
3. Чому ідеалізація реального об'єкта є необхідною вимогою для побудови моделі? Поясніть поняття ідеалізації.
4. У чому полягає наукова і практична цінність моделі, зокрема у галузі лінгвістики?
5. Схарактеризуйте наукове значення методу моделювання у галузі перекладознавства.

Практичні завдання

1. Змодельуйте процес семантико-синтаксичного аналізу речень, що реалізує модель «чорної скриньки», користуючись наведеними нижче прикладами. Для цього потрібно використовувати:

- цифри 1, 2, 3 і т. д., що позначають номер кожного слова у реченні;
- символи: –, ~, ⊥, || на позначення відношень між словами.

Приклад 1:

Вхід моделі:	Марина готує синові обід.
Нумерація слів:	Марина – (1); готує – (2); синові – (3); обід – (4)
Формальна структура:	(1) – (2) (2) ~ (3) (2) ⊥ (4)

Приклад 2:

Вхід моделі:	Син Марії швидко зробив домашнє завдання.
Нумерація слів:	Син – (1); Марії – (2); швидко – (3); зробив – (4); домашнє – (5); завдання – (6)
Формальна структура:	(1) (2) (1) – (4) (4) (3) (4) ⊥ (6) (5) (6)

З'ясуйте, які саме відношення позначено символами

Символ	Значення
–	
~	
⊥	

Побудуйте формальну структуру для поданих нижче речень за аналогією до наведених прикладів:

1. Він особливо цінував глибоку повагу дітей до літніх людей.
2. Машину директора ретельно оглядають механіки.
3. Хлопці дуже люблять фільми про шпигунів.
4. Дівчинка смажить біфштекс із баранини для свого брата.
5. Сестри Попелюшки шили сукні з мереживними воланами у найкращого кравця.
6. Широке використання людством техніки вимагає точної відповідності стандартів міжнародним нормам.

2. Визначте, якими саме граматичними конструкціями відображаються в реченнях чотири типи відношень, що розглядалися у завданні 1. Для цього можна використовувати позначення типу *N* (іменник), *V* (дієслово), *A* (прикметник), *Adv* (прислівник) тощо. Для позначення комбінації елементів використовувати знак «+» (наприклад, *N+V* позначає поєднання іменника з дієсловом).

Тип відношення	Конструкція	Приклад конструкції

3. Сформулюйте алгоритм, що характеризує принцип «чорної скриньки» під час семантико-синтаксичного аналізу речень у завданні 1.

8.3. Моделі у наукових дослідженнях

Як було зазначено у підрозділі 7.2., наукове дослідження не починається зі створення моделей та виконання експериментів з ними. Що являє собою наукове дослідження? Це інтелектуальна діяльність людини або людського колективу, що спрямована на отримання нових знань або напрацювань теоретичного та практичного характеру у відповідній проблемній галузі. Слід наголосити, що під «новими знаннями» можуть мати на увазі два моменти. Перший – це отримання щось нового, унікального, тобто того, чого не було досі в науці. А другий – подання нового погляду, нових підходів та нових рішень досі існуючої наукової проблеми.

Наукове дослідження може бути оформлено у вигляді курсової роботи, дипломного проекту, дисертації та монографії. Курсова робота – наукова праця, що здійснюється студентом у межах певної дисципліни для розвитку навичок дослідницької роботи, що включає вміння опрацьовувати літературу, фактичний матеріал дослідження та робити на основі отриманих результатів висновки та узагальнення. Така робота є свого роду попереднім етапом написання дипломної роботи, яка є самостійною науковою працею, у якій студент демонструє отримані знання і навички, а також здатність аналізувати й

систематизувати матеріал, уміння розв'язувати поставлену проблему, виробляти власні напрацювання і показати їхню реалізацію на практиці.

У майбутньому наукова проблема може розглядатися більш глибоко в дисертації, яка являє собою працю науково-кваліфікаційного характеру та характеризує власний теоретичний та практичний внесок автора у розв'язання поставленої наукової проблеми у певній галузі. Окремою науковою працею є монографія – це наукова публікація, присвячена глибокому вивченню однієї теми, що містить нові гіпотези, концепції або теорії. Вона має вузьку спрямованість, максимально докладно і всебічно розглядає проблему, демонструючи глибокі знання і вміння вченого проводити наукові дослідження.

Усі названі вище наукові праці мають мету, завдання, а також об'єкт та предмет дослідження. Чітке їх визначення є важливим попереднім етапом, оскільки від цього залежить успіх будь-якої наукової роботи. Мета роботи – це обґрунтоване уявлення про загальні результати наукового пошуку. Вона містить у собі загальний задум дослідження, тому її формулювання має бути чітким та лаконічним. Завдання визначають проміжні етапи дослідження для досягнення поставленої мети. Об'єкт дослідження – це процес або явище, що породжує проблемну ситуацію і взяте дослідником для вивчення. Предмет дослідження – це той аспект проблеми, досліджуючи який, ми пізнаємо цілісний об'єкт. Не менш важливим питанням є вибір методів. Одним із таких є моделювання, який можуть використовувати поряд з іншими лінгвістичними методами.

Далі відбувається власне дослідницька робота, під час якої будуються відповідні моделі. Юрій Апресян виокремлює такі етапи: 1) фіксування фактів, що потребують пояснення; 2) висування гіпотез; 3) реалізація гіпотез у вигляді моделей, що не тільки пояснюють вихідні факти, але і передбачають нові факти, раніше не спостережувані; 4) експериментальна перевірка моделі. Але ці етапи можуть варіюватися в різних лінгвістичних дослідженнях. Характер і кількість етапів можуть бути різними залежно від мети і завдань лінгвістичного моделювання. Наприклад, додаткові етапи між (3) та (4) також можуть

включати: 3.1) виокремлення основних елементів, що є визначальними для модельованого лінгвістичного об'єкта або явища; 3.2) встановлення взаємозв'язків між цими елементами моделі; 3.3) параметризація компонентів моделі та 3.4) вироблення правил функціонування моделі для отримання потрібних результатів у процесі наукової розвідки.

Для того, щоб модель дійсно була відповідником досліджуваного об'єкта вона також має бути:

1) не абсолютною копією, а спрощеним аналогом, наближеним образом, що відображає найсуттєвіші (з погляду мети моделювання) властивості оригіналу;

2) адекватною, тобто модель повинна якомога точніше відтворювати поведінку оригіналу та породжувати такі самі об'єкти, як і оригінал;

3) не складнішою за оригінал, але разом із цим вона має давати можливість отримати потрібну інформацію про оригінал швидше та ефективніше за інші прийоми дослідження;

4) універсальною, що дає змогу застосовувати модель не лише до одного, а до кількох класів об'єктів;

5) евристичною, тобто модель повинна мати можливість генерувати нові знання про оригінал.

Наприклад, метод моделювання застосовано у студентській науковій роботі, що присвячена розроблянню системи автоматичного дієвідмінювання (див. підрозділ 7.3), яка: 1) виявляє з англійського речення дієслова, 2) аналізує їхню граматичну форму в контексті, 3) перетворює їх у вихідну форму і 3) будує усі можливі граматичні форми (парадигму) для кожного виявленого дієслова. Систему реалізовано у вигляді комп'ютерної програми, що автоматично виконує усі лінгвістичні операції. Основні етапи дослідження включають: 1) підготовчий: визначення мети і завдань дослідження, вибір джерел; 2) аналіз лінгвістичного матеріалу для вибору основних параметрів моделі для розроблюваної системи; 3) побудова формальної моделі системи англійського дієвідмінювання і 4) програмна реалізація, що передбачає

вироблення алгоритму та інтерфейсу програми. Отже, на першому етапі за мету роботи взято створення програмного забезпечення для автоматичної побудови англійської дієслівної парадигми, морфологічну характеристику словоформи та початкову форму слова, що наявна в тексті. Об'єктом дослідження обрано явище дієслівної словозміни в англійській мові, а предметом – формальне моделювання дієслівної словозміни та його реалізація у вигляді комп'ютерної програми. На другому етапі здійснено аналіз лінгвістичного матеріалу для вибору основних параметрів моделі для розроблюваної системи. Для цього було здійснено класифікацію англійських дієслів за словозмінними властивостями, тобто за відповідними граматичними типами й граматичними групами (див. рис. 8.3).

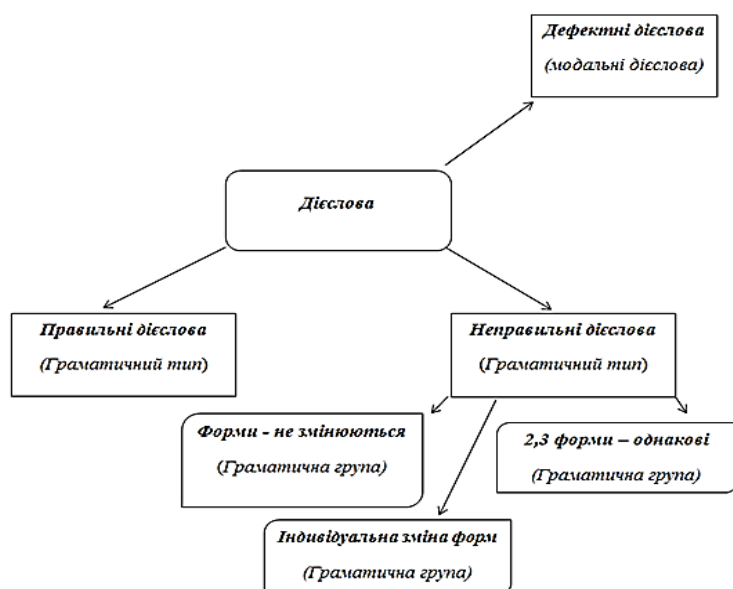


Рис. 8.3. Класифікаційна схема англійських дієслів.

За результатами проведеного лінгвістичного аналізу побудовано формальну модель самої системи мови, у якій відбувається явище словозміни (рис. 8.4). Зазначена система відповідає теорії системної тріади Володимира Широкова, про яку детально викладено у підрозділі 5.3. Зміст елементів модельованої системи подано у таблиці 8.2.

Таблица 8.2. Складники системної тріади.

Структура	Суб'єкт	Субстанція
Будова феномена словозміни	1) Морфологічний аналіз текстових словоформ;	1) Англійські речення

	2) Лематизація довільної текстової словоформи; 3) Формування повної словозмінної парадигми	2) Словоформи 3) Елементи словоформ
--	---	--

Подана модель (лексикографічна система), на основі якої здійснюється подальше створення комп'ютерної системи автоматичної побудови англійських дієслівних парадигм (див. підрозділ 7.3), умовно поділяється на ліву $A(PAR)$, яку складає множина дієслів у вихідній формі x_0 , та праву частину, що репрезентована множиною $P(PAR)$.

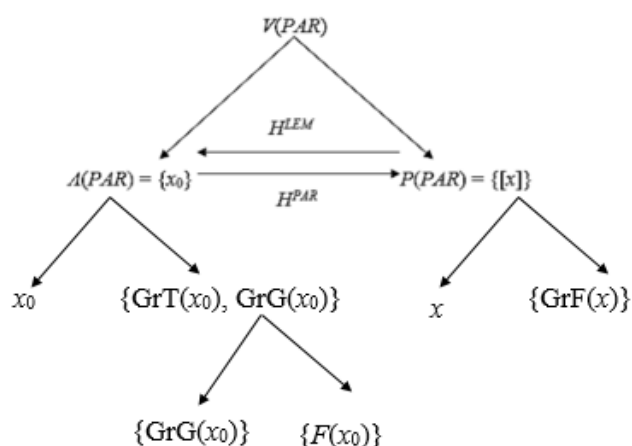


Рис. 8.4. Детальний вигляд моделі системи словозміни.

Своєю чергою, ліву і праву частини можна також поділити на окремі підсистеми. У лівій частині підсистема встановлює відповідність дієслова x_0 до певного граматичного типу $GrT(x_0)$ та граматичної групи $GrG(x_0)$, виокремлених під час лінгвістичного аналізу на другому етапі. Далі граматичній групі надається відповідний набір квазіфлексій $F(x_0)$. Права частина складається зі словоформ парадигми x та параметрів, що характеризують саму граматичну форму $GrF(x)$. На останньому етапі дослідження створено базу даних та інтерфейс програми на основі розробленої моделі.

Контрольні запитання:

1. Які переваги надає метод моделювання у лінгвістичних дослідженнях та в яких випадках його використовують?
2. Що необхідно зробити перед тим, як будувати ту чи іншу лінгвістичну модель?

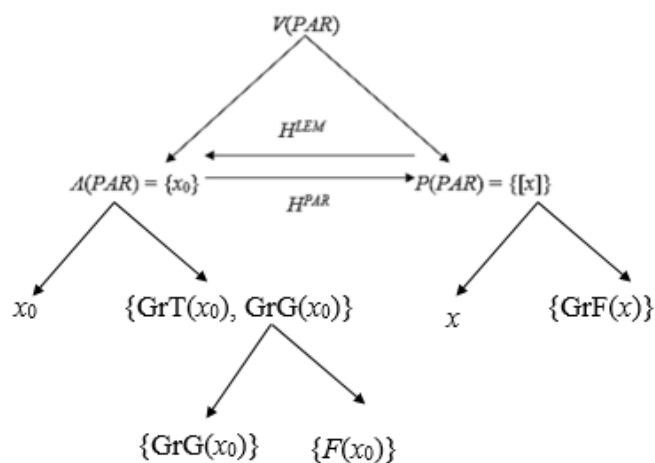
3. Чи є коректним стверджувати, що формальна модель мови є аналогом природної мови?

4. Які функції природної мови можуть бути предметом лінгвістичного моделювання?

5. Які лінгвістичні дослідження не вимагають методу моделювання?

Практичні завдання

Нижче подано модель, яка відображає у формальному вигляді систему словозміни:



Визначте усі можливі значення параметрів моделі для наведених іменників:

блат, болт, верблюд, гриб, депо, джміль, кермо, кіт, кольє, м'яч, мужва, народ, піп, піт, полон, світло, сором, спирт, страх, таксі, фарба, хміль, штани, шторм, шум, ягода.

Для цього потрібно:

– зробити класифікацію наведених іменників за їхніми словозмінними характеристиками, тобто розподілити їх по граматичних типах $GrT(x_0)$ та групах $GrG(x_0)$:

$GrT(x_0)$	$GrG(x_0)$	Іменники

– провідмінювати іменники та визначити набір квазіфлексій $F(x_0)$ для кожної групи іменників $GrG(x_0)$:

$GrG(x_0)$	$F(x_0)$

– подати параметри граматичної форми $\text{GrF}(x)$, що входить до словозмінної парадигми (на прикладі лише одного іменника з кожної граматичної групи)

x	GrF(x)

ТЕМА 9 КЛАСИФІКАЦІЯ ЛІНГВІСТИЧНИХ МОДЕЛЕЙ

План заняття:

1. Базова класифікація
2. Класифікація за об'єктом моделювання
3. Класифікація за способом побудови і аналізу мовних об'єктів

На сьогодні існує безліч лінгвістичних моделей, і тому подання вичерпної класифікації всіх їх можливих типів є доволі складним завданням. У цьому розділі подано основні критерії, за якими класифікуються лінгвістичні моделі, і коротко охарактеризовано особливості їх типів. Кожен із них має своє призначення, і тому ця інформація є визначальною для вибору необхідних моделей для досягнення мети теоретичного або прикладного дослідження.

9.1. Базова класифікація

Базова класифікація лінгвістичних моделей передбачає їх об'єднання за загальними критеріями: 1) охоплення структури мови, 2) типологічний статус, 3) гносеологічний статус, 4) мета наукового дослідження, 5) наукові методи, 6) функціональний статус, 7) використання матеріальних засобів створення, 8) предмет моделювання. Цілком можливе існування інших типів лінгвістичних моделей, що не входять до цієї класифікації. Слід наголосити, що класифікація в деяких випадках ускладнюється тим, що та чи інша модель може відповідати кільком критеріям.

За охопленням структури мови, моделі поділяються на **глобальні**, які *охоплюють мову в цілому*, та **часткові**, що *описують лише окремий її рівень*: фонетичний, морфологічний, лексико-семантичний та синтаксичний. За своєю сутністю, глобальна лінгвістична модель являє собою теоретичний опис загальних закономірностей будови й функціонування часткових моделей, по відношенню до яких вона виступає як метамодель²⁵. Прикладом глобальної моделі є теорія мовних рівнів, яку розробляли ранні структуралісти, такі як Еміль Бенвеїніст, Іван Бодуен де Куртене. До часткових належать, наприклад, фонологічна система англійської мови або система голосних і приголосних

²⁵ **Метамодель** – це модель, що описує структуру та принципи роботи іншої моделі.

фонем англійської мови тощо. Наприклад, Іван Бодуен де Куртене запропонував теоретичну концепцію, за якою в мові виокремлюються три рівні: «фонетична будова слів і речень», «морфологічна будова слів» і «морфологічна будова речень». Разом із цим науковець також виокремлював «зовнішню» (фонетичну), «позамовну» (семантичну) та «власне мовну» (морфологічну) сторони мови.

За типологічними характеристиками, моделі поділяються на **універсальні**, що *охоплюють або прагнуть охопити всі мови світу*, та **специфічні**, тобто *призначені для конкретної мови або групи мов*. Одним із прикладів універсальної моделі є відмінкова граматика Чарльза Філмора (див. підрозділ 6.2), за якою кожне дієслово у реченні керує певним набором відмінків залежних слів. Модель семантичних відмінків є універсальною, оскільки її можна застосувати до більшості мов світу. Специфічними моделями у цьому випадку можуть бути ті, що показують реалізацію семантичних відмінків, наприклад, в англійській або українській мовах. Ще одним прикладом універсальної моделі є генеративний синтаксис, що вивчає механізм побудови граматично правильних речень (див. підрозділ 4.1) за допомогою певного набору правил.

За гносеологічним статусом розрізняють **моделі мови**, що використовуються для *опису мови та окремих її аспектів* (фонологічних, граматичних, лексичних тощо) та **моделі лінгвістичних знань**, до яких відносяться *різні концепції щодо будови та функціонування мови, притаманні тим чи іншим лінгвістичним школам (Женевська лінгвістична школа, Копенгагенська школа, Празький лінгвістичний осередок та інші) або окремим ученим*. Моделі мови можуть бути реалізовані, наприклад, у вигляді словників, тезаурусів, граматик, різних таксономічних схем тощо. Прикладами моделей лінгвістичних знань є «Глосематика» (див. підрозділ 2.2), що репрезентує мову як мережу функцій; теорія Ноама Хомського про механізм засвоєння мови дитиною (див. тему 10), теорія Олександра Мельчука «Смисл ↔ Текст», що висвітлює механізм породження текстів із заданим змістом, теорія системи

мови Володимира Широкова, що розкриває механізм утворення мовних явищ різного типу (див. підрозділ 4.2).

За метою дослідження виокремлюють **теоретичні, описові та прикладні моделі**. **Теоретична модель** являє собою *теоретичне та схематичне уявлення про реальність за допомогою ідеалізованих абстракцій або об'єктів*. Її також розуміють як сукупність абстрактних об'єктів, що описують проблемне поле, яке викликає теоретичний інтерес з боку дослідника та репрезентує собою єдине логічне ціле. Наприклад, у підрозділі 6.3 було розглянуто відображення процесу перекладу за допомогою трансформаційної моделі. Ця модель оперує такими об'єктами, як речення мовою оригіналу, мінімальні ядерні структури, інваріантні структури, речення мовою перекладу. Зв'язок між цими об'єктами можна схарактеризувати таким чином: 1) речення мовою оригіналу розкладається на мінімальні ядерні структури, що містять у собі елементи його змісту; 2) на основі отриманих ядерних структур утворюється загальна структура, що відображає план змісту вихідного речення; 3) ця структура трансформується у речення мовою перекладу. **Описовою моделлю** називають *будь-який опис об'єкта*. Якщо, наприклад, об'єктом є лексична система, то її описовою моделлю можуть бути тлумачні словники та тезауруси. **Прикладні моделі** *призначені для розв'язання практичних завдань*. Приклад розробляння такої моделі наведено в підрозділі 7.3, у якому мова йде про створення системи автоматичної побудови англійських дієслівних парадигм.

Якщо говорити про методи інших наук, моделі можуть бути, наприклад, математичними, соціологічними, психологічними тощо. А залежно від матеріальних засобів створення, моделі поділяються на графічні, символні та комп'ютерні. Якщо говорити про сучасну структурну лінгвістику, для розв'язання того чи іншого завдання використовують математичні моделі, які згодом можуть бути реалізовані у комп'ютерному середовищі (у вигляді комп'ютерних моделей). Математичний апарат дає змогу чітко репрезентувати модельований феномен, зокрема його будову та механізм функціонування.

Соціологічні і психологічні методи є основою відповідних моделей перекладу²⁶.

Залежно від предмету моделювання, тобто від того, який аспект мовленнєвої діяльності розглядається, розрізняють **моделі аналізу, моделі синтезу, породжувальні та дослідницькі** моделі. Ці моделі належать до формальних, оскільки в них чітко прописані об'єкти, їх взаємозв'язки та правила поводження з ними. Такі моделі використовують для відтворення об'єкту або явища, які неможливо безпосередньо спостерігати (поняття «чорна скринька»).

Моделлю аналізу називають *модель, що містить кінцеву кількість правил, за якими можна проаналізувати безліч речень певної мови*. До них належать синтаксичні аналітичні моделі, які на «вході» отримують текст, а на «виході» утворюють для кожного речення запис його синтаксичної структури. Такі моделі реалізовано у вигляді синтаксичних аналізаторів²⁷, приклад роботи одного з яких показано на рис. 9.1.

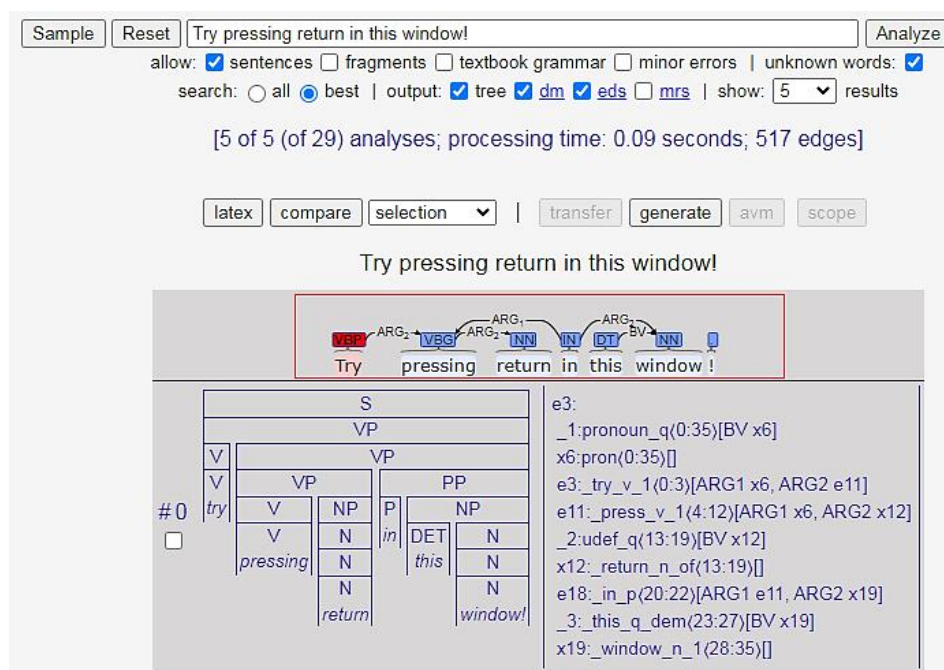


Рис. 9.1. Синтаксичний аналізатор LinGO English Resource Grammar (ERG).

²⁶ Детально про моделі перекладу:

<http://baltijapublishing.lv/omp/index.php/bp/catalog/download/100/2560/5427-1?inline=1>

²⁷ <https://mova.institute/>

<https://visl.sdu.dk/visl/en/parsing/automatic/dependency.php>

<http://erg.delph-in.net/logon>

Модель синтезу надає скінчену кількість правил, за якими можна побудувати велику кількість правильних речень. Синтаксичні синтетичні моделі використовують як вихідну інформацію запис синтаксичної структури речень, а на виході видають правильні речення цієї мови. Семантичні синтетичні моделі отримують на «вході» смисловий запис деякого речення спеціальною семантичною мовою і видають на «виході» множину речень природної мови, синонімічних цьому реченню (моделі, побудовані за теорією «Смисл ↔ Текст»). Моделі синтезу можуть використовуватися, наприклад, у системах автоматичного перекладу, реферування й анотування текстів, а також систем перефразування текстів. Приклад роботи системи перефразування текстів AISEO²⁸ показано на рис. 9.2.

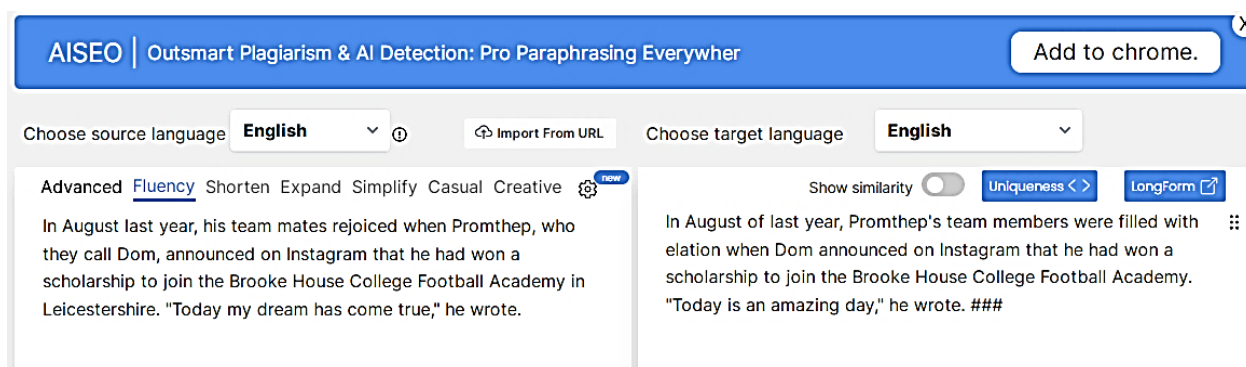


Рис. 9.2. Приклад роботи системи перефразування текстів.

Проміжну ланку між моделями аналізу й синтезу посідають **породжувальні моделі** – механізми, що можуть, по-перше, генерувати безкінечну множину правильних речень за допомогою кінцевого числа граматичних правил, а, по-друге, імітувати здатність відрізняти правильні речення від неправильних. Теоретичну основу для побудови цих граматик створив Ноам Хомський (див. теми 10-12).

Дослідницькі моделі – це такі моделі, що імітують діяльність науковця-лінгвіста. Залежно від того, яку інформацію вони отримують на «вході» і що саме вони видають на «виході», ці моделі поділяють на три класи (таблиця 9.1.).

²⁸ <https://aiseo.ai/tools/sentence-generator.html>

Таблиця 9.1. Характеристика дослідницьких моделей.

Клас моделі	«Вхід»	«Вихід»	Приклади моделей
Клас 1	Тексти певною мовою	Словник і граматики цієї мови	Дешифрувальні моделі
Клас 2	Текст + множина правильних фраз	Правильність або неправильність фраз	Експериментальні моделі
Клас 3	Текст + фрази + семантичні інваріанти	Правильність або неправильність фраз, ідентифікація фраз спільних або різних за змістом	Граматики, словники, опис того чи іншого лінгвістичного явища

Контрольні запитання:

1. Які типи лінгвістичних моделей ґрунтуються на принципі «чорної скриньки»?
2. Які типи лінгвістичних моделей можуть бути використані у галузі перекладознавства та комп'ютерної лінгвістики?
3. До якого типу лінгвістичних моделей належать словники? Які елементи включає модель словника?

Практичні завдання

1. Побудуйте описову модель лексико-семантичної підсистеми, що репрезентована дієсловами, утвореними від іменників та прикметників:

бешкетувати, глохнути, голодувати, байдикувати, пиляти, звірити, лютувати, потребувати, шуміти, сумувати, кам'яніти, рибалити, блаженствувати, боронувати, товстіти, бинтувати, стовбеніти, свердлити.

Для цього необхідно:

– розподіліть дієслова таким чином, щоб у межах кожної групи слова мали лише одну пояснювальну конструкцію, причому ця конструкція не може бути застосована до інших груп дієслів;

– навести до кожної групи пояснювальну конструкцію:

Група дієслів	Пояснювальна конструкція

Наприклад, пояснювальна конструкція може мати такий вигляд: «читач – людина, яка **читає**» або «їдальня – місце, де **їдять**».

2.* Побудуйте формальну модель, яка відображає набуття словом набору граматичних та лексичних властивостей, на основі інформації, що подана у наведеній нижче словниковій статті (<http://sum.in.ua/s/voda>):

ВОДА

1. Прозора, безбарвна рідина, що становить собою найпростішу хімічну сполуку водню з киснем.

2. *перев. мн.* Водна маса джерел, озер, річок, морів, океанів.

3. *тільки мн.* Лікувальні мінеральні джерела, а також курорт з такими джерелами.

4. *тільки мн., мед.* Навколоплідна рідина.

Щоб побудувати модель, необхідно:

– виокремити граматичні та лексичні властивості, що зафіксовані у словниковій статті;

– визначити компоненти, що складають модель, ґрунтуючись на аналізі словникової статті;

– визначити параметри до цих компонентів;

– подати модель у вигляді символного або математичного запису.

9.2. Класифікація за об'єктом моделювання

Мета цієї класифікації (див. рис. 9.3) – об'єднати лінгвістичні моделі, які відтворюють мовні процеси і явища, дослідницькі процедури та готові описи. Моделі, в яких як об'єкт виступають конкретні мовні процеси і явища, поділяють на ті, що: (а) імітують мовленнєву діяльність людини; (б) імітують дослідницьку діяльність лінгвіста; (в) являють собою метатеорії. Моделі групи (а) описують інтелектуальні процеси, що пов'язані з мовленнєвою діяльністю. У моделях групи (б) за об'єкт беруться процедури, що ведуть ученого до виявлення того чи іншого мовного явища. Ці моделі певною мірою імітують дослідницьку діяльність лінгвіста (їх також називають моделями дослідження). Першими, хто зробив серйозний крок у їх розробленні, були американські дескриптивісти, які намагалися створити універсальні алгоритми відкриття граматик природних мов (див. у підрозділі 2.3 про дистрибутивний аналіз).

Моделі (в) репрезентують готові лінгвістичні описи, а не мовленнєву чи дослідницьку діяльність (див. у підрозділі 10.1 про теорію формальних мов та формальних граматик Ноама Хомського). У цьому підрозділі мова буде йти про моделі групи (а).

До моделей, що імітують мовленнєву діяльність людини, належать **моделі граматичної правильності**, що відображають уміння відрізняти правильне від неправильного в мові, та **функціональні моделі**, що розкривають уміння співвідносити зміст мовлення (план змісту) з його формою (планом вираження). Прикладом моделей граматичної правильності є граматики Ноама Хомського, які розглядаються в темах 10 і 11. А прикладом функціональних – модель «Смисл ↔ Текст», основоположником якої є Ігор Мельчук і яку буде розглянуто в темі 12.



Рис. 9.3. Класифікація моделей за об'єктом.

Залежно від типу інформації на «вході» та на «виході» моделі граматичної правильності поділяють на **розпізнавальні** та **породжувальні**. Розпізнавальна модель одержує на «вході» деякий відрізок тексту природною мовою або його абстрактну репрезентацію штучною мовою і дає на «виході» відповідь, чи є цей відрізок граматично правильним або неправильним. Породжувальна модель є зворотною по відношенню до розпізнавальної. Цей тип моделей використовують для породження речень на запит. Образно кажучи, натисканням кнопки буде згенеровано деякий ланцюжок певною мовою.

Наведемо простий приклад, що демонструє різницю між розпізнавальною та породжувальною моделями. Нехай існує модель такого вигляду:

$$\{V_N, V_T, R\},$$

де V_N – множина елементів речення, V_T – множина значень, які можуть бути присуджені елементи речення, і R – правила, за якими елементи речень набувають свого значення, інакше кажучи – правила побудови речення. Нехай $V_N = \langle S, N, V \rangle$, літери позначають речення (S), іменник (N) та дієслово (V); а $V_T = \langle \text{т, н, п, с} \rangle$ – символами умовно позначено слова «тварини», «не», «п'ють», «спиртне». Якщо це породжувальна модель, то правила (R) включатимуть:

$$R = \{S \rightarrow N + V + N \quad V \rightarrow \text{н} + V \quad N \rightarrow \text{т} \quad N \rightarrow \text{с} \quad V \rightarrow \text{п}\}$$

Застосовуючи наведені правила, модель видає ланцюжок у вигляді речення $\langle \text{т, н, п, с} \rangle$:

$$\begin{aligned} S &\rightarrow N + V + N \rightarrow \text{т} + V + N \rightarrow \text{т} + \text{н} + V + N \rightarrow \text{т} + \text{н} + \text{п} + N \\ &\rightarrow \text{т} + \text{н} + \text{п} + \text{с} \end{aligned}$$

У випадку з розпізнавальними моделями, набір правил моделі (R) матиме такий формальний вигляд:

$$R = \{N + V + N \rightarrow S \quad \text{т} \rightarrow N \quad \text{н} + V \rightarrow V \quad \text{с} \rightarrow N \quad \text{п} \rightarrow V\}$$

Як видно, розпізнавальна граматика для наведеного прикладу містить зворотні правила. Аналіз правильності речення за допомогою правил розпізнавальної граматики наведено на рис. 9.5.

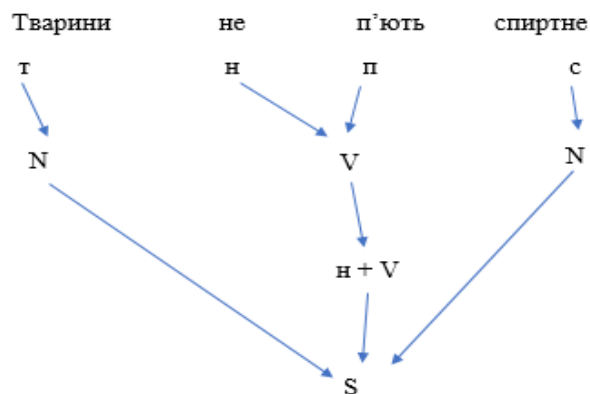


Рис. 9.4 Схема аналізу граматичної правильності речення.

За результатом аналізу модель на «виході» дає, що подане речення є граматично правильним, оскільки воно відповідає наведеним граматичними

²⁹ Що буде, якщо поміняти місцями іменники (N)?

правилам побудови ланцюжка. Але якщо поміняти місцями слова на кшталт «Тварини спиртне не п'ють», то це речення, хоча і буде граматично правильним для української мови, але сама побудова речення не буде відповідати заданим правилам моделі. У той же час ця модель може згенерувати речення, яке відповідатиме заданим правилам і буде формально граматично вірним, але повністю втратить свій сенс або семантику. Саме тому було розроблено подальші моделі – функціональні.

Як приклад функціональної моделі, розглянемо детально модель «Смисл ↔ Текст», за якою мова репрезентується як засіб здійснення її носіями двох інтелектуальних операцій: 1) повідомлення своїх думок іншим людям, тобто кодування смислів (змісту), що закладені у текстах (побудова текстів, синтез); 2) розуміння думок інших людей, тобто виконання зворотної операції з вилучення смислів (змісту) зі сприйманих текстів (розуміння текстів, аналіз). Модель можна уявити як логічний пристрій, що імітує зазначені операції в їх найпростіших проявах, пов'язаних виключно зі знанням мови (словника та граматики). Подібна модель деякої конкретної мови – це складно організована сукупність правил, суто механічне застосування яких має в ідеалі забезпечувати:

- перехід від заданого тексту (або текстів) певною мовою до формального опису сенсу цього тексту, тобто до його семантичного подання, тобто змістового запису;

- перехід від заданого змісту, тобто семантичного подання, до відповідного тексту (або текстів) певною мовою.

Наведемо приклад породження низки синонімічних фраз, однією з яких є, наприклад: «Орвелл упевнений у тому, що його активність у політиці робить його твори кращими». Семантична структура, що задає спільний зміст фраз, показана на рисунку 9.5.

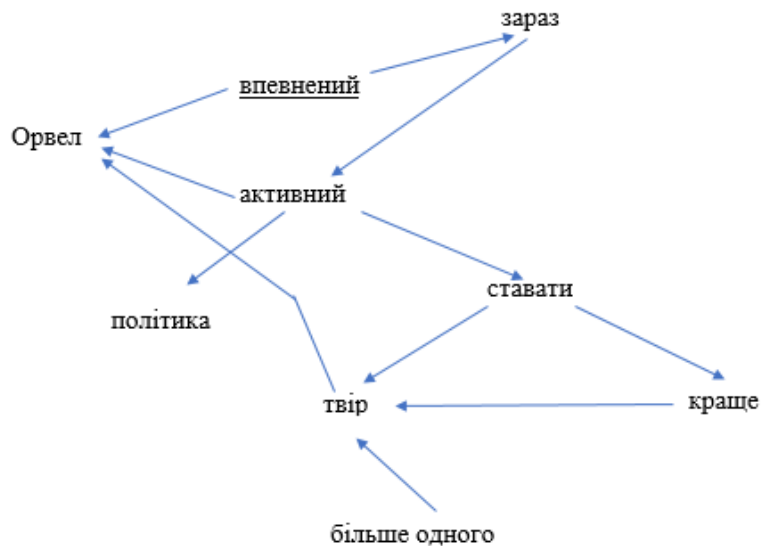


Рис. 9.5. Семантична структура фрази.

Підкреслення слова «упевнений» означає, цей елемент семантичної структури є комунікативно-домінантним. Наведену структуру можна реалізувати у вигляді таких фраз:

1. Орвелл не сумнівається в позитивному впливі його політичної активності на якість його творів.
2. Орвелл не сумнівається, що те, що він активний у політиці, робить його твори кращими.
3. Орвелл не має сумнівів у тому, що його політична активність позитивно позначається на якості його творів.
4. Орвелл впевнений у тому, що його твори стають кращими завдяки його політичній діяльності.
5. Те, що якість його творів покращується внаслідок його політичної діяльності, не викликає в Орвелла жодних сумнівів.
6. Орвелл переконаний, що його політична діяльність справляє позитивний вплив на його творчість.
7. Орвелл упевнений, що якість його творів підвищується завдяки його активності в політиці.
8. На переконання Орвелла, його твори робляться кращими завдяки його політичній заангажованості.

Модель «Смисл ↔ Текст» найповніше розкриває феномен володіння мовою, який передбачає здібність: а) обирати мовні одиниці (не тільки лексичні), що виражають заданий зміст; б) правильно поєднувати мовні одиниці з певними значеннями; в) перифразувати висловлювання зі збереженням їхнього змісту (семантичного інваріанта).

Залежно від того, який аспект мовленнєвої діяльності моделюється – слухання чи говоріння, – функціональні моделі поділяють на **аналітичні** та **синтетичні**. Повна аналітична модель певної мови отримує на «вході» деякий відрізок тексту (зазвичай не менший, ніж висловлювання) і дає на «виході» його змістовий запис (семантичну репрезентацію) семантичною метамовою (тобто його тлумачення). Повна синтетична модель деякої мови, будучи протилежною до повної аналітичної моделі, на «вході» отримує семантичний запис (тобто уявлення деякого фрагмента смислу), а на «виході» формує множину синонімічних текстів цією мовою, що репрезентують цей смисл. Моделі аналізу та синтезу становлять важливий складник деяких комп'ютерних систем перекладу, а також різних систем, що ґрунтуються на технології штучного інтелекту (наприклад, питально-відповідальні системи).

Контрольні запитання:

1. Схарактеризуйте різницю між моделями граматичної правильності та функціональними моделями.
2. Визначте спільні і відмінні риси, що мають між собою породжувальні та розпізнавальні моделі.
3. Схарактеризуйте прикладне значення моделі «Смисл ↔ Текст» у комп'ютерній лінгвістиці та перекладознавстві.

Практичні завдання

1. *Сформулюйте граматичні правила, щоб модель породжувала речення зі структурою типу:*

Qualified managers provide adequate support.

2. *Побудуйте максимально можливу кількість спільних за змістом речень за поданою семантичною мережею:*



3. Побудуйте можливі ідентичні за змістом тексти перекладу для поданих нижче англійських речень з оповідання «Doctor in the House»:

1. It is the physical contact with the examiners that makes oral examinations so unpopular with the students.

2. The written answers have a certain remoteness about the examiners, and mistakes and omissions, like those of life, can be made without the threat of immediate punishment.

3. There were six candidates from other hospitals waiting to go in with me, all of them in their best clothes.

Для цього необхідно:

– побудувати семантичну мережу, в якій потрібно вказати зв'язок елементів, що складають зміст речення;

– визначити комунікативно-домінантний елемент у семантичній мережі;

– побудувати синонімічні речення англійською мовою, так щоби зміст залишався таким самим;

– подайте переклад побудованих речень у тому числі засобами машинного перекладу.

9.3. Класифікації за способами побудови або аналізу мовних об'єктів

Крім двох наведених класифікацій лінгвістичних моделей існують також інші – за способами побудови або аналізу мовних об'єктів. Так, за способом побудови розрізняють **індуктивні**, що розробляються на основі аналізу певного мовного матеріалу для формування гіпотези щодо будови та функціонування лінгвістичного об'єкта або явища, та **дедуктивні**, які будуються на основі якоїсь гіпотези та перевіряють її істинність на певному мовному матеріалі.

Зазначені типи в ідеалі мають доповнювати один одного. Індуктивні моделі можуть охоплювати як окремі мовні явища, так і цілі підсистеми і навіть мовні системи. Дослідження універсальних або типологічних рис мов дає змогу розробляти дедуктивні моделі – штучно створювані системи, що гіпотетично являють собою праобрази реальних систем мови. Порівняльна характеристика моделей подана у таблиці 9.2.

Таблиця 9.2. Індуктивні та дедуктивні моделі.

Індуктивні моделі	Дедуктивні моделі
Прийом, на якому ґрунтується побудова	
Індукція	Дедукція
Основа для побудови	
Фактичний матеріал	Гіпотетичні ситуації
Прийом опрацювання матеріалу	
Аналіз	Синтез
Призначення	
Показати будову та / або принципи функціонування мовних феноменів	Показати істинність сформульованих гіпотез

Як приклад індуктивної моделі, можна навести модель утворення атрибутивних словосполучень (наприклад: *високий будинок, пройдений шлях, це цікаве газетне повідомлення* тощо), яка наведена у статті «Про моделювання в синтаксисі як метамовну проблему»³⁰. Спочатку мовленнєві ланцюги замінюють їхніми образами, наприклад: *нова книжка* – **Пк Ім**; *мій друг* – **За Ім**; *прочитаний лист* – **Дк Ім**; *моя біла сукня* – **За Пк Ім**; *велика кам'яна хата* – **Пк Пк Ім**; *це цікаве газетне повідомлення* – **За Пк Пк Ім** тощо. Тут варто зазначити, що *моя біла сукня* – це конкретне словосполучення, а **За Пк Пк Ім** – його образ. Якщо звести усі можливі образи у єдиний вираз, то можна отримати модель, у якій позиційно еквівалентні класи слів виписані в стовпчик за зменшенням їхньої частоти, а факультативні (не обов'язково наявні в цій конструкції) класи беруть у круглі дужки, кількість яких указує різний ступінь факультативності:

³⁰ http://enpuir.npu.edu.ua/bitstream/handle/123456789/29335/Popova_148-156.pdf?sequence=1

$$\begin{pmatrix} \text{Пк} \\ (\text{За}) \\ \text{За} \end{pmatrix} \begin{pmatrix} \text{Пк} \\ \text{За} \\ \text{Дк} \end{pmatrix} \text{Ім}$$

За способом аналізу мовних об'єктів розрізняють **статичні** та **динамічні** моделі. Статичні моделі репрезентують лінгвістичні об'єкти (зокрема, їхню будову) в статистиці та використовують описовий метод, а динамічні відтворюють механізм породження того чи іншого лінгвістичного об'єкта та створюються засобами математичного (формального) моделювання. До статичних належать **структурні** та **класифікаційні** моделі: перші відображають складники об'єкта, а другі – різновиди цих складників. Такі моделі використовуються у таксономічній лінгвістиці (див. підрозділ 5.2) у вигляді блок-схем, діаграм, таблиць тощо. До динамічних належать формальні граматики (див. підрозділ 10.1) та модель «Смисл ↔ Текст» (див. підрозділ 12.3).

Контрольні запитання:

1. Як можуть взаємно доповнюватися індуктивні та дедуктивні моделі?
2. Чи можуть під час певного дослідження застосовуватися у той самий час статичні і динамічні моделі? Наведіть приклади.

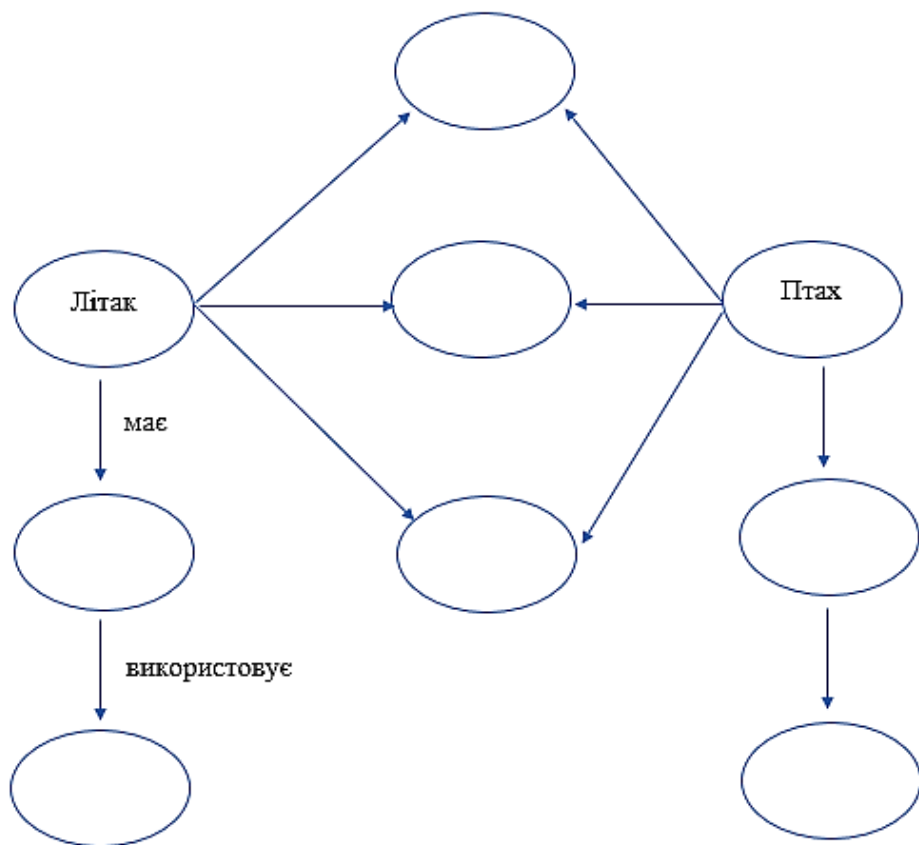
Практичні завдання

1. Побудуйте модель, що розподіляє подані нижче слова за певними лінгвістичними ознаками:

гора, хлопчик, дівчина, машина, дерево, трава, залізо, старий, кошеня, горбик.

Потрібно виокремити не менше п'яти таких ознак. З'ясуйте, чи існують відношення між окремими ознаками (наприклад, одна ознака є найбільш загальною, а інша – вужчою).

*2. Побудуйте модель у вигляді семантичної мережі, що відображає спільні та відмінні елементи значень для слів **літак** і **птиця***



Овали позначають елементи значення, а стрілки – відношення між елементами. Відношення можуть бути різними, наприклад:

- елемент класу (належить до класу);
- атрибутивні зв'язки (має, використовує, вміє тощо);
- значення властивості (має властивість чогось);
- зв'язок типу «частина – ціле»;
- функціональні зв'язки (виробляє, впливає);
- кількісні (більше, менше, дорівнює);
- просторові (далеко від, близько від, над);
- часові (раніше, пізніше, протягом);
- логічні зв'язки (і, або, не) тощо.

МОДУЛЬ 4 ФОРМАЛЬНІ ГРАМАТИКИ І СЕМАНТИКИ

ТЕМА 10 ФОРМАЛЬНІ МОВИ І ФОРМАЛЬНІ ГРАМАТИКИ

План заняття:

1. Поняття формальної мови та формальної граматики
2. Породжувальна граMATика як один із типів формальних граматик
3. Передумови виникнення породжувальної граматики

Цей розділ присвячено формальним мовам, тобто природним мовам, які описані з використанням мови математики й логіки. Функціонування формальної мови забезпечується за допомогою формальної граматики, набору правил утворення ланцюжків, що можуть мати вигляд певних синтаксичних конструкцій. Як було зазначено раніше, теорія формальних граматик Ноама Хомського має в основі припущення, що в мозку людині наявний «лінгвістичний механізм», який відповідає за опанування мови. Деякі положення цієї теорії на сьогодні залишаються актуальними та використовуються у розв'язанні завдань прикладної та комп'ютерної лінгвістики.

10.1. Поняття формальної мови та формальної граматики

Мова, у першу чергу – це система позначень і правил для передавання повідомлень. Розрізняють мови природні і штучні. Прикладом штучних мов є не лише мови міжнародного спілкування на кшталт Есперанто, Інтерлінгва, Волапюк тощо, а й інші системи, які використовуються як засіб комунікації, такі як азбука Морзе, елементами якої є символи {., –} або двійкова система для передавання інформації у комп'ютерах, що послуговується символами {0, 1}. Різновидом штучних є формальні мови, до числа яких належать мови програмування. Формальна модель, що репрезентує природну мову, характеризується узагальненістю (абстракцією) та чіткістю суджень. Розглянемо детально, що являють собою ці характеристики.

Узагальнення або абстрагування – *це результат розумової діяльності, який відображає загальні ознаки або якості тих чи інших об'єктів або явищ дійсності. Якщо мова йде про математичні об'єкти (число, множина, функція),*

то їх утворення відбувається шляхом узагальнення реальних об'єктів. Вивчаючи який-небудь об'єкт, математик помічає лише деякі його властивості, а від решти відволікається. Наприклад, абстрактний математичний об'єкт «число» може насправді позначати кількість дерев у саду або кількість молекул у краплі води. Але головним є те, щоб про дерева або молекули води можна було говорити як про елементи сукупності. Така «ідеалізація» реальних об'єктів передбачає одну важливу властивість: математика часто оперує нескінченними сукупностями, тоді як у реальності цих сукупностей не існує. Якщо говорити про систему мови як ідеалізований об'єкт, то це може бути не лише реальна мова, а й та, яка навіть не існує в природі. Головним є те, щоб ця система мови забезпечувала передачу інформації, використовуючи відповідні алгоритми (правила) перетворення даних та певні формати їх репрезентації. Прикладом неіснуючої мови може бути та, яка породжує речення на кшталт «Глокая куздра штеко будланула бокра и курдячит бокрёнка»³¹. Побудоване системою мови речення репрезентує зміст повідомлення у такому форматі:

$$Subj_{[char]} + [Act_{[char]} + Obj_{[char]}] + [Act + Obj_{[char]}],$$

тобто: «певний суб'єкт з відповідними характеристиками виконав певним способом дію по відношенню до суб'єкта з певними характеристиками, а потім почав виконувати дію по відношенню до другого об'єкту зі своїми характеристиками». У Таблиці 10.1 наведено масив характеристик ($[char]$) для кожного учасника комунікативної ситуації, що репрезентована аналізованим реченням.

Таблиця 10.1. Характеристики учасників комунікації.

$Subj_{[char]}$	Значення характеристик
Age	Adult
Gender	Female
Attribute	Глокая
$Act_{[char]}$ (1)	Значення характеристик
Status	Finished
Time	Past

³¹Цю фразу запропонував Лев Щерба і використовував на вступних лекціях із загального мовознавства у 1928 р.: https://en.wikipedia.org/wiki/Glokaya_kuzdra

Manner	Штеко
$Obj_{[char]}(1)$	Значення характеристик
Age	Adult
Gender	Masculine
$Act_{[char]}(2)$	Значення характеристик
Status	Continuing
Time	Present
$Obj_{[char]}(2)$	Значення характеристик
Age	Young
Gender	Masculine

Це один із прикладів відображення в абстрактному вигляді комунікативної функції мови, що репрезентує у повідомленні тих учасників комунікації, які можуть не існувати в реальності, та наділяти їх певними характеристиками. Перейдемо до розгляду іншої властивості формальної моделі – чіткість суджень. У математиці цей критерій перевірки гіпотез на істинність відбувається за допомогою чітких міркувань, що підкоряються певним правилам.

Формальною мовою називають *математичну модель реальної мови*, при цьому під реальною мовою розуміють певний спосіб комунікації суб'єктів між собою. Для спілкування суб'єкти використовують кінцевий набір знаків (символів), що вимовляються або виписуються в чіткому часовому порядку, інакше кажучи утворюють лінійні послідовності у вигляді слів або речень. І в цьому відношенні розглядається лише комунікативна функція мови, яка вивчається та моделюється з використанням математичних методів, а інші функції до уваги не беруться. Формальні мови створюються для розв'язання конкретних проблем за допомогою комп'ютера (як простий приклад, мовою програмування можна розв'язати завдання із визначення коренів квадратного рівняння). У сучасній структурній лінгвістиці науковці використовують формальні мови для моделювання тих чи інших лінгвістичних явищ³², що існують у системі природної мови, і застосовують для цього математичні

³² Об'єктом класичної теорії формальних мов є синтаксис, математична модель якого описує механізми породження і розпізнавання «правильно побудованих» ланцюжків.

методи. Але спочатку необхідно виокремити з мови лише ті її властивості, які є важливими для вивчення, а потім ці властивості чітко визначити. Отримана таким чином абстракція називатиметься формальною мовою – математичною моделлю реальної мови (див. рис. 10.1). Зміст конкретної математичної моделі залежить від того, які властивості планується наразі виділити і вивчати.

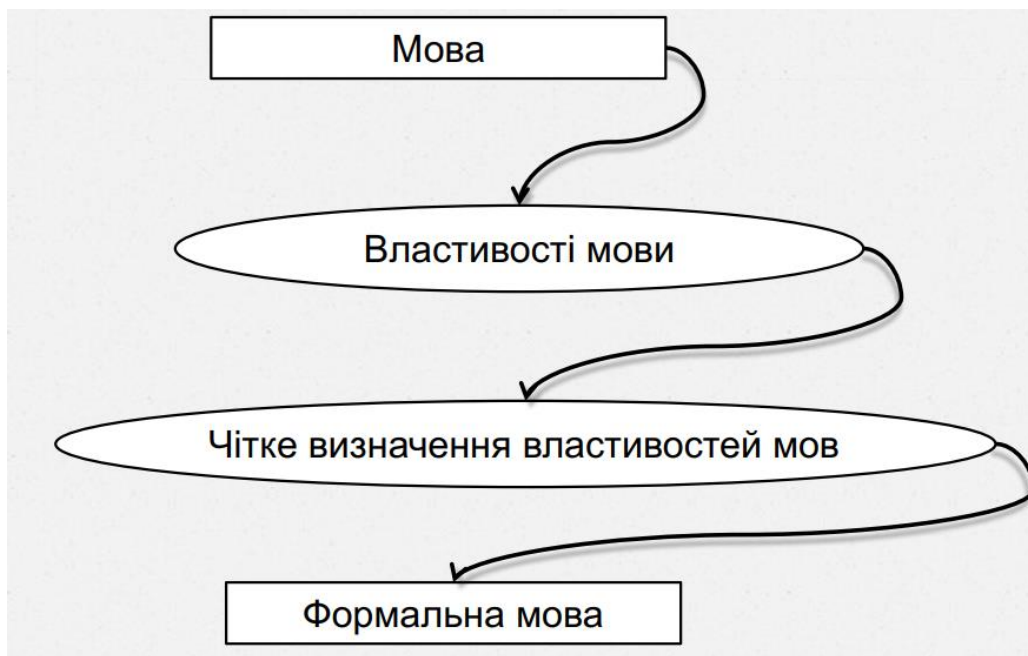


Рис. 10.1. Схема створення формальної мови.

Нехай існує деякий алфавіт, який умовно позначено X . Словом (ab , bc , abc тощо) в алфавіті X називається послідовність символів цього алфавіту. Набір допустимих слів становить лексику мови W . Слова можуть об'єднуватися в складніші конструкції – речення S , яких може бути безліч. Речення будуються зі слів і словосполучень за правилами синтаксису, який являє собою опис правильних речень³³. Алфавіт, лексика і синтаксис повністю визначають набір припустимих конструкцій мови та внутрішні взаємозв'язки між ними. Набір правил синтаксису утворює граматику мови L . Правила синтаксису можуть описувати або процедуру отримання правильних речень, або процедуру розпізнавання «правильності» речень (тобто їхньої приналежності даній мові).

³³ Правильне речення означає побудоване за заданими правилами граматики. Але граматично правильний не завжди означає семантично коректний, пор., приклад речення за Ноамом Хомським: *Colourless green ideas sleep furiously* (див. за посиланням https://en.wikipedia.org/wiki/Colorless_green_ideas_sleep_furiously).

У першому випадку мова йде про породжувальну граматику, а в другому – про розпізнавальну.

Але в теорії формальних мов видається важливим вивчити закони розташування слів поруч одне з одним, тобто синтаксичні властивості текстів. Тому формальна мова задається як множина послідовностей, складених з елементів скінченного алфавіту, як це описано нижче (див. рис. 10.2).

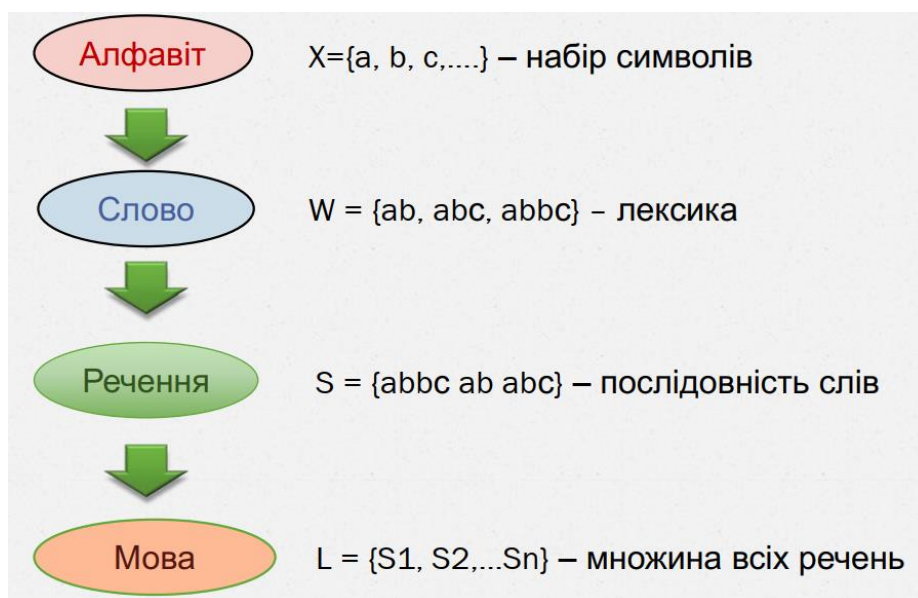


Рис. 10.2. Формальна мова та її складники.

Розглянемо, як приклад, формальну мову, яка являє собою множину коректних математичних формул, що включають у себе натуральні числа, дужки і знаки арифметичних дій. Алфавітом цієї мови буде множина такого вигляду:

$$\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -, *, /, (,)\}$$

Наприклад, рядок $((2 + 3) * (5 + 7))$ буде словом такої мови, а рядок $+2(3 - 1($ не буде, хоч і складається з допустимих символів, тому що він містить очевидні помилки.

Як можна описати цю мову? Ми знаємо, що числа складаються з цифр і можуть об'єднуватися за допомогою знаків арифметичних дій у вирази, які треба поміщати в дужки: $(2 + 3)$. Також можна додавати, віднімати, множити й ділити – тобто об'єднувати у складніші формули. Своєю чергою, вирази можуть об'єднуватися у формули. Одним із прикладів формули є $((1 + 2) * ($

$(5 - 3)$), яку складають два простих вирази $(1 + 2)$ та $(5 - 3)$. І це можна записати у вигляді такого символного запису: ((ВИРАЗ) ЗНАК (ВИРАЗ)). При цьому замість ФОРМУЛА можна використати будь-який простий або складний вираз, а замість ЗНАК – будь-який арифметичний знак. Вираз $(2 + 2)$ матиме такий символний запис: (ЧИСЛО ЗНАК ЧИСЛО).

Побудова ланцюжків у вигляді математичних формул відбувається за встановленими правилами. Результатом їх послідовного використання є отримання із запису ((ВИРАЗ) ЗНАК (ВИРАЗ)) конкретної математичної формули типу $((1 + 2) * (5 - 3))$. Під час використання правил відбувається заміна елемента запису на інший або на символ із алфавіту Σ . Отже, правило складається з лівої та правої частин. У лівій вказано елемент, що підлягає заміні, а в правій – елемент, що підставляється замість того, що в лівій частині. Так, наприклад, правило ЦИФРА $\Rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$ означає, що елемент ЦИФРА можна замінити на будь-яку цифру від 0 до 9. Наведемо інший приклад: ЧИСЛО $\Rightarrow$ ЦИФРА. Це правило означає, що замість елемента ЧИСЛО можна підставити елемент ЦИФРА. Уведемо такі правила формальної мови:

1. ФОРМУЛА $\Rightarrow ((\text{ВИРАЗ}) \text{ЗНАК} (\text{ВИРАЗ}))$
2. ФОРМУЛА $\Rightarrow$ ЧИСЛО
3. ФОРМУЛА $\Rightarrow$ ВИРАЗ
4. ЗНАК $\Rightarrow + \mid - \mid * \mid /$
5. ЧИСЛО $\Rightarrow$ ЦИФРА $\mid$ ЧИСЛО ЦИФРА
6. ЦИФРА $\Rightarrow 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$
7. ВИРАЗ $\Rightarrow (\text{ЧИСЛО} \text{ЗНАК} \text{ЧИСЛО})$.

Так, правило (1) потрібно читати так: «замість елемента ФОРМУЛА можна підставити ланцюжок ((ВИРАЗ) ЗНАК (ВИРАЗ))». Аналогічним чином читають усі інші правила. Одне число це також коректна формула, і це можна подати у вигляді запису (2). Для того, щоб узяти формулу в дужки, передбачено правило (3). За правилом (5), числом може бути як цифра, так і комбінація цифр. Якщо

до числа додати праворуч ще одну цифру, то ми знову отримаємо число. Варто зазначити, що запис (4) можна було записати у вигляді кількох правил:

ЗНАК $\Rightarrow +$

ЗНАК $\Rightarrow -$

ЗНАК $\Rightarrow *$

ЗНАК $\Rightarrow /$

Але задля скорочення запису в один рядок використовують символ «|», який не входить до алфавіту Σ формальної мови. Аналогічним чином записано правила (5) і (6).

Наведемо приклад породження формули $(12 + 5)$, що передбачає послідовне використання правил (3) – (7) – (4) – (5) – (6) – (5) – (5) – (6) – (6):

ФОРМУЛА $\Rightarrow$ ВИРАЗ $\Rightarrow$ (ЧИСЛО ЗНАК ЧИСЛО) $\Rightarrow$ (ЧИСЛО + ЧИСЛО) $\Rightarrow$ (ЧИСЛО + ЦИФРА) $\Rightarrow$ (ЧИСЛО + 5) $\Rightarrow$ (ЧИСЛО ЦИФРА + 5) $\Rightarrow$ (ЦИФРА ЦИФРА + 5) $\Rightarrow$ (1 ЦИФРА + 5) $\Rightarrow$ $(12+5)$.

Перейдемо тепер до поняття **формальної граматики**. У широкому сенсі формальними граматами, або просто граматами, називають будь-які «автоматичні пристрої» у вигляді числень або алгоритмів, що дають змогу «задавати»³⁴ мови. За способом задання Ноам Хомський розрізняє три типи граматик: розпізнавальні, перелічувальні та породжувальні, мова про які йшла у підрозділі 9.2, коли розглядалася класифікація лінгвістичних моделей за об'єктом моделювання. Розпізнавальні граматики призначені для ідентифікації певного ланцюжка (речення) заданої мови, перелічувальні – для генерації нескінченної кількості ланцюжків, а породжувальні – для породження ланцюжків на запит користувача. В усіх трьох випадках граматика мови описує закони побудови ланцюжка, які зазвичай називають синтаксичними закономірностями. Тобто, можна сказати, що граматика надає теоретичну базу для побудови формального опису синтаксичних закономірностей мови, що

³⁴ **Задання мови** – це встановлення вихідних синтаксичних конструкцій та правил, за якими можна породжувати граматично правильні ланцюжки у вигляді речень або виразів.

уможлиблює створення алгоритму для опрацювання ланцюжків мови комп'ютером.

Одним із різновидів формальної граматики є породжувальна граматика, яка дає змогу виводити (породжувати) ланцюжки мови з деякого початкового ланцюжка за допомогою певних правил заміни (або підстановки), як це було показано на наведеному вище прикладі утворення математичних формул.

Породження являє собою *покроковий процес, у якому на кожному кроці з ланцюжка, вже отриманого на попередньому кроці (зокрема, з початкового), можна шляхом застосування до нього правил заміни отримати новий ланцюжок у вигляді речення.* Варто зазначити, що теорія формальних граматик у корні відрізняється від попередніх лінгвістичних теорій своєю метою і завданнями. Насамперед, ця теорія ставить за мету не збір, опис та узагальнення мовних даних. На думку її автора, Ноама Хомського, метою має стати відповідь на питання, яким чином людина засвоює мову. Іншими словами, теорія формальних граматик – це не теорія мови, а теорія засвоєння мови.

У підрозділі 10.2 детально буде розглянуто питання, що являє собою породжувальна граматика.

Контрольні запитання:

1. Що є предметом формальної теорії мов? Що означає побудувати ланцюжок мови?

2. Опишіть процес побудови ланцюжка мови. Коли можна вважати завершеним процес утворення ланцюжка?

3. Що являє собою правило формальної граматики?

Практичні завдання

1. Які з перелічених ланцюжків формальної мови можна отримати з літери D , застосовуючи будь-яке число разів правила 1 – 5:

bb, bbc, bbcd, bbcd, bd, cbd, bdc, b, cdb, cb, cc, ccd, ccdbb, ccccd?

1. $D \rightarrow AB$	3. $B \rightarrow b$	5. $D \rightarrow d$
2. $A \rightarrow CD$	4. $C \rightarrow c$	

2. Укажіть у таблиці, які з перелічених вище ланцюжків належать та які не належать формальній мові:

Належать формальній мові	Не належать формальній мові

3. Можна доповнити подані вище правила 1 – 5 правилами 6, 7 і 8 такого вигляду:

b → дієслово третьої особи однини теперішнього часу

c → прикметник чоловічого роду однини

d → іменник чоловічого роду однини

Побудувати наведені нижче фрази, використовуючи правила 1 – 8:

Молодий хлопець співає

Молодий вродливий хлопець співає, танцює

Сильний молодий вродливий хлопець співає, танцює, сміється.

10.2. Породжувальна граматики як один із типів формальних граматики

Породжувальна граматики, як і інші типи граматики, відображає будову і принцип функціонування системи мови у вигляді формальної моделі (див. підрозділ 4.2). Слід уточнити, що функціонування системи мови у теорії породжувальних граматики розглядається у синтаксичному аспекті та, за задумом Ноама Хомського, встановлює правила, за допомогою яких мовці породжують речення. Породжувальна граматики ґрунтується на положенні про те, що володіння мовою має два складники: 1) компетенцію (англ. *competence*), тобто знання сукупності правил, за якими будують речення, та 2) використання (англ. *performance*), інакше кажучи вміння будувати речення, враховуючи позамовні чинники, такі як ситуація, стиль мовлення, характер взаємостосунків зі співрозмовником тощо. **Мовна компетенція** за Ноамом Хомським – це система уявлень про граматику мови, що притаманні людині від народження, інакше кажучи, властиві людині як біологічному виду незалежно від її досвіду, місця існування тощо. Говорячи про породжувальну граматику, учений уявляє її як систему правил, яка деяким експліцитним і добре визначеним чином приписує реченням структурні описи. Граматики сама по собі не є моделлю для

того, хто говорить або слухає, вона прагне охарактеризувати в найбільш нейтральних термінах знання мови, що дає основу для дійсного використання мови мовцем-слухачем. У цьому плані вона лише незначною мірою враховує ситуативні умови використання мови – її правила призначені для опису мовної компетенції, що давало підстави для критики теорії Ноама Хомського. Зокрема, серйозними недоліками теорії вважалися постулат про вроджений характер мовної здібності, недостатня увага живому мовленню з усіма її можливими відхиленнями від правил граматики тощо. Однак, варто зазначити, що дослідження, які проводилися в рамках породжувальної граматики, зробили вагомий внесок у розвиток сучасної лінгвістики. Так, якщо до Ноама Хомського переважав емпіричний аналіз розрізнених фактів, що належали різним мовам, то теорія породжувальних граматик змусила лінгвістів звернутися до дослідження питань про те, як саме влаштована мова загалом, як діє її механізм у процесах породження та розуміння мовлення. Це також сприяло вдосконаленню методики дослідження мовного матеріалу. Тобто, для цього стали залучатися не тільки власне лінгвістичні методи, а й методи інших наук, зокрема формальної логіки. Це дало можливість глибше проникнути в механізм взаємодії словесних значень у межах речення. Теорія породжувальних граматик детально висвітлена у роботі Ноама Хомського «Синтаксичні структури» (1957 р.).

Перейдемо до розгляду питання, як саме була побудована породжувальна граматика. Перші породжувальні (а точніше, трансформаційно-породжувальні) граматики склалися з двох компонентів – базового та трансформаційного. Перший включав:

1) набір синтаксичних символів на позначення самого речення (*S*) та його окремих елементів (*NP* – група підмета, *VP* – група присудка, *DET* – детермінант тощо);

2) необмежену множину слів (*собака, кішка, людина, рудий, кудлата, бігти, їсти* тощо) та окремих морфем (наприклад: *-є* – закінчення третьої особи

однини теперішнього часу, *-ли* – закінчення третьої особи множини минулого часу);

3) набір правил (правила бази), за якими відбувається побудова речення з елементів (1) і (2) та які являють собою правила підстановки одного елемента іншим: $S \rightarrow NP + VP$ (тобто речення складається з групи підмета і групи присудка); $NP \rightarrow A + N$ (своєю чергою, група підмета складається з іменника і прикметника).

Набір правил бази дає змогу описати глибинну структуру речення і створити на її основі множину граматично правильних речень (поверхневих структур). Перехід від глибинної структури до поверхневої відбувається шляхом послідовного використання правил підстановки із заданого набору. Так, елемент S задає глибинну структуру речення $NP + VP$, після чого замість групи підмета NP вводиться сполучення іменника і прикметника $A + N$, а після – замість прикметника A підставляється слово *big*, а іменника N – *brother*. Аналогічним чином група присудка розгортається в дієслово та, наприклад, займенника: $V + PRON$, а після цього виконуються підстановки $V \rightarrow \textit{cares}$ та $PRON \rightarrow \textit{you}$. У результаті ми отримуємо поверхневу структуру речення: *Big brother cares you*.

Другий компонент породжувальної граматики – трансформаційний – містить правила розташування слів у реченні (трансформаційні правила), до яких належать правила узгодження, керування, додавання певного елемента речення (див. підрозділ 6.3). Ноам Хомський вводить у цей компонент правила перестановки, додавання, вилучення та заміни елементів речення, а також правил перетворення одних синтаксичних конструкцій в інші. Одним із прикладів перестановки є перетворення стверджувального речення у питальне: *A unicorn is in the garden* $\rightarrow$ *Is a unicorn in the garden?* Тобто, переглянувши стверджувальне речення зліва направо, мовець переставляє дієслово-зв'язку *is* на початок. Додавання елементів у структуру речення – коли потрібно, наприклад, змінити мету висловлення, але при цьому зберегти змістову основу: *батько прийшов* – *батько не прийшов* – *чи батько прийшов* (запитання). **Еліпс**

– вилучення елемента зі складу речення, тобто прийом «згортання речення»: *Там на півдні зараз тепло – Там зараз тепло*. **Пасивізація** – створення пасивних синтаксичних конструкцій: *John hit the ball → The ball was hit by John*. **Негація** – побудова заперечувального речення: *John hit the ball → John did not hit the ball*. **Субституція** – заміна одного елемента речення іншим: *John hit the ball → He hit the ball*. Усе це разом, за теорією породжувальних граматик, становить вроджену здатність до використання мови, тобто здатності будувати граматично правильні речення.

Працею Хомського, що розглядає питання семантики та її зв'язків із синтаксисом, є «Аспекти теорії синтаксису» (1965 р). У ній вперше було зроблено спробу ввести семантичний компонент у граматику. Під семантичним компонентом учений розуміє сукупність правил семантичної інтерпретації, що приписують значення глибинним структурам. Знання мови включає в себе імпліцитну здатність розуміти необмежену кількість речень. З огляду на це породжувальна граматика має бути системою правил, які можуть ітерувати³⁵ і породжувати при цьому нескінченно велику кількість структур. За Ноамом Хомським, цю систему правил можна розкласти на три основні компоненти породжувальної граматики: синтаксичний, фонологічний та семантичний. Останні два мають виконувати суто інтерпретаційну функцію, тобто вони не беруть участі у породженні структур речень. Синтаксичний компонент вказує для кожного речення **глибинну структуру**, що визначає його семантичну інтерпретацію, та **поверхневу структуру**, що надає йому фонетичну інтерпретацію. Перша опрацьовується семантичним компонентом, друга – фонологічним. Глибинна структура речення набуває семантичної (змістової) інтерпретації завдяки семантичному компоненту. На вхід до фонологічного компоненту надходить поверхнева структура, яка на виході отримує фонологічну інтерпретацію. Кінцевим результатом «роботи» граматики стає співвіднесення семантичної інтерпретації з фонетичною. Опосередкованою ланкою в цьому співвіднесенні є синтаксичний компонент граматики, який

³⁵ У цьому контексті **ітерація** означає повторне використання правил граматики.

здійснює побудову поверхневої структури, тобто конкретного речення (див. рис. 10.3).

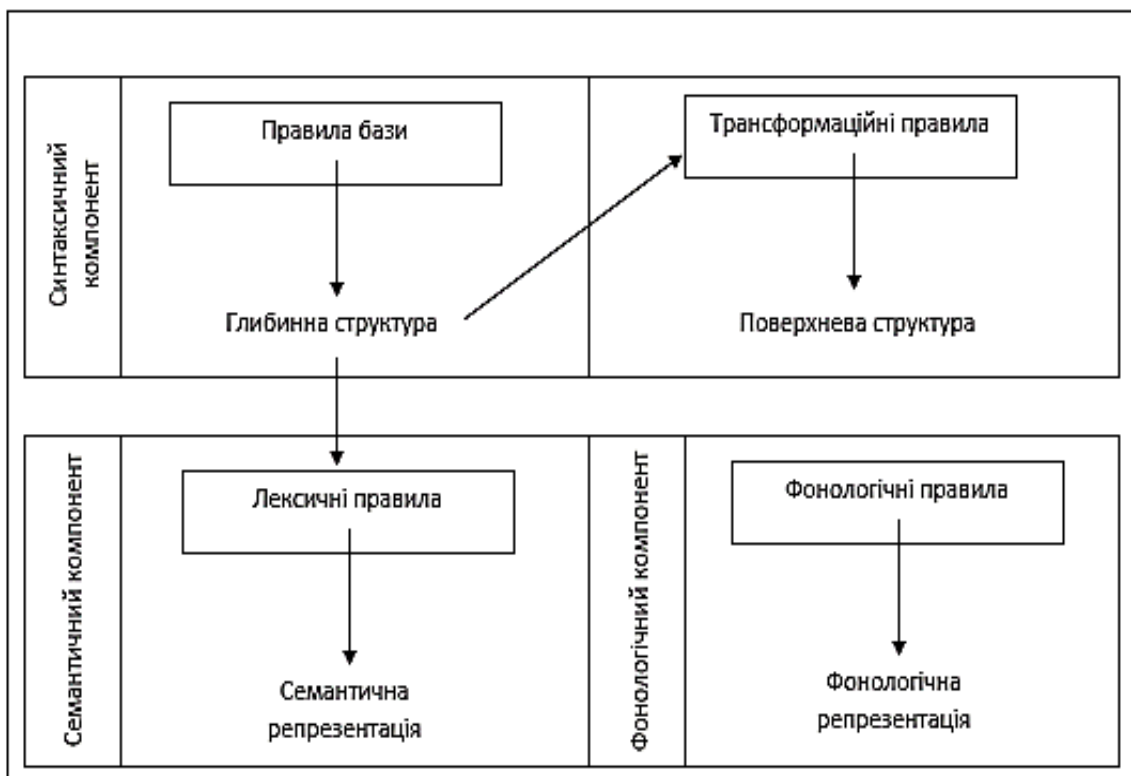


Рис. 10.3. Схема будови трансформаційно-породжувальної граматики.

Роль семантичного компонента породжувальної граматики висвітлена у роботі Ноама Хомського «Аспекти теорії синтаксису». У ній **глибинні структури** визначаються як *структури, породжувані базовим компонентом*. Семантична інтерпретація речення залежить лише від його лексичних одиниць і граматичних функцій та відносин, репрезентованих у таких структурах. Глибинна структура потрапляє на вхід семантичного компонента й отримує семантичну інтерпретацію, яка потім перетворюється відповідно до правил трансформації у поверхневу структуру, де правила фонологічного компонента дають фонетичну інтерпретацію.

Наведемо приклад, що демонструє роботу семантичного компонента породжувальної граматики. На малюнку 10.4 подано глибинну структуру речення «Хомський створив породжувальну граматику». Ця глибинна структура відображає семантичний зміст S , тоді як поверхнева структура S задає його фонетичну форму.

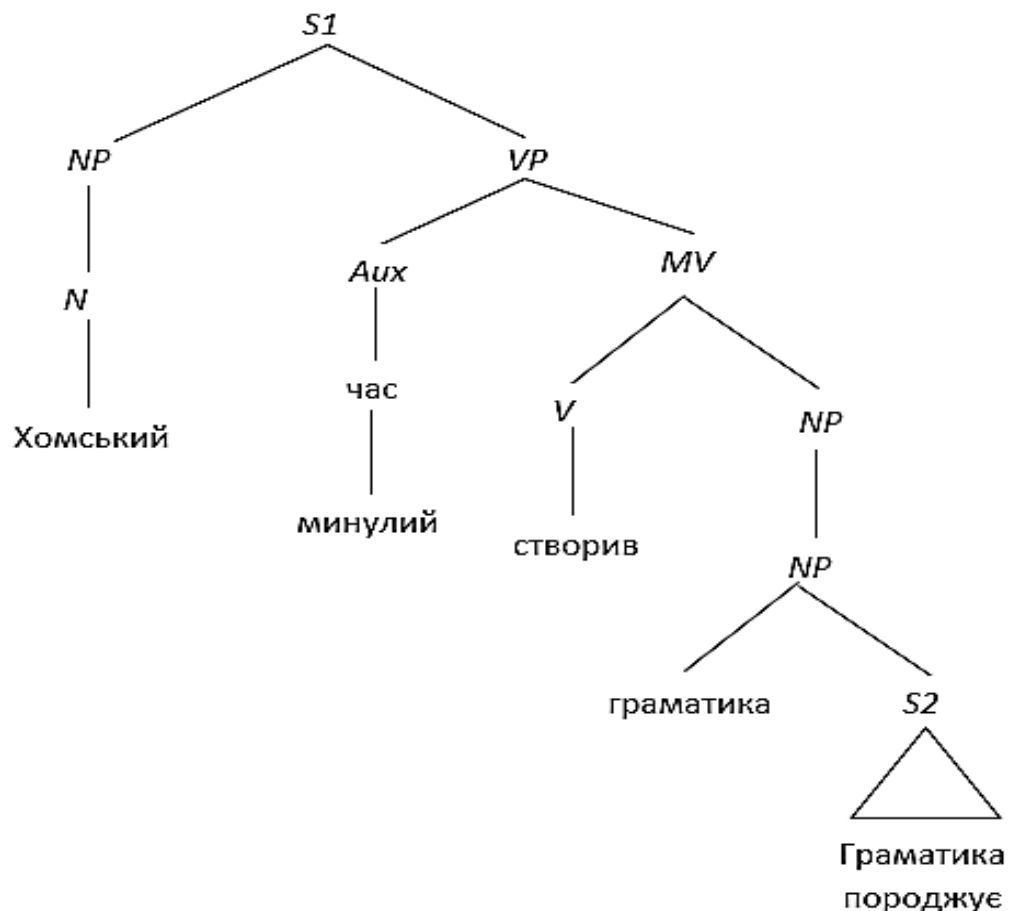


Рис. 10.4. Схема глибинної структури речення.

Для породження цього речення будуть використовуватися такі правила бази:

- | | |
|------------------------------|---|
| 1. $S1 \Rightarrow NP + VP$ | речення складається з групи підмета NP і групи присудка VP |
| 2. $NP \Rightarrow N$ | група підмета NP складається з іменника N |
| 3. $VP \Rightarrow Aux + MV$ | група присудка VP складається з допоміжної частини Aux та основного дієслова MV |
| 4. $MV \Rightarrow V + NP$ | основне дієслово MV складається з дієслова V та групи підмета NP |
| 5. $NP \Rightarrow NP + S2$ | група підмета NP розкладається на іменну групу NP та відносне речення S2 |

За наведеними правилами отримуємо глибинну структуру, яку можна подати у вигляді дерева безпосередніх складників (див. рис. 10.4). Правила бази застосовуються рекурсивно, поки не буде отримано ланцюжок із термінальних символів (див. підрозділ 10.4). У поданому прикладі термінальними символами глибинної структури є «Хомський», «створив», «граматику», «породжує». Так, наприклад, нетермінальний символ *N* правила (2) замінюється на термінальний

символ «Хомський». Групу підмета (5) складають група підмета «граматика» та відносне речення S2 «що (= граматики) породжує». Ця конструкція у поверхневій структурі відображається як атрибутивне словосполучення. Конструкція S2 розкладається аналогічним чином, що й речення S1. Глибинна структура S1 набуває інтерпретації семантичним компонентом, тобто термінальні елементи лексикалізуються – набувають статус слів із певним лексичним значенням та відповідними обмеженнями у сполучуваності. Зазначена лексична інформація зберігається у словникових статтях словника породжувальної граматики.

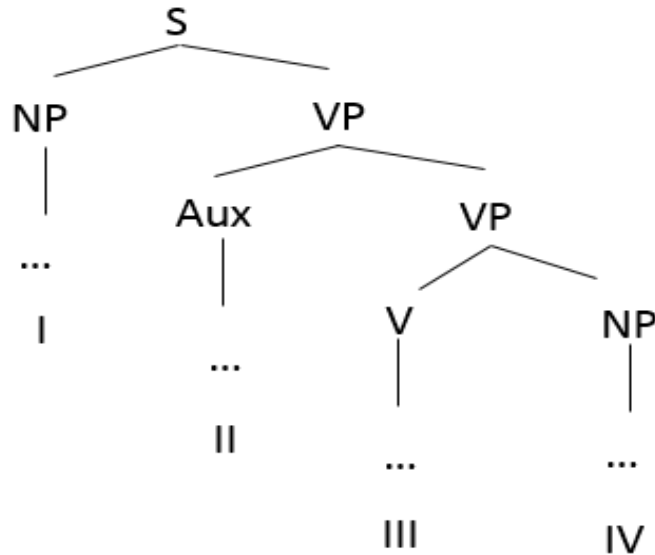
Трансформаційний компонент породжує поверхневі структури зі структур, отриманих у результаті дії базових правил (наприклад, «Хомський, котрий створив породжувальну граматику», «Хомський створив граматику, що породжує», або «Хомський, котрий створив породжувальну граматику, що породжує»). Якщо глибинна структура складається із системи речень, які вставлені одне в одне, то трансформаційні правила застосовуються циклічно, починаючи з найглибше вставлених речень і закінчуючи головним реченням. Після опрацювання трансформаційним компонентом, речення надходить до фонологічного компонента, що забезпечує фонетичне оформлення речення у вигляді ланцюжка фонетичних символів. На цьому етапі процес побудови речення завершується.

Контрольні запитання:

1. Схарактеризуйте призначення теорії породжувальної граматики Ноама Хомського.
2. Що дала теорія породжувальних граматик сучасній лінгвістиці?
3. Які слабкі сторони можна виокремити у теорії породжувальних граматик?
4. За допомогою чого забезпечується перехід від глибинної до поверхневої структури?
5. Що є мовною компетенцією та мовною діяльністю у розумінні Ноама Хомського?

Практичні завдання

1. Доповніть глибинну структуру термінальними та нетермінальними елементами в I, II, III та IV таким чином, щоб можна було побудувати наведені нижче речення:



- (1). Any specialist should follow the instructions.
- (2). The patient was given first aid.
- (3). A grey horse chewed hay.
- (4). My pilots have passed medical tests.

2. Нижче подано англійські речення, для яких потрібно побудувати систему правил бази і подати його у вигляді ланцюжка:

- (1). I work.
- (2). Try pressing “Return” in this window.
- (3). Abrams became a manager.
- (4). Abrams will interview Browne by Monday.
- (5). Kim knows when to sleep.

Правила бази складаються з лівої та правої частин. Ліва містить нетермінальний елемент, який може бути замінено іншими, застосовуючи інше правило. Праву частину можуть скласти як термінальні, так і нетермінальні елементи.

3. Побудуйте схеми глибинної структури для речень із завдання 2 у вигляді дерева безпосередніх складників.

10.3. Історичний погляд на формування універсальної граматики

Як вже було обговорено вище, створення породжувальних граматик пов'язано з іменем Ноама Хомського – центральної фігури в лінгвістиці³⁶. У своїх напрацюваннях універсальної граматики учений спирався на ідеї та розвідки своїх попередників, які розглядатимуться нижче.

Варто зазначити, що історії мовознавства відомі ранні спроби створення універсальної граматики. Взагалі, що таке універсальна граматика? Чому вона так називається? **Універсальна граматика** – термін, яким у ряді лінгвістичних теорій позначається передбачуваний *набір правил чи принципів, властивих будь-якій природній мові*. Подібні правила не визначають мову повністю: вони допускають значну варіантність, але певною мірою обмежують її. У сучасній когнітивній науці універсальна граматика розуміється як притаманна людині можливість засвоїти рідну мову.

Яскравим прикладом, що демонструє прагнення знайти універсальні категорії та принципи в мовах, є «**Граматика Пор-Рояля**». Повна її назва – «Всезагальна і раціональна граматика, що містить основи мистецтва мовлення, викладені ясно і природно; раціональні підстави того, що є спільним для всіх мов, як і головних відмінностей між ними; а також численні зауваження про французьку мову». Основу теорії, викладеної у «Граматичі Пор-Рояля», складає опис «операцій розуму», за допомогою яких носії мови сприймають, аналізують та осмислюють навколишню дійсність. На думку авторів «Граматики...», зазначені операції можна збагнути винятково через мову та виразити лише за її допомогою. Для виявлення універсальних моментів, обов'язково присутніх у природніх мовах, автори використовують переважно структури французької мови, залучаючи для порівняння приклади з давньогрецької, латинської, давньоєврейської, іспанської, італійської, англійської та німецької мов. Тим самим автори дотримуються індуктивного методу доказу. Однак індукція «Граматики...» вирізняється неповнотою, і тому отримані результати не можуть

³⁶ З біографією та науковими поглядами Ноама Хомського можна детально ознайомитися тут: <https://www.britannica.com/biography/Noam-Chomsky/Linguistics>.

претендувати на універсальність. «Граматика...» являє собою радше аналітичний опис французької мови у зіставленні з деякими іншими мовами. Тому у подальших напрацюваннях вчені виходили з того, що універсальна граматика ґрунтується на природі речей.

На залежності граматики від природи речей наголошує Мартін Дакійський у своїй роботі **«Про способи позначення»**³⁷: «...усі національні мови є граматично ідентичними. Це пояснюється тим, що вся граматика запозичена з речей... і як природа речей однакова для тих, хто розмовляє різними мовами, такими є і способи буття та способи розуміння; і, отже, способи позначення однакові, а отже, однакові способи граматичної побудови або мовлення. І тому вся граматика однієї мови однакова з граматикою в іншій мові».

Такого самого погляду дотримувався англійський фізик Ісаак Ньютон у своїй праці **«Про універсальну мову»**: «діалекти окремих мов сильно розрізняються настільки, що загальну мову неможливо вивести з них настільки правильно, як із природи самих речей, яка єдина для всіх народів і на основі якої вся мова була створена із самого спочатку»³⁸.

Усі наведені вище ідеї лягли в основу наукових розвідок Ноама Хомського. Свою першу знакову роботу, яка називалася **«Синтаксичні структури»**, Хомський опублікував ще 1957 року та сформулював у ній свою теорію універсальної граматики. Спираючись на попередні дослідження, Хомський запропонував певну структуру, що регулювала побудову речень у природній мові, тобто універсальну граматику. Це те, що об'єднує всі існуючі мови – набір правил і структурних блоків, завдяки яким можливий переклад з однієї мови іншою, а також швидке засвоєння будь-якої мови в дитячому віці. Мова – це не слова, а структурна ієрархія. Основа мови – сукупність синтаксичних правил, за якими будується будь-яке висловлювання. Хомський

³⁷ <https://www.cambridge.org/core/books/abs/syntactic-theory-in-the-high-middle-ages/modistic-grammar/6A63E9B9E1FA2D1C8991C217853A3C36>

³⁸ Newton I. Of an Universall Language//Modern Language Review (переклад українською мовою цього фрагменту виконано Володимиром Широковим у праці «Лінгвістично-інформаційні студії...» Т.1. К.: Український мовно-інформаційний фонд НАН України. С. 55. URL: https://movoznavstvo.org.ua/files/tom_1_B5_print.pdf).

припустив, що в усіх людей є вроджена мовна здатність – приблизно така сама, як здатність ходити на двох ногах, але ще більш унікальна.

Розглядати ідеї Хомського варто, відповідаючи на ті запитання про мову, які він сформулював:

- 1) як ми впізнаємо свою рідну мову?
- 2) яким чином маленька дитина поступово засвоює рідну мову і починає нею говорити?
- 3) як пояснити перехід від дитини, яка не знає мови, до дитини, яка знає мову?
- 4) крім того, що взагалі означає «знати мову»?

Уявімо собі папугу. Вона може вимовляти якісь речення, але при цьому важко говорити, що вона знає мову. Вона просто може вимовляти деякі речення. А що ж означає знання мови? Як проявляється знання мови? І як ми можемо сказати, що хтось знає мову? Хомський відштовхується від відповіді, яку вже було сформульовано до нього. Це так звана біхевіористська відповідь на те, як люди вивчають мову. З точки зору біхевіоризму, ми вивчаємо мову за допомогою стимулів і реакцій, які отримуємо із зовнішнього середовища. Тобто володіння мовою – це суто результат навчання. Але людина, як зауважив Хомський, уже в самому ранньому дитинстві здатна породжувати висловлювання, яких ніколи не чула. Припустимо, дитині кажуть: «мама мила раму». У свою чергу, дитина самостійно трансформує це речення та створює схожі висловлювання: «рама мила маму» або «мама мила маму» і так далі. Набір висловлювань, які можна породити природною мовою, потенційно нескінченний. Це безмежне поле можливостей просто не може бути результатом навчання.

Звідси випливає висновок: мова не є простим набором реакцій на подразнення, а людська здатність «отримувати необмежену кількість речень з обмеженої кількості слів» закладена у людині від початку, на генетичному рівні. А оскільки мов існує безліч, логічно припустити, що людині притаманне знання не якоїсь певної мови, а єдиних для всіх мов принципів і правил, які

Ноам Хомський і називає «універсальною граматикою». Сама думка, що мова є певною мірою вродженим і властивим тільки людині умінням, зрозуміло, не нова: про це писали Френсіс Бекон і Чарльз Дарвін.

Дитина отримує реакцію від свого оточення, тобто дорослий коригує її поведінку і мову, завдяки чому вона починає розуміти, як говорити правильно. Хомський поставив таке запитання: а чи достатньо стимульної інформації отримує дитина для вивчення рідної мови? Бо якщо помістити якусь іншу істоту, наприклад, мавпу в ті самі умови, в яких людина вивчає рідну мову, мавпа не заговорить. Навіть папугу доведеться тренувати набагато більше.

Ноам Хомський запропонував таке пояснення: мова – це деяка вроджена здатність. Людина, на відміну від інших істот, має у своєму мозку деяку структуру, яка дозволяє їй засвоювати мови. Саме завдяки наявності цього особливого «пристрою», дитина за короткий термін опановує таку складну систему як мова. Цю притаманну тільки людині систему і загальну структуру, за посередництвом яких відбувається засвоєння рідної мови, Хомський назвав універсальною граматикою. Отже, **універсальна граматика** – *це деяка система правил і принципів, за допомогою якої будуються осмислені речення будь-якої мови у світі.*

Крім того, Хомський сформулював поняття породжувальної граматики, про яку мова йшла у підрозділі 10.2. Його важливою тезою також є те, що будь-яка породжувальна граматика або принципи для кожної конкретної мови є похідними від певної загальної універсальної граматики. Тобто, під впливом певного середовища деякі універсальні принципи розвиваються у систему тієї чи іншої мови, якою користуються представники окремої культури або народу. І головний принцип теорії універсальної граматики полягає в тому, що за всією поверхневою відмінністю мов на планеті Земля стоїть певна загальна глибока схожість їхньої структури. З погляду породжувальної та універсальної граматики існують загальні правила і принципи, на яких будується всяка можлива мова взагалі.

Контрольні запитання:

1. У чому полягає основна ідея теорії породжувальних граматики Ноама Хомського?
2. За рахунок чого дитина за відносно короткий проміжок часу засвоює мову?
3. Які характерні помилки притаманні людині під час вивчення рідної мови?
4. Які недоліки у теоріях попередників Ноама Хомського ви можете навести?

Практичні завдання

1. Розглянемо слово «обличчя» у сенсі, в якому можна проілюструвати реченням: «Маруся втерла вологе обличчя рукавом». Опишемо це значення як сукупність семантичних ознак «передня частина (1) голови людини (2)». У цьому визначенні вказано на зв'язок «обличчя» з головою. Це виключає уживання слова в цьому значенні у реченні на кшталт «Обличчя руки було дуже брудним». Аналогічним чином пропишіть семантичні ознаки для слів:

голова, паркан, пустир, око.

Подайте речення, в яких враховано прописані семантичні ознаки.

2. Подано речення зі словом «ліс», що були згенеровані трансформаційно-породжувальною граматиною, які є граматично правильними, але в той самий час невірними з приводу семантики:

- (1). Сьогодні я йшов по лісу футбольного поля.
- (2). Зелений ліс трави літає недалеко від річки.
- (3). Поле – це ліс посеред міста.
- (4). Польовий ліс використовується для будівництва.

Пропишіть семантичні ознаки слова «ліс» для семантичного компонента трансформаційної граматики, які мають бути враховані під час побудови речень та не допускати побудову семантично невірних речень.

Для виконання завдання можна використовувати словник СУМ 11 за посиланням <http://sum.in.ua/> або словник «Горох» <https://goroh.pp.ua/>.

ТЕМА 11 ПОБУДОВА МОДЕЛЕЙ ПОРОДЖУВАЛЬНОЇ ГРАМАТИКИ

План заняття:

1. Елементи моделі породжувальної граматики
2. Класифікація породжувальних граматик
3. Види трансформацій в межах граматики

Породжувальні граматики становлять окремий напрям сучасної лінгвістики, завданням якого вважається моделювання інтелектуальних процесів, пов'язаних з використанням мови. З огляду на це мова розглядається як механізм, здатний формувати складні лінгвістичні об'єкти (речення) за певним шаблоном та відповідно до заданих правил. Зауважимо, що система і структура як невід'ємні атрибути мови розглядаються так, як це описано в підрозділі 5.3. Апарат теорії породжувальних граматик Ноама Хомського застосовують для опису як природних, так і штучних мов (наприклад мов програмування). У цьому розділі детально розглядаються питання побудови моделей породжувальної граматики різного типу. У центрі уваги перебуває типологія породжувальних граматик Ноама Хомського, і кожний аналізований тип передбачає свій підхід до створення моделі.

11.1. Елементи моделі породжувальної граматики

У 1956 р. Ноам Хомський запропонував модель породжувальної граматики, якою цілком зручно користатися для задання як природних, так і штучних мов. Одна з переваг цієї моделі полягає в тому, що можливо зіставити кожний породжуваний ланцюжок (висловлювання) з його синтаксичною структурою. За визначенням ученого, породжувальна граMATика (G) є четвіркою об'єктів³⁹:

$$G = \{S, P, T, N\}$$

Розглянемо детально кожний з цих елементів. Першим іде **початковий (базовий) символ S** , який задає вихідну структуру ланцюжка. У попередній темі ми розглядали початкову структуру типу $S \Rightarrow NP + VP$, за допомогою якої

³⁹ Приклади згенеровано за допомогою Інтернет-ресурсу English Resource Grammar, який доступний за посиланням: <http://erg.delph-in.net/logon>.

можна описати речення, що мають групу підмета та групу присудка, наприклад:

(1) *He worked for an hour.*

(2) *More programmers than engineers arrived.*

(3) *Abrams hired Browne.*

Проте існують інші вихідні структури, які можна задавати початковим символом породжувальної граматики, такі як $S \Rightarrow VP$ (а), $S \Rightarrow ADV + S$ (б), $S \Rightarrow PP + S$ (в), $S \Rightarrow V + VP$ (г) та інші:

(4а) *Try pressing this button!*

(5а) *List the manager of any consultant.*

(6а) *Don't hire a programmer.*

(7б) *Evidently Chiang showed Devito his office.*

(8в) *Hired without an interview, Devito became a manager.*

(9г) *Was Abrams interviewed?*

(10г) *Has Abrams been interviewing?*

Наступний елемент породжувальної граматики Хомського являє собою **множину «правил породження» P** або продукцій. Кожне правило складається з пари елементів (N , книга) і задає можливість заміни лівого елемента на правий під час породження ланцюжка. Саме тому, правила зазвичай записують у вигляді формул типу $N \Rightarrow \text{книга}$, вказуючи конкретно, що на що можна замінювати. Множина правил у граматиці має бути непорожньою і скінченною, і зазвичай позначається латинською літерою P . Для запису правил граматики можна використовувати символи двох алфавітів: алфавіту термінальних елементів (терміналів), як у прикладі (а), та алфавіту нетермінальних елементів (нетерміналів), як у прикладі (б):

(а) $N \Rightarrow PRON$, де N і $PRON$ – нетермінальні елементи

(б) $PRON \Rightarrow \text{він}$, де $PRON$ – нетермінальний елемент; він – термінальний елемент.

Третім елементом породжувальної граматики є **множина терміналів**, яку позначають латинською літерою T . Термін «термінальний» означає, що такий

елемент не може бути замінений на інший під час породження ланцюжка, бо подальший процес формування структури висловлювання буде перервано. (terminate). У свою чергу, **нетермінальні елементи N** – четвертий елемент породжувальної граматики – використовуються на проміжних етапах процесу породження. Призначенням **нетерміналів** є задання алгоритму породження ланцюжка, який може бути найрізноманітнішим і зазвичай залежить від типу граматики, де він використовується. Приклади нетермінальних елементів наведено у таблиці 11.1.

Таблиця 11.1. Нетермінальні елементи породжувальної граматики.

Нетермінальний елемент	Що означає
NP	група підмета
VP	група присудка
ADJ	прикметник
DET	детермінант
N	іменник
V	дієслово
PP	прийменникова група
P	прийменник

Перелік нетермінальних елементів граматики не вичерпується наведеною таблицею. Під час побудови ланцюжка (речення) вони позначають ті чи інші частини мови та підлягають заміні іншими термінальними або нетермінальними елементами. Так, наприклад, нетермінальний елемент N може бути замінений одним із наданих терміналів {книга, собака, людина, гриб}. Надамо деякі приклади правил, за допомогою яких можна утворити ті чи інші ланцюжки:

Приклад 1.

- 1) $S \Rightarrow NP$
- 2) $NP \Rightarrow N$ *lady*
- 3) $NP \Rightarrow DET + N$ *this lady*
- 4) $NP \Rightarrow ADJ + N$ *young lady*

lady + VP + PP => This + young + lady + V + N + PP => This + young + lady + knows + N + PP => This + young + lady + knows + me + PP => This + young + lady + knows + me + P + N => This + young + lady + knows + me + by + N => This + young + lady + knows + me + by + name.

Таким чином, граматика Хомського являє собою, фактично, кінцеве число правил, які можна інтерпретувати як «елемент x можна замінити на елемент y ». Це універсальний механізм породження одних символічних ланцюжків з інших, причому з одного й того самого ланцюжка можна породити кілька різних структур, іноді навіть нескінченну їх кількість. Якщо після довільного числа підстановок, виконаних відповідно до правил граматики, породжений ланцюжок складається лише з терміналів, то процес породження вважається завершеним, а сам ланцюжок є структурою мови (висловленням), яка сформована породжувальною граматикою Хомського.

Контрольні запитання:

1. Чи може бути мова нескінченною, тобто породжувати нескінченну кількість ланцюжків?
2. З яких елементів складається модель породжувальної граматики Хомського? Поясніть їхні функції.
3. Опишіть процес породження ланцюжка. Наведіть власні приклади.
4. Із чого складаються правила породжувальної граматики? Наведіть власні приклади.

Практичні завдання

1. Побудуйте механізм формування українського речення у вигляді формальної граматики, яка має породжувати висловлення типу:

«жінки + шукали + милих + братів».

Надайте зміст елементів породжувальної граматики, для того щоб вона формувала заданий ланцюжок:

- Початковий символ (або аксіома): S ;
- Словник нетермінальних елементів: N ;
- Словник термінальних елементів: T ;

– Опишіть правила, за якими має відбуватися побудова речень: *P*.

2. Вкажіть, яким чином з вихідної структури *S* відповідно до цих правил можна отримати ланцюжок:

ЖІНКИ + ШУКАЛИ + МИЛИХ + БРАТІВ.

Скількома способами це можна зробити? Наведіть відповідні послідовності перетворень.

3. Вкажіть, яким чином з вихідної структури *S* відповідно до цих правил можна отримати ланцюжок:

ЗЕЛЕНІ + КОНІ + МАХАЛИ + ГОЛОВАМИ + КОПИТАМИ + ХВОСТАМИ.

Скількома способами це можна зробити? Наведіть відповідні послідовності перетворень.

11.2. Типологія породжувальних граматик

Формальні граматиками, виходячи з опису моделі у підрозділі 11.1, – це абстрактні системи, що дають змогу за однаковими процедурами отримувати граматично правильні речення певної мови разом із описом їхньої структури. За допомогою граматик можливо розпізнавати або породжувати множини ланцюжків, які інтерпретуються зазвичай як множини правильних речень, та зіставляти ці множини з описом їхньої синтаксичної структури у вигляді дерев безпосередніх складників (див. підрозділ 4.1). Відповідно до цього, виокремлюються породжувальні та розпізнавальні граматиками (див. підрозділ 9.2). Надалі ми будемо робити акцент на типології саме породжувальних граматик.

Породжувальні граматиками можуть бути різними: залежно від виду правил побудови виокремлюють низку основних типів породжувальних граматик. Важливими для лінгвістичних додатків є граматиками складників (див. підрозділ 10.2), у яких на кожному кроці виведення замінюється лише один символ, тому в них із кожним виведенням асоціюється так зване дерево виведення. Це робить граматиками складників хорошим інструментом для опису природних та штучних мов. Для опису природних мов слід використовувати трансформаційні

граматики, в яких трансформуються не ланцюжки, а безпосередньо синтаксичні структури – дерева складників або дерева підпорядкування (див. підрозділ 4.1).

Ноам Хомський запропонував класифікацію формальних граматик, яка виокремлює 4 основні типи за способом формулювання правил: тип 0, тип 1 (контекстно-залежні), тип 2 (контекстно-вільні) та тип 3 (регулярні). Цю класифікацію також називають ієрархією Хомського (див. рис. 11.1), де правила формулюються за певними структурами. Тож розглянемо детально кожен із виокремлених типів, починаючи з граматики найпростішого типу (тип 3) та закінчимо універсальною граматикою (тип 0). Зазначимо, що граматики типів 3 – 1 використовують для штучних мов (мов програмування) та мають обмежені можливості щодо їх застосування до природних мов. Але з метою демонстрації принципу їхньої роботи та подальшого вивчення в межах дисципліни «Математична та комп'ютерна лінгвістика», у прикладах будуть використані речення природною мовою.

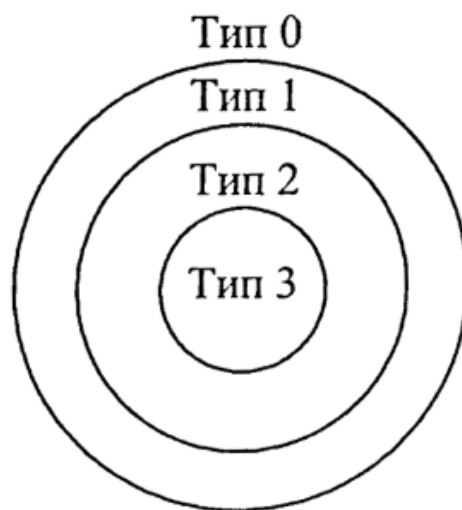


Рис. 11.1. Ієрархія Хомського.

Граматика типу 3 або регулярна граматика – *формальна граматика, у якій ліва частина правил містить один нетермінальний елемент, а права частина – або термінальний, або термінальний та нетермінальний елементи, або порожній рядок (ϵ).* Прикладом регулярної граматики може бути граматика, що породжує нашу «мову спілкування», наприклад, з тим чи іншим побутовим прибором, наприклад мікрохвильовою пічкою:

- (1) Oven $\Rightarrow$ power_on Do
- (2) Do $\Rightarrow$ meat_unfreeze Do
- (3) Do $\Rightarrow$ meat_fry Do
- (4) Do $\Rightarrow$ power_off

У поданому прикладі «power_on», «meat_unfreeze», «meat_fry» – термінальні елементи, тобто кнопки, що натискає користувач на панелі. Нетермінальними елементами граматики є Oven і Do, тобто свого роду синтаксичні конструкції примітивної мови програмування. При цьому Oven є вихідним елементом граматики, з якого починається породження речень мови. Наведемо приклад речення, що описує команду на включення пічки, розморожування та смаження м'яса:

Oven $\Rightarrow$ power_on Do $\Rightarrow$ power_on meat_unfreeze Do $\Rightarrow$ power_on meat_unfreeze meat_fry Do $\Rightarrow$ power_on meat_unfreeze meat_fry power_off.

Побудована граматика дає змогу породжувати речення (або команди) з виконання кількох операцій. Отже, у «мові спілкування» з мікрохвильовою пічкою, так само як із будь-яким іншим кінцевим автоматом, кількість можливих речень нескінченна. Незважаючи на всю простоту, примітивна мова, що є результатом роботи регулярної граматики, є доволі продуктивною, завдяки чому породжувана мова наближується до структури будь-якої природної мови, у якій кількість можливих речень також нескінченна.

Граматика типу 2 або контекстно-вільна граматика – породжувальна граматика, правила якої дозволяють замінити нетермінальний елемент у лівій частині на будь-яку комбінацію термінальних та нетермінальних елементів без контекстних обмежень. Одним із прикладів цього типу є граматика, яка будує будь-які математичні вирази (див. підрозділ 10.1). Для математичного виразу немає різниці у контексті або сенсі. Наведемо приклад⁴⁰ контекстно-вільної граматики, що може будувати англійські речення:

$$G = \{S, P, T, N\}$$

$$S = S;$$

⁴⁰ Приклад взято з: <https://people.cs.umass.edu/~mccallum/courses/inlp2007/lect5-cfg.pdf>

$T = \{\text{that, this, a, the, man, book, flight, meal, include, read, does}\};$

$N = \{S, NP, NOM, VP, Det, Noun, Verb, Aux\};$

$P:$

(1) $S \Rightarrow NP + VP$

(8) $VP \Rightarrow Verb$

(2) $S \Rightarrow Aux + NP + VP$

(9) $VP \Rightarrow Verb + NP$

(3) $S \Rightarrow VP$

(10) $Det \Rightarrow \text{that} \mid \text{this} \mid \text{a} \mid \text{the}$

(4) $S \Rightarrow \epsilon$

(11) $Noun \Rightarrow \text{book} \mid \text{flight} \mid \text{meal} \mid \text{man}$

(5) $NP \Rightarrow Det + NOM$

(12) $Aux \Rightarrow \text{does}$

(6) $NOM \Rightarrow Noun$

(13) $Verb \Rightarrow \text{books} \mid \text{includes} \mid \text{reads}$

(7) $NOM \Rightarrow Noun + NOM$

Перші чотири правила задають базову синтаксичну структуру, яку можуть скласти: група підмета і група присудка (1); або допоміжне дієслово та група підмета і група присудка (2); або лише група підмета (3); або порожній рядок (4). Базова структура S у цій граматиці – єдиний елемент граматики, що може бути порожнім. Побудуємо речення «The man reads this book» за поданим набором правил граматики:

$S \Rightarrow NP + VP \Rightarrow Det + NOM + VP \Rightarrow The + NOM + VP \Rightarrow The + man + VP \Rightarrow$
 $The + man + Verb + NP \Rightarrow The + man + read + NP \Rightarrow The + man + read + Det +$
 $NOM \Rightarrow The + man + read + this + NOM \Rightarrow The + man + read + this + Noun \Rightarrow The$
 $+ man + read + this + book.$

Слід зазначити, що замість $Noun$ може бути підставлений будь-який термінальний елемент $man, book, flight, meal$, який буде виступати у якості агента ($the\ man\ reads, the\ meal\ reads...$). Але ці речення будуть граматично правильними та будуть повністю відповідати правилам граматики типу 2.

Грамматика типу 1 або контекстно-залежна грамматика – породжувальна грамматика, у якій можливість заміни ланцюжка елементів може визначатися контекстом. Нехай існує породжувальна грамматика такого вигляду:

$$G = \{S, P, T, N\}$$

При цьому ліва частина одного або декількох правил може містити нетермінальні елементи, оточені «контекстом», тобто послідовністю елементів,

які в тому самому вигляді наявні у правій частині відповідних правил. При цьому сам нетермінальний елемент замінюється непорожньою послідовністю елементів у правій частині, а «контекст» фактично являє собою попередньо заданий термінальний елемент.

Наприклад, побудуємо контекстно-залежну граматику та сформулюємо для неї правила таким чином, щоб вона породжувала речення типу «батько миє машину»:

$$S = S$$

$$T = \{\text{батько, миє, машину}\}$$

$$N = \{\text{NP, VP, Verb, Sentence}\}$$

P :

$$(1) S \Rightarrow \text{NP} + \text{VP}$$

$$(2) \text{VP} \Rightarrow \text{Verb} + \text{NObj}$$

$$(3) \text{Verb} \Rightarrow \text{ContextVerb} [\text{миє}] \quad \text{контекст}$$

$$(4) \text{ContextVerb } P \Rightarrow \text{ContextVerb} [\text{миє}] + \text{NObj} \quad \text{NP} = \text{Subj} [\text{миє}]$$

$$\text{NObj} = [\text{миє}] \text{ Obj}$$

$$(5) \text{NP} \Rightarrow \text{NSubj}$$

$$(6) \text{NSubj} \Rightarrow \text{батько}$$

$$(7) \text{NObj} \Rightarrow \text{машину}$$

Прокоментуємо наведені вище правила граматики. Правило (1) демонструє базову структуру речення, що складається з групи підмета та групи присудка. Правило (2) задає структуру групи присудка і позначає, що певна дія відбувається над якимось об'єктом. Правило (3) вводить контекст, а саме дієслово *миє*, від якого залежатимуть інші елементи речення. Правило (4) конкретизує групу присудка відповідно до контексту, тобто дієслова *миє* та його керування. Правило (5) задає структуру групи підмета, яка має сполучатися із групою присудка. Правила (6) і (7) визначають заміну нетермінальних елементів термінальними.

Наведемо процедуру породження речення «батько миє машину» за допомогою правил контекстно-залежної граматики:

$S \Rightarrow NP + VP \Rightarrow NP + Verb + NObj \Rightarrow NP + \text{миє} + NObj \Rightarrow NSubj + \text{миє} + NObj \Rightarrow \text{батько} + \text{миє} + NObj \Rightarrow \text{батько} + \text{миє} + \text{машину}.$

Подана модель демонструє, що Ноам Хомський намагався ввести деякі елементи семантики у свою модель породжувальної граматики, але йому це неповністю вдалося, бо мови та граматики були розраховані на використання із специфічними обмеженими машинами (кінцеві автомати, машина Тьюрінга).

ГраMATика типу 0 – породжувальна граMATика, яка не накладає жодних обмежень на будову її правил. Цей класичний тип граматики можна застосовувати як до штучних (формальних), так і до природних мов. Приклади та детальний аналіз граматики типу 0 наведено у підрозділі 11.1.

Звичайно, що існують інші типи граматик. Наприклад: **домінаційна граMATика**, що породжує безліч ланцюжків, інтерпретованих як речення разом із їхніми синтаксичними структурами у вигляді дерев підпорядкування. **ГраMATика синтаксичних груп** породжує множину речень разом з їхніми синтаксичними структурами.

Існують інші концепції граматики – наприклад, **Δ -граMATика**, – що слугує не для породження речень, а для їх представлення у вигляді дерев: дерев підпорядкування або дерев складових. Як окремий тип слід відзначити **граматики Монтегю**, що призначені для одночасного опису синтаксичних та семантичних структур речення, використовуючи складний математико-логічний апарат. Такі граматики більш підходять для опису штучних мов, аніж природних.

Безумовно, теорія породжувальних граматик Ноама Хомського зробила великий внесок у лінгвістику, зокрема вона надала теоретико-методологічний апарат, що підвищує експліцитність лінгвістичного опису та сприяє розробці методик для створення моделей синтаксису. Крім цього завдяки теорії Хомського у лінгвістиці була сформована техніка формалізації опису мовного матеріалу для полегшення автоматизації процесів його опрацювання засобами комп'ютерних систем. Проте, у 1970-і рр. вплив теорії породжувальних граматик значно послаблюється через її слабкі сторони, зокрема, недооцінку

ролі семантичного компонента. Проблемі врахування семантичного компонента у побудові речень присвячена низка праць, таких як теорія «Смисл ↔ Текст», теорія семантики переваги та теорія концептуальної залежності, які розглядаються у темі 12.

Контрольні запитання:

1. За яким принципом визначається належність граматики до того чи іншого типу у ієрархії Хомського?
2. Чим відрізняється побудова правил у регулярній, контекстно-незалежній, контекстно-вільній та універсальній граматиках?
3. Яким чином задається вплив контексту на побудову речення у граматиці типу 1?

Практичні завдання

1. Побудуйте контекстно-вільну граматику, так щоб вона породжувала речення:

Mary saw Bob.

Mary ate an apple.

Did Mary eat an apple.

Mary.

Eat Mary.

Схарактеризуйте зміст елементів породжувальної граматики, для того щоб вона формувала речення:

- Початковий символ (або аксіома): S ;
- Словник нетермінальних елементів: N ;
- Словник термінальних елементів: V ;
- Опишіть правила, за якими має відбуватися побудова речень: P .

2. Побудуйте контекстно-залежну граматику та схарактеризуйте зміст її елементів:

- Початковий символ (або аксіома): S ;
- Словник нетермінальних елементів: N ;
- Словник термінальних елементів: V ;

– Опишіть правила, за якими має відбуватися побудова речень: *P*.

Граматику потрібно побудувати так, щоб вона могла формувати речення:

Собака їла м'ясо.

Дівчина написала листа.

Контекст у граматиці має бути прописаний таким чином, щоб не виходили речення типу:

М'ясо з'їло комаху.

Лист написав дівчину.

3. Побудуйте граматику типу 0, так щоб вона породжувала речення на кшталт:

Кішка спіймала мишу.

Схарактеризуйте зміст елементів цієї породжувальної граматики:

– Початковий символ (або аксіома): *S*;

– Словник нетермінальних елементів: *N*;

– Словник термінальних елементів: *V*;

– Опишіть правила, за якими має відбуватися побудова речень: *P*.

11.3. Трансформації в межах генеративної лінгвістики

Генеративна теорія спрямована на опис мовної здібності людини, яка проявляється опосередковано у мовленнєвій реалізації. Тому, завданням теорії є виявлення:

– глибинних ментальних структур;

– варіантів переходу від глибинних структур до поверхневих, які реалізовано у вигляді синтаксичних структур природної мови.

Виходячи з вищесказаного, генеративна теорія висуває на перший план синтаксис. При цьому основними проблемами лінгвістичного аналізу мови є: відокремлення граматично правильних послідовностей від граматично неправильних та подальше дослідження структури правильних послідовностей.

Будь-яка граматика аналізованої мови проектує випадкову сукупність спостережуваних висловлювань на множину граматичних висловлювань. З огляду на це граматика відображає поведінку носія мови, який на базі свого

мовного досвіду здатний відтворити та зрозуміти безкінечну кількість нових речень. Метод подання синтаксичної структури речення заснований на виведенні складних синтаксичних структур з більш простих за допомогою невеликого набору правил перетворення (трансформацій). Вихідним положенням є те, що різні мовні одиниці, які належать до однієї системи, можуть перетворюватися одна в одну за певними правилами. Трансформація – це формальна процедура (алгоритм), мета якої виявити змістовні особливості, глибинну семантику. Говорячи про трансформації, слід зазначити, що теорія оперує поняттями глибинної та поверхневої структури.

Глибина структура (ГС) – це спосіб абстрактного опису семантичної структури речення, абстрактна формула, утворена найбільш загальними, універсальними елементами змісту. Як правило, одна глибинна структура реалізується в декількох поверхневих. Значно рідше трапляються випадки, коли одній поверхневій структурі відповідають дві або більше глибинних. **Поверхнева структура (ПС)** – засіб вираження глибинної синтаксичної семантики граматичними класами слів, тобто частинами мови. Так, фраза (1) має дві глибинні семантичні структури (2) і (3):

- | | |
|-----------------|------------------------|
| (1) Його арешт. | (2) Він заарештував. |
| | (3) Його заарештували. |

Як приклад перетворення глибинної структури у поверхневу при породженні мови Хомський розглядає речення (4), яке, на його думку, складається з двох глибинних (5) і (6):

- | | |
|-------------------------|-------------------|
| (4) Мудра людина чесна. | (5) Людина мудра. |
| | (6) Людина чесна. |

Хомський вважав, що в основі механізму породження фраз лежить перетворення глибинних структур у поверхневі (див. рис. 11.2).

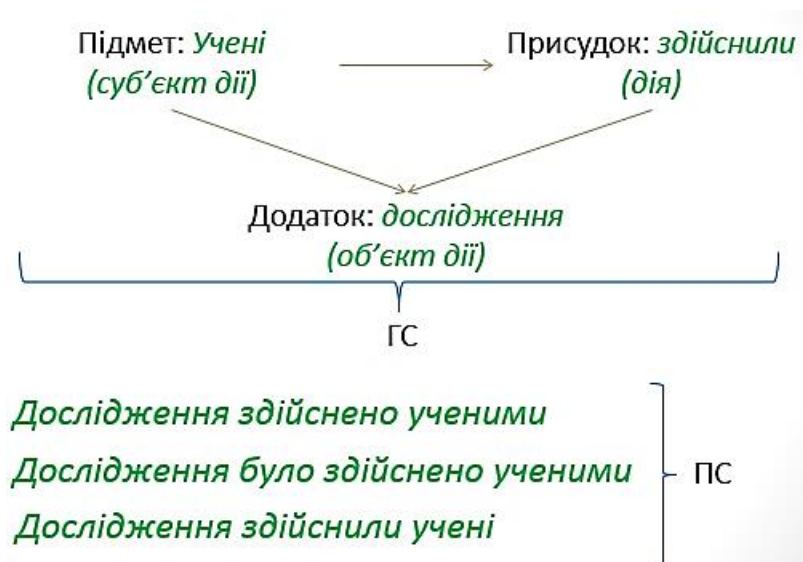


Рис. 11.2. Глибинна структура поверхневій структури.

Як було зазначено в темі 4, основу синтаксичної системи будь-якої мови утворює набір елементарних, ядерних або вихідних речень, які можуть відображати певні ситуації, наприклад:

- ситуація «порушення стану фізичної рівноваги», пор. *Вітер хитає (стрясє) будинок*;
- ситуація «спричинення руху об'єкта у напрямку до суб'єкта», пор. *Світло тягне метелика*;
- ситуація «порушення стану емоційної рівноваги», пор. *Син турбує (хвилює, засмучує, радує, турбує) мати*;
- ситуація «викликання негативних емоцій», пор. *Він злить (ображає, сердитий) мене* тощо.

Ядерними прийнято вважати елементарні, тобто прості стверджувальні речення, а усі інші речення є похідними. Ядерні речення складаються з дієслова у формі дійсного способу теперішнього часу активного стану, а також слів, що реалізують його валентності. **Валентність** – це здатність слова визначати кількість і якість залежних від нього словоформ, зумовлена його семантичними і граматичними властивостями. Так, властивості дієслова «керувати» передбачають його поєднання з іменними формами називного й орудного відмінка у позиціях підмета і додатка.

Будь-яке ядерне речення можна трансформувати у безліч синтаксичних варіацій за допомогою лексичних, синтаксичних, семантичних засобів. Так, речення *Хворий кашляє* перетворюється в похідне словосполучення *кашель хворого*. Подібні трансформації також слугують для побудови різних похідних речень, елементами яких і виступають отримані словосполучення (*Лікаря турбує кашель хворого*). До трансформацій, під час яких відбувається зміна значення вихідного речення, належать модальна (*Водій закриває двері – Водій хоче закрити двері*) і фразова трансформація (*Хворий кашляє – Хворий починає кашляти*). У подібних випадках в похідних реченнях з'являються операційні модальні і фразові дієслова. Будь-який складний синтаксичний тип виходить з одного або декількох ядерних речень шляхом застосування до них упорядкованого набору обов'язкових і факультативних трансформацій (*Він дивиться на картину. Картина намальована художником – Він дивиться на картину, яка була намальована художником*). Всі трансформи, одержувані від однієї ядерної конструкції, повинні зберегти вихідні семантичні зв'язки (*Учень написав твір – Твір написано учнем – Написаний учнем твір*).

Трансформаційний процес може бути багатоступеневим та різноспрямованим. У результаті послідовного застосування різних трансформацій утворюються ланцюжки речень, пов'язаних трансформаційними відношеннями (*Водій закриває двері. – Водій починає закривати двері. – Водій хоче почати закривати двері*). Це означає, що з кожного вихідного речення можна побудувати кілька розгалужених ланцюжків речень, які в сукупності утворюють так зване «дерево похідності».

Найпоширенішими видами трансформацій є:

1) Заміна одного елемента іншим (субституція): *Я побачив, що машина зупинилася / Я побачив це.*

2) Перестановка елементів (пермутація): *Урок розпочався / Розпочався урок (?)*.

3) Додавання елемента (ад'юнкція): *Машина зупинилася / Машина не зупинилася, машина вже зупинилася.*

4) Виключення елемента (еліпсис): *Мешканці говорили, що вітер був сильний / Говорили, що вітер був сильний.*

5) Перехід до іменника (номіналізація): *Хлопець сміється – сміх хлопця.*

6) Перехід до прикметника (ад'єктивація): *Обід був смачним – смачний обід; Склад викладачів – викладацький склад.*

Трансформації або зовсім не міняють змісту ядерних речень, або змінюють його незначно і тільки певним чином. Фактично, використовуючи лише невелику кількість трансформацій можливо отримати значну кількість різноманітних синтаксичних структур. Повний список трансформацій тієї чи іншої ядерної конструкції неможливо визначити, він постійно оновлюється носіями певної мови.

Контрольні запитання:

1. Що лежить в основі трансформаційного процесу? Як співвідносяться між собою поняття *трансформації* і *трансформи*?

2. Чи є поняття *трансформації* та *перекладу* тотожними? Чи є між ними різниця? Наведіть свої приклади.

3. Для розв'язання яких лінгвістичних завдань можна використовувати метод трансформаційного аналізу?

Практичні завдання

1. Побудуйте можливі трансформації з поданих нижче ядерних речень та визначте глибинну структуру, що лежить в основі цих трансформацій:

- Ненависть скривила його обличчя.
- Сніг проломив павільйон.
- Туга стискає серце.
- Гуркіт грому трясє хатинку.
- Очікування гнітить його.

2. Подано речення, у яких орудний відмінок останнього іменника має різні значення:

1. Студент дивує всіх чудовими успіхами.
2. Він чимчикував лісовою стежкою.

3. Я знав її маленькою дівчинкою.
4. Він читав книгу теплим вечором.
5. Машина промайнула чорною блискавкою.
6. Він рубав дерева гострим топірцем.
7. Я вважав його великим майстром.
8. Він виїхав нічним потягом.

Трансформуйте ці речення, зберігаючи їх первинний сенс, але замість орудного відмінка потрібно застосувати іншу граматичну конструкцію. Дозволяється трансформувати усе речення, а не лише конструкцію з орудним відмінком. Визначте глибинну конструкцію цих речень.

3. Нижче подано словосполучення типу $N_1 + N_2$ (іменник у називному відмінку + іменник у родовому відмінку):

- | | |
|---------------------|-----------------------|
| 1. допомога сусідів | 6. книга сестри |
| 2. прання білизни | 7. рукав сорочки |
| 3. шум вітру | 8. лікування пацієнта |
| 4. відро води | 9. дім батька |
| 5. радість зустрічі | 10. стакан молока |

Розкрийте значення родового відмінка у поданих словосполученнях, застосувавши відповідні трансформації:

Трансформація	Розшифрування
$N_1 + V$	суб'єкт виконує дію
$N_4 + V_{pl}$	над суб'єктом виконують дію
$N_1 + \text{частина} + N_2$	суб'єкт є частиною чогось цілого
$N_1 + \text{належить} + N_4$	суб'єкт належить іншому суб'єкту
$N_1 + \text{під час} + N_2$	суб'єкт проявляє себе під час події

Перебудуйте наведені словосполучення $N_1 + N_2$ відповідно до правил трансформації.

ТЕМА 12 МОДЕЛІ У ГЕНЕРАТИВНІЙ СЕМАНТИЦІ

План заняття:

1. Модель «Смисл ↔ Текст»
2. Модель концептуальної залежності
3. Модель семантики переваги

Одним із подальших напрямків розвитку теорії породжувальних граматик (кінець 1960-х – початок 1970-х рр.) стала генеративна або породжувальна семантика. Представники цього нового напрямку ставили за мету побудувати моделі мови за схемою «від значення до тексту» і навпаки – «від тексту до значення». Одним із відомих напрацювань у галузі формальної семантики є теорія відмінкової граматки Чарльза Філмора. Крім відмінкової граматки також відомими є модель семантики преференцій, модель концептуальної залежності та модель «Смисл ↔ Текст», які детально розглянуті нижче. У межах породжувальної семантики вчені проводили дослідження логічних аспектів граматичної структури та семантики речення. Отримані результати досліджень використовували в прикладній лінгвістиці й психолінгвістиці.

12.1. Модель «Смисл ↔ Текст»

Модель «Смисл ↔ Текст» (далі МСТ) спирається на розуміння природної мови як механізму, що перетворює задані сенси на відповідні їм тексти, а задані тексти – на відповідні їм сенси. Краще розробленим у моделі є перехід від сенсів до текстів, тобто мовний синтез. Як було зазначено в підрозділі 9.3, основною ідеєю моделі є те, що мова є засобом, за допомогою якого її носії виконують дві інтелектуальні операції: 1) повідомляють свої думки іншим людям, тобто кодують певні сенси текстами, що їх виражають (генерація текстів, синтез); та 2) розуміють думки інших людей, іншими словами, виконують зворотну операцію вилучення смислів зі сприйманих текстів (розуміння текстів, аналіз). У підрозділі 9.2 було наголошено, що модель являє собою складно організовану сукупність правил, суто механічне застосування яких має в ідеалі забезпечувати перехід від поданого тексту до формального опису сенсу, а також перехід від поданого змісту до відповідного тексту або

текстів. З погляду своєї внутрішньої будови МСТ є багаторівневою, де кожен рівень відповідає за ту чи іншу форму репрезентації текстів. МСТ включає морфологічний, поверхнево-синтаксичний та глибинно-синтаксичний рівні.

Розглянемо приклад репрезентацій речення «Різке потепління спричинило танення льодовиків» на цих рівнях.

На морфологічному рівні МСТ речення набуває вигляду морфологічної структури, тобто послідовності лексем із приписаною до них інформацією про граматичну форму, а саме: відмінок (наз., род., дав. тощо), число (одн., множ.), вид (док., недок.) та рід (чол., жін., сер.). Отже, морфологічна структура аналізованого речення матиме такий вигляд:

РІЗКИЙ, одн., назв., сер. // ПОТЕПЛІННЯ, одн., назв // СПРИЧИНЯТИ, мин., док., дійсн., одн., сер. // ТАЯННЯ, одн., знах. // ЛЬОДОВИК, множ., род.

На поверхнево-синтаксичному рівні речення перетворюється у дерево залежностей, де вузлами є імена лише семантично змістовних лексем з набором морфологічних характеристик. Самі вузли пов'язані одним із відношень підпорядкування. Гілки являють собою поверхнево-синтаксичні відношень між словоформами. Набір відношень не є універсальним і в кожній мові встановлюється окремо. Приклад відображення аналізованого речення у вигляді дерева залежностей показана на рисунку рис. 12.1.

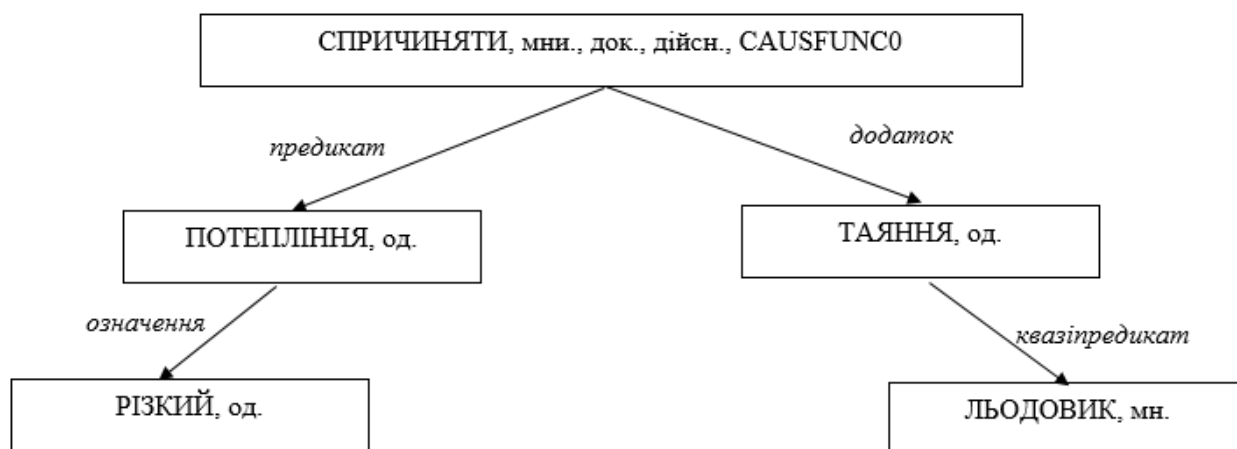


Рис. 12.1. Поверхнево-синтаксична структура речення.

На глибинно-синтаксичному рівні речення також перетворюються у дерево залежностей, вузлами якого є імена лексем з морфологічними характеристиками. Вузли дерева пов'язані певним типом відношення, а саме підпорядкування. Крім того для них також вказуються лексичні функції, наприклад CAUSFUNC0 на вершині дерева, як це показано на рисунку 12.1.

Лексичні функції (ЛФ), що грають ключову роль у теорії МСТ – *це абстрактне значення, яке виражається різними лексичними засобами*. Так, лексична функція Magn (від лат. magnus «великий») відповідає значенню інтенсифікації, високому ступеню прояву ознаки. За теорією МСТ, у мовах світу можна виокремити кілька десятків значень високого рівня абстракції, таких як «високий ступінь», «початок», «припинення», «каузація», «ліквідація», «маніфестація» тощо, кожне з яких виражається великим класом слів. При цьому вибір конкретного слова для вираження певного значення цілком залежить від того слова, з яким воно поєднується в тексті, тобто від контексту.

Наведемо приклад реалізації ЛФ «каузація», що позначається як CAUSFUNC0 – «бути причиною того, що дещо існує»:

– CAUSFUNC0(радість) = викликати(радість), дарувати(радість), приносити(радість);

– CAUSFUNC0(біда) = накликати(біду), обернутися(бідою), наробити(біди), потрапити в(біду).

Безпосередньо до семантики належать опис семантичного рівня і правил переходу до синтаксичного рівня, а точніше, до глибинно-синтаксичного. Система правил, що забезпечує такий перехід у моделі, називається **семантичним компонентом**. До семантичних компонентів МСТ належать тлумачно-комбінаторний словник і система перефразування, а також лексичні функції та модель керування. Опис семантичного рівня тексту передбачає наявність семантичної мови, яку складають:

– словник, що включає елементарні семантичні одиниці (семи), проміжні семантичні одиниці, що відповідають словам природної мови, та символи-характеристики комунікативної організації;

– правила утворення, за допомогою яких з одиниць словника будуються семантичні репрезентації;

– правила перетворення, що задають рівносильність двох семантичних репрезентацій (шляхом згортання фрагментів графа в проміжні одиниці та розгортання останніх).

Роботу семантичного компонента забезпечує тлумачно-комбінаторний словник (ТКС), за допомогою якого виокремлюються «узагальнені слова»⁴¹. ТКС містить різну інформацію про слово, розподілену за кількома зонами, головними з яких є тлумачення, модель керування, стандартні та нестандартні ЛФ. Для кожного слова в ТКС в ідеалі наводяться всі слова, пов'язані з ним за сенсом. Це, по-перше, його парадигматичні варіанти, або заміни, тобто ті мовні засоби, що замінюють певне слово у певних контекстах за певних умов, і, по-друге, його синтагматичні параметри, тобто ті, що ідіоматично виражають різні сенси при цьому слові. Усі ці засоби називаються **лексичними корелятами**, зв'язки слів із якими описуються за допомогою ЛФ. У свою чергу, набір ЛФ дає змогу впорядкувати опис сполучуваності слів і відокремити його від опису значень. Апарат стандартних ЛФ використовується також під час формулювання лексичних правил системи перифразування, характеристики деяких наведені у таблиці 12.1.

Таблиця 12.1. ЛФ, що використовуються у МСТ.

ЛФ	Позначає	Приклади
Magn	Високий ступінь / інтенсивність, максимальний прояв якості	Magn(аналіз) = глибокий, повний, ретельний
AntiMagn	Мінімальний прояв якості	AntiMagn(аналіз) = поверхневий, скупий
Bon	Позитивну якість	Bon(пропозиція) = вигідна Bon(людина) = приємна
AntiBon	Негативну якість або негативне ставлення	AntiBon(людина) = невихована, дурна, неадекватна AntiBon(праця) = марна
Func	Виконувану дію	Func(кіт) = їсть, спить, грається
Oper	Дію, виконуваного над	Oper(завдання) = виконувати

⁴¹ У західній лінгвістиці використовується термін *concept*.

	предметом	Oper(контроль) = здійснювати
Constr	Дія, в результаті якої предмет починає існувати	Constr(проект) = розробляти, створювати
Conv	Те саме відношення, але в іншому плані	Conv(покупати) = продавати Conv(боятися) = лякати
Syn	Заміщає ключове слово на інше, близьке за значенням	Syn(ходити) = шкутильгати, плентатися
Anti	Протилежний предмет, стан, властивість або якість	Anti(друг) = ворог Anti(поганий) = хороший
Correl	Співвідносне поняття, неантонімічно протиставлене лексичному елементу	Correl(фізичний) = моральний Correl(державний) = приватний

Індокси, що використовуються у назвах функцій, позначають номери актантів, учасників відповідних ситуацій (нульовий індекс позначає безособові Oper і Labor або Func, що не має доповнень). Так, наприклад, у ЛФ Oper1 (*допомога*) індекс позначає першого учасника ситуації «1», а сама ЛФ набуває значення *надавати*, а для другого учасника (індекс «2») ЛФ Oper2(*допомога*) – значення *отримувати*. Окрім простих функцій, існують також і складні, які отримують за допомогою комбінування простих, наприклад CausOper1 (*небезпека*) = *наражати на*. Набір лексичних функцій дає змогу впорядкувати опис сполучуваності слів і відокремити його від опису значень. Апарат стандартних лексичних функцій використовується також під час формулювання лексичних правил системи перифразування. МСТ включає 56 лексичних правил перифразування, разом з якими використовуються також синтаксичні, тобто такі перетворення синтаксичних конструкцій, що не змінюють сенсу всього висловлення.

Розглянемо нижче деякі правила та наведемо приклади їх застосування на конкретних текстах. Синонімічна заміна слова (C_0) у реченні без синтаксичних перетворень у вигляді формального правила може бути представлена таким чином:

$$C_0 \Leftrightarrow Syn(C_0)$$

(1) *Явище багатозначності $[C_0^1]$ розглядають багато $[C_0^2]$ мовознавців $[C_0^3]$. $\Leftrightarrow$ Явище полісемії $[Syn(C_0^1)]$ розглядають чимало $[Syn(C_0^2)]$ лінгвістів $[Syn(C_0^3)]$.*

Конверсивна⁴² заміна слова (C_0), що передбачає не лише заміну певного слова, а й граматичної форми оточуючих слів $[D]$ та / або усієї конструкції речення може бути представлена таким формальним правилом:

$$C_0 \Leftrightarrow Conv(C_0)$$

(2) *Брат $[D_1]$ старший $[C_0]$ за сестру $[D_2]$ на два роки. $\Leftrightarrow$ Сестра $[D_2]$ молодша $[Conv(C_0)]$ за брата $[D_1]$ на два роки.*

(3) *Учитель $[D_1]$ навчає $[C_0]$ школярів $[D_2]$ математиці $[D_3] \Leftrightarrow$ Школярі $[D_2]$ вивчають $[Conv(C_0)]$ з учителем $[D_1]$ математику $[D_3]$*

Деривативна заміна слова (C_0), тобто заміна слова з утворенням похідних слів із тих, що складають речення, представляється у вигляді правила:

$$C_0 \Leftrightarrow Adv(C_0)$$

(4) *Виявилося $[C_0]$, що хлопець не вчинив жодної шкоди. $\Leftrightarrow$ Як виявилося $[Adv(C_0)]$, хлопець не вчинив жодної шкоди.*

(5) *Вони завершили $[C_0]$ виступ привітанням учасників конференції $\Leftrightarrow$ На завершення $[Adv(C_0)]$ виступу вони привітали учасників конференції.*

Заміна з розкладанням, під час якої може виникати потреба вставити слово-зв'язку (для наочності у правилі взято в кутові дужки), виглядає таким чином:

$$C_0 \Leftrightarrow Pred(C_0)$$

(6) *Він приїхав на роботу завчасно $[C_0] \Leftrightarrow$ Його приїзд на роботу $[Pred(C_0)]$ завчасний.*

$$Pred(C_0) \Leftrightarrow Copul(C_0) \ 2 \rightarrow C_0$$

(7) *Він приїхав на роботу завчасно $[C_0] \Leftrightarrow$ Його приїзд на роботу $\langle$ був $\rangle$ $[Copul(C_0)]$ завчасним $[C_0]$.*

(8) *Він викладає англійську мову $\Leftrightarrow$ Він $\langle$ працює $\rangle$ $[Copul(C_0)]$ викладачем $[C_0]$ англійської мови.*

⁴² **Конверсія** – спосіб вираження лексико-семантичних відношень зворотних за своїм значенням мовних одиниць у вихідному і похідному висловлюваннях.

Розкладання з використанням ЛФ Oper, Func і Labor, де C_0 позначає дієслово, має такий формальний вигляд:

$$C_0 \Leftrightarrow //S_2/A_2(C_0) \ 2 \leftarrow Oper_1(S_2/A_2(C_0))$$

(9) *Оксана кохає $[C_0]$ Романа. $\Leftrightarrow$ Оксана закохана $[//A_2(C_0)]$ у Романа.*

(10) *Романа лікував $[C_0]$ Петро. $\Leftrightarrow$ Роман був $[Oper_1(S_2/A_2(C_0))]$ пацієнтом $[//S_2(C_0)]$ Петра.*

$$C_0 \Leftrightarrow S_0(C_0) \ 2 \leftarrow Oper_1(S_0(C_0))$$

(11) *Учені досліджували $[C_0]$ поведінку людей в екстремальних умовах. $\Leftrightarrow$ Учені проводили $[Oper_1(S_0(C_0))]$ дослідження $[S_0(C_0)]$ поведінки людей в екстремальних умовах.*

$$C_0 \Leftrightarrow S_0(C_0) \ 2 \leftarrow Func_1(S_0(C_0))$$

(12) *Він радіє $[C_0]$. $\Leftrightarrow$ Його переповнює $[Func_1(S_0(C_0))]$ радість $[S_0(C_0)]$.*

$$C_0 \Leftrightarrow S_0(C_0) \ 3 \leftarrow Labor_{13}(S_0(C_0))$$

(13) *Петро запозичив $[C_0]$ у банку невелику суму. $\Leftrightarrow$ Петро взяв $Labor_{13}(S_0(C_0))$ у позику $[S_0(C_0)]$ невелику суму у банку.*

Правила моделі МСТ також передбачають метафоричні та метонімічні розкладання й заміни:

$$X(C_0) \rightarrow C_0 \Leftrightarrow X(S_1(C_0)) \rightarrow //S_1(C_0), \text{ де } C_0 - \text{іменник, а } X - \text{певна ЛФ.}$$

(14) *Стаття $[C_0]$ аналізує питання $[X = Oper_1(C_0)]$... $\Leftrightarrow$ Автор $S_1(C_0)$ статті $[C_0]$ аналізує $[X = Oper_1(C_0)]$...*

$$C_0 \Leftrightarrow Equip(C_0) \ 2 \rightarrow C_0, \text{ де } C_0 - \text{іменник.}$$

(15) *Місто $[C_0]$ святкує Новий Рік. $\Leftrightarrow$ Мешканці $[Equip(C_0)]$ міста $[C_0]$ святкують Новий Рік.*

$X(C_0) \rightarrow C_0 \Leftrightarrow X(Figur(C_0)) \rightarrow Figur(C_0) \rightarrow C_0$, де C_0 – іменник, а X – певна ЛФ.

(16) *Війна $[C_0]$ охопила $[X = Func_2(C_0)]$ цілий континент. $\Leftrightarrow$ Пожежа $[Figur(C_0)]$ війни $[C_0]$ охопила $[Func_2(Figur(C_0))]$ цілий континент.*

Взаємозаміна лексичних корелятивів⁴³, що передбачає деякі перетворення структури речення, формалізується таким чином:

$Oper_1(C_0) \Leftrightarrow Oper_2(C_0)$

(17) *Банк надав $[Oper_1(C_0)]$ позику $[C_0]$ клієнтові $\Leftrightarrow$ Клієнт отримав $[Oper_2(C_0)]$ позику $[C_0]$ від банку.*

$Func_1(C_0) \Leftrightarrow Func_2(C_0)$

(18) *Йому належить $[Func_1(C_0)]$ ідея $[C_0]$ розділити всі наукові знання на окремі галузі $\Leftrightarrow$ Його ідея $[C_0]$ полягає $[Func_2(C_0)]$ у розділенні всіх наукових знань на окремі галузі.*

Автор моделі МСТ Ігор Мельчук зазначає, що запропонований ним перелік не претендує на повноту, оскільки в майбутньому можуть бути виокремлені нові ЛФ, які потребуватимуть створення нових правил перефразувань. Крім цього, правила виявляються доволі приблизними і мають певні неточності. Майже всі правила можна використовувати лише в обмежених змістових контекстах, пор: *Робочі мовчки пересували мішки $\neq$ Мішки мовчки пересувалися робочими*; або *Він спокійно купив у Петра книгу $\neq$ Петро спокійно йому продав книгу*. У правилах не враховуються перетворення, що виходять за межі простого речення, наприклад: *Ця книга розповідає про героя. $\Leftrightarrow$ Він купив цю книгу, що розповідає про героя*. Також не враховуються явища, належні до поверхневого синтаксису, такі як номіналізація, ад'єктивація: *прагнув завершити $\Leftrightarrow$ прагнув до завершення*. Нарешті, правила не уточнюють граматичних характеристик компонентів правила. Так, замінюючи дієслово на іншу синтаксичну конструкцію слід враховувати у деяких випадках вид дієслова (доконаний або недоконаний): *появився своєчасно $\Leftrightarrow$ поява була своєчасною, але появлявся своєчасно $\Leftrightarrow$ появи були своєчасними*.

Модель МСТ розроблялася для застосування, у першу чергу, в галузі машинного перекладу, де за її допомогою Ігор Мельчук намагався створити «діючу» модель мови. Передбачалося, що програми машинного перекладу

⁴³ **Лексичні кореляти** – всі слова, які пов'язані з ключовим словом парадигматичними та синтагматичними відношеннями.

функціонуватимуть на основі цієї моделі, але подальший їх розвиток у 2000-х рр. відбувався вже за рахунок новітніх інформаційних технологій, саме тому значення МСТ в історії лінгвістики складно оцінити через незначну кількість її послідовників та слабкий інтерес із боку сучасних лінгвістів. На Заході ця теорія відома лише в межах вузької групи найближчих учнів та співробітників ученого. Критики цієї моделі не вважають її лінгвістичною теорією, оскільки не містить «правил» і «узагальнень», як це спостерігається в пізніших працях Хомського. Надалі теорія МСТ знайшла своє продовження у фундаментальному творі Юрія Апресяна «Інтегральний опис та системна лексикографія».

Контрольні запитання:

1. У чому полягає сутність МСТ та на який принципах вона ґрунтується?
2. Яку роль грає тлумачно-комбінаторний словник як компонент моделі МСТ?
3. Яке наукове значення, на вашу думку, має теорія МСТ?

Практичні завдання

1. Наведіть сполучення, які можна утворити з поданими нижче словами, використовуючи апарат лексичних функцій:

авторитет, агресія, висота, влада, грамотність, дощ, друг, дружба, кіно (фільм), мистецтво, шлюб.

Приклад виконання завдання: страх

ЛФ	Можливі значення ЛФ
Magn(страх)	бурхливий, забурний, забурунний, панічний
Oper(страх)	відчувати, терпіти, натерпітися
CausFunc(страх)	викликати, вселяти, наганяти
Func(страх)	охоплювати, підкрататися, піддаватися
FuncMagn(страх)	пробирати
FuncAntiMagn(страх)	ворушитися (в душі)
LiquFunc(страх)	побороти, подолати
Syn(страх)	боязкість, фобія

2. *Перефразуйте подані речення без порушення сенсу, надаючи усі можливі варіанти:*

1. Територія Франції менше, ніж територія Техасу.
2. Величезний пес темно-сірого кольору неспокійно бігав по подвір'ю.
3. Всі члени сім'ї відчували сильну тугу через раптовий виїзд дядька.
4. Найбільший інтерес викликали переговори, пов'язані з торгівлею.
5. До кімнати увійшов хлопчик з широким обличчям, з вузькими плечима.
6. Візьмемо кут А, цей кут є гострим.
7. Школярі змайстрували двадцять маленьких столів.
8. Зробимо наближене обчислення даної функції.
9. У пробірці утворилася густа маса, схожа на тісто.

3. *Перефразуйте речення у завданні 2 з використанням систем перефразування тексту (як приклад, можна взяти систему PARAPHRASER: <https://www.paraphraser.io/uk/paraphrase-tool>).*

12.2. Модель преференційної семантики

Модель преференційної семантики (англ. preference semantics model – PSM) створена для аналізу змісту тексту. Вона має можливість відображати значення слів та словосполучень з урахуванням їхнього контексту. Стала вперше відомою у 1972 р., хоча вона розроблялася ще з 1964 р. Її автором є британській фахівець в галузі штучного інтелекту та опрацювання природної мови Йорік Уілкс⁴⁴. Зокрема, його доробки знайшли своє застосування в автоматизованих питально-відповідальних системах, а також чат-ботах. PSM – це теорія мови, за якою зміст певного тексту відображається складною семантичною структурою, побудованою з тих чи інших компонентів⁴⁵. На відміну від інших семантичних теорій, вона надає можливість описати у формальному вигляді мовні такі феномени, як лексичні неоднозначності, метафоричні та узуальні значення для їх подальшого машинного опрацювання. Семантика переваги припускає, що такі особливості вживання природної мови

⁴⁴ <https://www.theguardian.com/technology/2023/jun/09/yorick-wilks-obituary>

⁴⁵ Детально про теорію семантики переваги: <https://core.ac.uk/download/pdf/82487541.pdf>

є стандартними та розповсюдженими, і жодна теорія мови не може їх ігнорувати. Основна ідея теорії полягає в тому, що у семантичному плані певний предикат віддає перевагу (преференцію) певному суб'єкту та/або об'єкту дії. Наприклад:

Cat drinks milk. (1)

John ran quickly. (2)

Так, у прикладі (1) предикат «drink» віддає перевагу живій істоті як суб'єкту дії (кіт, людина, дитина). У прикладі (2) предикат «run» віддає перевагу нульовому об'єкту дії, тобто дія не передбачає об'єкта.

Розглянемо інші приклади, що запропонував Йорік Уілкx:

My car drinks gasoline (3)

John ran his life (4)

У реченні (3) відбувається порушення, оскільки «drinks» віддає семантичну преференцію істоті як суб'єктові дії. Щодо (4), предикат віддає семантичну преференцію нульовому об'єкту дії, як це було зазначено вище. Предикат «run» може передбачати об'єкт, якщо це буде метафора, пор. «to run my life». За теорією Уілкx, у прикладах (3) і (4) мають місце метафори, які інтерпретуються, відповідно: «машина використовує бензин» і «Джон прожив своє життя».

Для аналізу природної мови автор моделі розробив систему з дев'яти фреймів, кожен із яких виконує певну функцію:

(1) 80-100 семантичних «елементів» або примітивів на позначення елементів речення / тексту;

(2) семантичні формули для репрезентації значень слів, які входять до речення / тексту;

(3) прості шаблони для репрезентації змісту речення, яке в загальному плані описує ту чи іншу ситуацію;

(4) шаблони для репрезентації тексту або речень, у яких відображаються характеристики учасників комунікативної ситуації (суб'єкт дії, власне дія та об'єкт дії);

(5) надшаблони для репрезентації значень прийменників, зокрема для відображення відмінкової неоднозначності прийменників;

(6) умовиводи на основі здорового глузду, які використовуються для розуміння більш складних речень;

(7) семантичні блоки для відображення певного абзацу тексту або навіть цілих текстів;

(8) тезаурусна структура, яка забезпечує ієрархічну структуру із формул та псевдотекстів;

(9) псевдотексти, що репрезентують структуру тексту, а також умовний, фактографічний аспект «фреймових структур».

Методика Уілкса використовує процедуру фрагментації, щоб розбити текст на частини, а потім відобразити їх у міжмовній структурі, яка називається семантичним блоком. Цей блок складається з шаблонів, пов'язаних між собою надшаблонами та умовиводами на основі здорового глузду. Кожен з цих трьох елементів складається з формул, а над формулами та підформулами розміщені предикати та функції. Своєю чергою, формули складаються з елементів (або примітивів). Розглянемо детально наведені вище фрейми та розкриємо їхню роль у здійсненні семантичного аналізу з використанням PSM.

Семантичні примітиви (фрейм 1) поділяються на чотири основні групи – відмінки, дієслова, іменники та означення плюс два «різні елементи». Теорія Уілкса нараховує 21 відмінковий примітив, включаючи SUBJ (відмінок агенса) і OBJE (відмінок об'єкта) та *⁴⁶DIRE (загальний відмінок напрямку); 34 примітиви дії, включаючи GET (отримання якоїсь речі або речовини) і GIVE (відмова від якоїсь речі або речовини); 30 примітивів іменників, включаючи SPREAD (розширення) та GRAIN (будь-яка структура); 16 визначальних елементів, включаючи GOOD (морально правильний або схвалений), MAL (вказує на належність суб'єкта до чоловічої статі) та FEM (вказує на належність суб'єкта до жіночої статі). Відмінкові, субстантивні (іменникові) та дієслівні

⁴⁶ Зірочка «*» позначає клас примітивів. Наприклад, *ANI – клас істот, до якого входять MAN, FOLK, BEAST тощо.

примітиви можна розглядати як невеликі, ієрархічно організовані семантичні мережі, в яких примітиви виступають як вузли.

На рисунку 12.2 показана мережа для субстантивних примітивів, а в таблиці 12.2 наведено перелік із 11 субстантивних примітивів разом із розшифруванням.

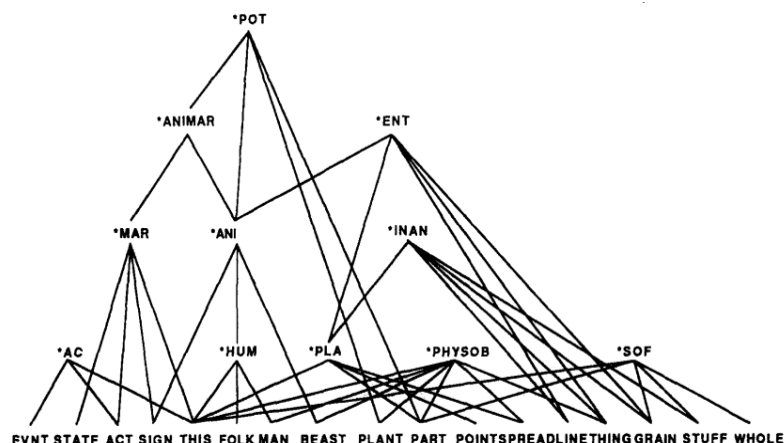


Рис. 12.2. Семантична мережа субстантивних примітивів у PSM.

Таблиця 12.2. Субстантивні семантичні примітиви.

Клас	Примітиви
*POT	будь-яка «дієздатна» сутність (тобто, *ANI, STATE, PART, ACT, PLANT)
*ANIMAR	поєднання примітивів *ANI та *MAR
*ANI	будь-яка жива істота (тобто, THIS, MAN, FOLK, BEAST або SIGN)
*MAR	будь-які «знакоподібні» об'єкти (тобто, THIS, SIGN, ACT, STATE)
*ENT	будь-яка сутність (тобто, THIS, POINT, FOLK, MAN, GRAIN, PART, THING, BEAST, SIGN, SPREAD, або LINE)
*INAN	будь-який неживий об'єкт (тобто, THIS, PART, GRAIN, STUFF, THING, LINE, SPREAD)
*AC	будь-який «визначник дій» (тобто, THIS, ACT, EVNT)
*HUM	будь-яка людська істота (тобто, THIS, MAN, FOLK)
*PLA	будь-який «визначник місця» (тобто, THIS, POINT, SPREAD, PART)
*PHYSOB	будь-який фізичний об'єкт (тобто, THIS, THING, MAN, BEAST, SPREAD, PLANT, PART)
*SOF	будь-який «піддатливий» суб'єкт (тобто, THIS, PART, WHOLE, GRAIN, STUFF)

Із семантичних примітивів будують **семантичні формули** (фрейм 2), що відображають окремі значення слів. Формули утворюються з підформул за допомогою синтаксису залежностей. Підформула – це будь-яка ліво-права пара, два члени якої є або примітивом, або іншою підформулою. Правий елемент є керуючим, а лівий – залежним. Крайній правий елемент семантичної формули – власне це її заголовок, що репрезентує «головний елемент» у значенні, наприклад: чи іменник відноситься до ЛЮДИНИ або РЕЧІ, чи дієслово позначає акт МИСЛЕННЯ або ДІЇ. Нижче наведено приклад семантичної формули (у її спрощеному вигляді), яка описує значення дієслова «drink»:

((*ANI SUBJ) (((FLOW STUFF) OBJE) (MOVE CAUSE))))

Можна сказати, що ця формула означає дію, тому що її заголовок, CAUSE, є елементом дії. Наведена формула інтерпретується таким чином: пиття – це причинно-наслідкова дія, яку переважно здійснюють живі істоти (*ANI SUBJ) по відношенню до рідини ((FLOW STUFF) OBJE). Відмінок SUBJ показує бажаного виконавця дії, а відмінок OBJE – бажаний об'єкт або пацієнта. У таблиці 12.3 наведено елементи семантичної формули (підформули), а також семантичні відмінки разом із їхніми значеннями.

Таблиця 12.3. Відмінки та їхні значення у семантичній формулі «drink».

Підформула	Відмінок	Значення відмінка
(*ANI SUBJ)	SUBJ	*ANI
((FLOW STUFF) OBJE)	OBJE	(FLOW stuff)
(MOVE CAUSE)	CAUSE	MOVE

Наведемо ще одну семантичну формулу для опису значення дієслова «fire at», семантичні відмінки якої разом із їхніми значеннями надано у таблиці 12.4:

((*HUM SUBJ) ((*ANI OBJE) ((STRIK GOAL) ((THING INST) ((THING MOVE) CAUSE)))))

Таблиця 12.4. Відмінки та їхні значення у семантичній формулі «fire at».

Підформула	Відмінок	Значення відмінка
(*HUM SUBJ)	SUBJ	*HUM
(*ANI OBJE)	OBJE	*ANI
(STRIK GOAL)	GOAL	STRIK
((THING MOVE) CAUSE)	CAUSE	THING MOVE

Пізніше Уїлкс зазначив, що, оскільки не існує реальної різниці між примітивами та словами (перші є лише підмножиною других), допускається використовувати формули «змішаного типу», в які вставляються англійські слова. Наприклад, слова «пістолет» і «куля» вставляються в семантичну формулу «стріляти», як це показано нижче:

(((*HUM SUB J) ((*ANI OBJE) ((STRIK GOAL) ((gun INST) ((bullet MOVE) CAUSE))))))

Ці формули були описані як «вхідні», оскільки вони згадуються одна в одній, тобто «пістолет» і «куля» мають власні семантичні формули в інших частинах системи.

Простий шаблон (фрейм 3) – це структура, що містить «сутність» повідомлення. Кожен такий шаблон містить три примітиви – трійку «агент-дія-об’єкт». Як приклад, розглянемо речення типу: «The crook drank some beer». Із наведеного речення отримано заголовки для семантичних формул «crook», «drink» та «beer». Припустимо, що «crook» має два значення, менш вірогідне з яких – пастуша клюка, з головами [crook (MAN)] та [crook (THING)]. Заголовок «пити» – CAUSE, а заголовок «beer» – (FLOW STUFF). Пропонується два варіанти інтерпретації:

[MAN CAUSE (FLOW STUFF)]

[THING CAUSE (FLOW STUFF)]

Шаблон (фрейм 4) – це конструкція, що базується на слотах агента, дії та об’єкта для трьох семантичних формул, які своєю чергою можуть мати залежні формули, так що вся конструкція відображає можливе «повідомлення» і може містити будь-яку кількість формул. Кожному фрагменту речення (або словосполученню) відповідає певний шаблон. Наявність більш ніж одного шаблону на один фрагмент свідчить про неоднозначність, яку можна зменшити, дослідивши: 1) внутрішню «узгодженість» шаблонів та 2) зв’язки між шаблонами для послідовних фрагментів тексту. Алгоритм розширення шаблону переглядає формули в кожному шаблоні та аналізує їхні відмінкові частини,

щоб визначити, чи задовольняються їхні преференції. Тут [квадратні дужки] позначають семантичну формулу для слова-значення.

Так, [crook (MAN)] позначає формулу людського значення слова «шахрай», де MAN – головний примітив. Для речення «The crook drank some beer»:

- [drink] віддає преференцію живій істоті як агенту;
- [crook (MAN)] може задовольнити цю преференцію, але [crook (THING)] не може;
- [drink] віддає преференцію рідині як об'єкту (в обох варіантах інтерпретації).

Аналізоване речення можна подати у вигляді двох шаблонів (стрілки → або ← позначають задоволені преференції):

[crook (MAN)] → [drink] ← [beer (FLOW STUFF)]

[crook (THING)] → [drink] ← [beer (FLOW STUFF)]

Перший має найбільшу кількість задоволених преференцій, або більшу «семантичну щільність», у порівнянні з другим. Саме тому обирають перший. У першому задовольняються дві преференції: «шахрай» як жива істота – агент дії, та рідина як об'єкт дії. У другому шаблоні задовольняється лише одна преференція: об'єкт дії (рідина).

Надшаблони (фрейм 5) призначені для розв'язання проблеми відмінкової неоднозначності прийменників. Надшаблони – це шаблони, що охоплюють два шаблони, впорядковані від конкретного до загального. Так, у наведених нижче прикладах прийменник *by* виражає різні відмінкові відношення: (5) – DIRECTION, у (6) – INSTRUMENT відмінок, а в (7) – знову DIRECTION:

He left Comano by the autostrada. (5)

He left Comano by car. (6)

He left Comano by following the arrows. (7)

Кожний надшаблон відображає відповідний відмінок, де [] позначає умовного агента:

(*ANI) (MOVE) (WHERE POINT) → [] [] (WHERE LINE)

(*ANI) (MOVE) (WHERE POINT) → [] [] (*REAL)

(*ANI) (MOVE) (WHERE POINT) → [] (*DO) (WHERE SIGN)

Три елементи перед стрілкою зіставляються з шаблоном, який репрезентує частину речення, що стоїть перед прийменником. Три елементи після стрілки зіставляються з шаблоном, що репрезентує речення за прийменником.

Умовиводи на основі здорового глузду (фрейм 6) потрібні для розуміння більш складних речень, як це показано нижче. На кроці (1) вводиться висловлювання, у цьому прикладі воно репрезентує семантичну неоднозначність слова «them». Крок (2) – вихідне висловлювання трансформується у запис. На кроці (3) висловлювання перетворюється у шаблон, з якого можна зрозуміти, що «them» означає «women», а не «soldiers»:

(1) The soldiers fired at the women and we saw several of them fall.

(2) (1 (THIS STRIK) (*ANI 2) <-> ((*ANI 2) (NOTUP BE) DTHIS)

(3) [1 strikes animate2] <-> [animate2 falls].

Семантичний блок (фрейм 7) являє собою абзац (або навіть цілий текст), який складають шаблони, надшаблони та умовиводи на основі здорового глузду, кожен з яких у свою чергу складається з формул, що містять предикати та функції, які знаходяться над формулами та під формулами. Формули, у свою чергу, складаються з елементів: елементів семантичної репрезентації, які являють собою відповідні текстові елементи, як показано в Таблиці 12.5.

Таблиця 12.5. Текстові елементи та елементи семантичної репрезентації.

Елементи семантичної репрезентації	Відповідні текстові елементи
Формула	Значення англійського слова
Шаблон	Речення англійською
Семантичний блок	Абзац або текст англійською

Надшаблони та умовиводи на основі здорового глузду разом об'єднують усі ці шаблони в єдиний семантичний блок. У свою чергу, семантичні елементи існують як примітиви, з яких складаються всі вищезгадані складні елементи.

Тезаурусна структура (фрейм 8) забезпечує лексичним матеріалом процес репрезентації семантичних примітивів. Він встановлює ієрархію формул та псевдотекстів. Тезаурус, так само як і тезаурус Роже⁴⁷, – це спосіб угруповання англійських слів у синонімічні ряди, де слова належать до однієї частини мови. Ці ряди згруповані під однією з близько тисячі рубрик, які, в свою чергу, згруповані під десятьма дуже загальними розділами. Ці загальні розділи Роже можна ідентифікувати з такими примітивами, як MAN, THING, і, можливо, з іншими примітивами, що відповідають назвам розділів Тезауруса Роже, наприклад, #Abstract Relations (GP~IN), #Space (WHERE), #Matter (STUFF), #Intellect (THINK), #Volition (GOAL) тощо. Так, розділ #Volition включає заголовок, наприклад, #22 propulsion, під яким іде підзаголовок #221 firer, приєднаний до якогось ряду слів «firer»: gun, bow, rifle, howitzer ...

Псевдотекст (фрейм 9) – це структура фактичної та функціональної інформації про концепт або предмет». Псевдотексти для загальних примітивів, таких як MAN (які потім можуть бути прирівняні до дуже загальних заголовків у верхній частині тезауруса, в даному випадку «Людство») будуть складатися з дуже загальних стверджувальних форм, таких як:

MAN HAVE THING

MAN THINK SIGN

MAN WANT THING

Псевдотексти використовуються лише в випадку порушення преференцій. Показовим прикладом, де відбувається таке порушення, є речення (3), в якому для його усунення використовується проекція, алгоритм якої порівнює шаблони:

[my + car drink gasoline]

Аналізуючи речення, модель порівнює шаблони з псевдотексту «car», шукаючи найближчий збіг, в результаті чого вибирає [ICengine (USE) #liquid], оскільки

⁴⁷ Детально про тезаурус Роже можна ознайомитися за посиланням:
https://en.wikipedia.org/wiki/Roget%27s_Thesaurus#:~:text=Roget's%20Thesaurus%20is%20composed%20of,clusters%22%20or%20semantically%20linked%20words.

автомобіль має двигун внутрішнього згоряння, а бензин – це рідина. Далі (USE) проектується на «drink» у формальному записі речення, яке перетворюється у:

[my + car use gasoline].

У PSM сам термін *преференція* набуває різних тлумачень, залежно від того про що йде мова. **Перша інтерпретація:** *преференція* – це обмеження вибору. Ці переваги прописані у семантичних формулах (фрейм 2) і надшаблонах (фрейм 5), а також у простих шаблонах (фрейм 3). Отже, у наведеній вище семантичній формулі (фрейм 2) для «пити» перевагою агента є примітив *ANI, а перевагою об'єкта – (FLOW STUFF).

Друга інтерпретація: *преференція* – це булеве значення, яке або виконується, або порушується. У моделі порушення переваг оцінюється за допомогою алгоритму розширення шаблону, що оперує ієрархіями примітивів (фрейм 1). Якщо ієрархії розглядати як семантичні мережі з дугами, що позначають включення класів, алгоритм перевіряє відсутність порушення переваги шляхом побудови мережових шляхів. Задоволеними перевагами вважаються ті шляхи через семантичні мережі примітивів, які позначають випадки включення класів. Відповідно, порушені переваги – це шляхи, що позначають виключення класу. Щоб проілюструвати це, повернемося до речення (3), у якому для дії «пити» задовольняється *преференція* істоти як агента дії («сгоок»). Такому агенту дії відповідають два семантичні примітиви, *ANI та MAN, що показані на рисунку 12.3. Вони є елементами семантичної мережі субстантивних примітивів, показаної на рисунку 12.2. Дуги мережі між *ANI та MAN позначено жирною лінією. Дуга позначає включення, тобто чоловіки є живими істотами, отже, *преференція* *ANI задовольняється.

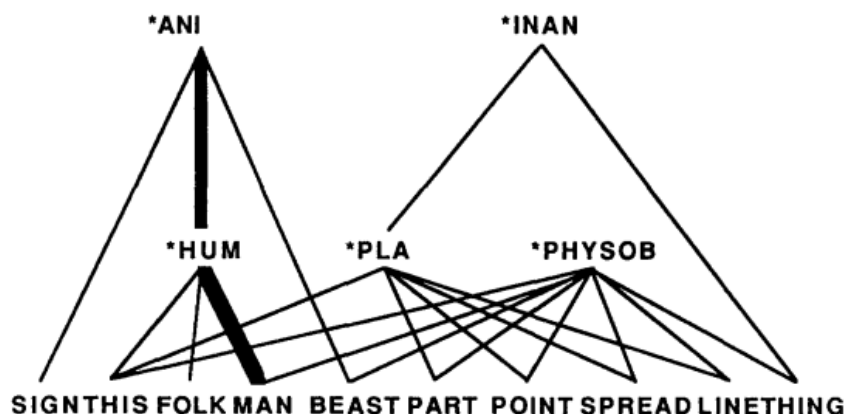


Рис. 12.3. Дуги семантичної мережі, що позначають включення класів.

Третя інтерпретація: *преференція* – це вибір шаблону, який має найбільшу семантичну щільність, тобто в якому задоволена максимальна кількість преференцій, які відповідають суб'єкту та об'єкту дії (див. вище приклад вибору шаблону для речення «The crook drank some beer»).

Четверта інтерпретація: *преференція* – це вибір найбільш конкретного варіанту семантичної репрезентації тексту з використанням надшаблонів (фрейм 5) у випадку семантичної неоднозначності прийменників: можливі варіанти упорядковуються від конкретного до загального, і кожен з них перевіряється по черзі, доки не буде знайдено відповідність.

Теорія преференційної семантики Уїлкса припускає, що зв'язність тексту (дискурсу) можна обчислити знизу вгору, виходячи з семантики окремих речень. Зв'язність обчислюється з точки зору преференцій шаблонів для різних семантичних формул. Теорія надалі розвивалася послідовниками Уїлкса (Давід Картер, Артур Кейтер, Сюймін Хуан та інші), які змогли реалізувати її у вигляді комп'ютерної програми «Meta 5». Вона аналізує речення, розрізняє сім типів семантичних зв'язків між парами значень слів у цих реченнях і знімає певну лексичну неоднозначність. Програма складається з лексикону, що містить фрейми близько 500 значень слів, обмеженого переліку граматичних правил і семантичних процедур, що містять процеси зіставлення та відсіювання.

Контрольні запитання:

1. У чому полягає теорія преференційної семантики Йоріка Уїлкса?

2. Який сенс автор теорії вкладає в слово *преференція*?
3. Які чинники зумовили появу такого напрямку як породжувальна семантика?

Практичні завдання

1. Подайте речення у вигляді шаблонів (див. фрейм 4) та визначте, які задовольняються при цьому *преференції*.

The policeman interrogated the crook.

The chopper drinks gasoline.

Для виконання цього завдання потрібно визначити:

*Що можуть означати слова **crook** та **chopper**?*

*Якому суб'єкту та об'єкту дії віддається у випадку дії, позначеної дієсловом **interrogate**?*

2. За допомогою графу на малюнку 12.2 визначте, яким семантичним примітивам можуть відповідати слова:

policeman, crook, chopper, gasoline.

3. Побудуйте прості шаблони (див. фрейм 3) для українських дієслів, що використовуються у поданих нижче реченнях:

- рубати: Дяк заставляв Тараса рубати дрова;
- бити: Побила б тебе лиха година;
- поливати: Дерев'яну підлогу поливали водою;
- будувати: Будемо будувати дім.

Для побудови шаблонів рекомендується використовувати семантичні примітиви з таблиці 12.2.

12.3. Модель концептуальної залежності

Ще одним доробком в галузі систем штучного інтелекту, призначених для опрацювання природної мови є модель концептуальної залежності (conceptual dependency theory), розроблена групою дослідників Єльського університету під керівництвом Роберта Шенка 1975 р. Модель Шенка – це спосіб відображення висловлювань природної мови таким чином, щоб із них можна було зробити відповідні умовиводи. Таке відображення не залежить від мови та будується

з примітивів, які є не словами, а базовими концептами. Примітиви, у свою чергу, об'єднуючись один із одним, формують значення слів. Як приклад, розглянемо ситуацію, репрезентовану реченням «I gave this man a book».

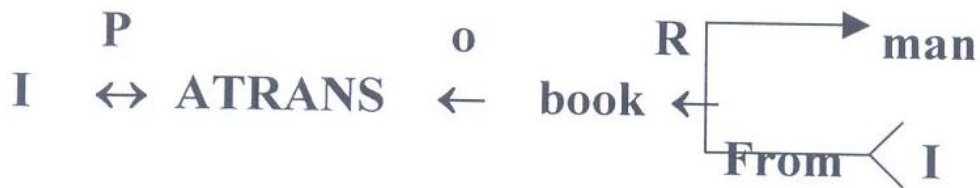


Рис. 12.4. Концептуальне відображення речення.

У поданій схемі подвійні стрілки позначають два можливих варіанти зв'язку між агенсом (суб'єктом дії) і дією. Символом *P* позначено минулий час; *ATRANS* – одна з примітивних (базових) дій, що використовується у теорії для позначення передачі володіння; *o* вказує на відношення об'єктного відмінка; *R* вказує на відношення об'єктного відмінка до реципієнта. Теоретичний апарат дає змогу відображати багато глибинних зв'язків природної мови, таких як відношення між дієсловом та його суб'єктом або дієсловом та його об'єктом. Відображення таких відношень відбувається за рахунок відмінкової структури. Розберемо речення «John cut bread with a knife», схема якого подана на рисунку 12.4. У цьому реченні *John* є агентом, тобто суб'єктом дії (*cut*), *bread* – об'єктом дії, а *knife* – інструментом здійснення дії.

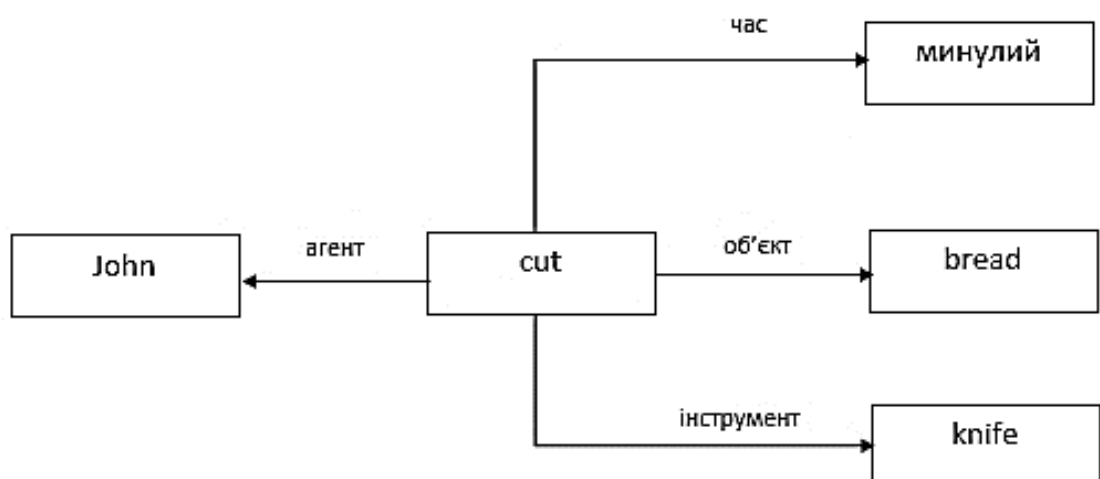


Рис. 12.5. Схема відношень, репрезентованих у реченні.

Показовим для моделі Шенка є заперечення впливу синтаксису і морфології на семантику. Основою семантичної репрезентації є мережа

концептуалізацій – подібний до графу об’єкт, вершини якого позначені символами семантичних одиниць чотирьох класів (наведені у таблиці 12.6), запропонованих Робертом Шенком.

Таблиця 12.6. Класи семантичних одиниць.

Клас	Елементи класу	Частина мови
PP	імена фізичного об’єкта	іменник
ACT	імена дій	дієслово
PA	характеристики об’єкта	прикметник
AA	характеристики дій	прислівник

Модель концептуальної залежності забезпечує структуру, в якій можуть бути репрезентовані знання, та набір будівельних блоків, з яких можна побудувати репрезентації. Типовий набір примітивів-дій наведено у таблиці 12.7.

Таблиця 12.7. Примітиви-дії.

Примітиви-дії	Що означає	Приклади дій
ATRANS	Передача абстрактного відношення	give
PTRANS	Передача фізичного місцезнаходження об’єкта	go
PROPEL	Застосування фізичної сили до об’єкта	push
MOVE	Рух частини тіла його власником	kick
GRASP	Хапання предмета агенсом	catch, grab
INGEST	Поглинання об’єкта	eat
EXPEL	Видавання звуків	cry
MTRANS	Передавання ментальної інформації	tell
MBUILD	Створювати нову ментальну інформацію	decide
SPEAK	Вимовляння звуків	speak
ATTEND	Спрямованість органу чуття на стимул	listen

Для відображення відтінків сенсу передбачено кілька модифікаторів відношень: *p* – минулий час, *f* – майбутній час, */* – заперечення, *k* – тривалість, *ts* – початок передачі, *tf* – закінчення передачі, *c* – умова, *?* – запитання, ∞ – нескінченість, *nil* – теперішній час. Для позначення відношень або залежностей використовуються стрілки $\leftarrow$ (вказує напрямок залежності) та $\Leftrightarrow$ (позначає відношення агент – дієслово). Характер дій позначається літерними маркерами

O (позначає об'єктне відношення), *D* (напрямок об'єкта під час виконання дії), *R* (реципієнт) та *I* (інструмент).

Розглянемо деякі приклади використання примітивів, які наводить Роберт Шенк для пояснення своєї теорії на реальних висловленнях. Найабстрактнішим із глобальних примітивів-дій є *ATRANS*, який може оперувати з відношеннями, що мають місце між багатьма реальними об'єктами. Як приклад, висловлення «John gave Mary a book» подається у вигляді концептуальної схеми (рис.12.6).

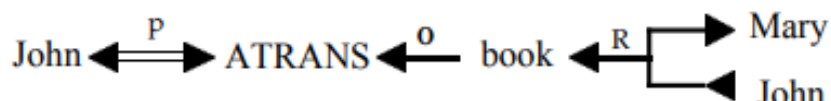


Рис. 12.6. Концептуальна схема висловлення «John gave Mary a book».

де *O* позначає об'єкт передачі, а *R* – реципієнта дії. Символ *P* вказує на те, що факт передачі мав місце в минулому. Так само можна подати речення «Mary took a book from John» (рис. 12.7).

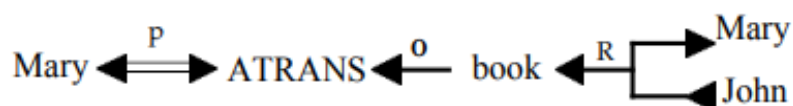


Рис. 12.7. Концептуальна схема висловлення «Mary took a book from John».

Дещо складнішою в інтерпретації Шенка є концептуальна схема речення «John ate a frog», у якій використовується примітив-дія *INGEST*. Цей примітив може репрезентувати дії, що виражені дієсловами *eat*, *drink*, *swallow*, *chew* та інші. Наведемо схему для аналізованого висловлення (рис. 12.8).

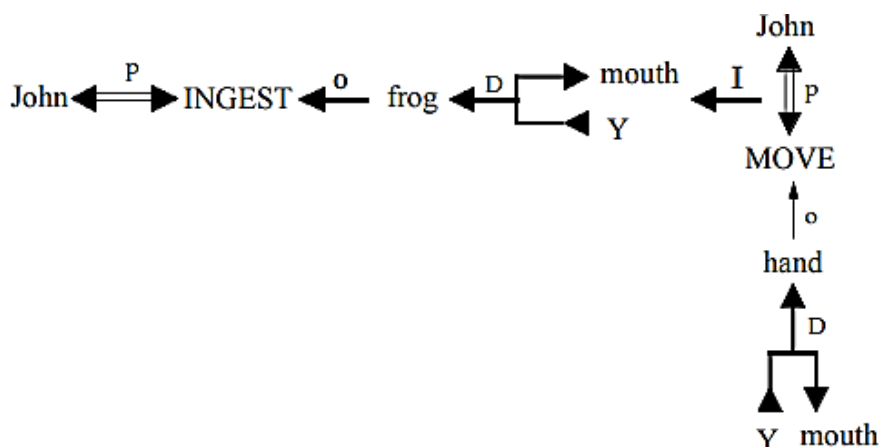


Рис. 12.8. Концептуальна схема висловлення «John ate a frog».

Подана схема отримує таку семантичну інтерпретацію: суб'єкт *John* виконує дію, що пов'язана з поглинанням об'єкта *frog* та має конкретний напрямок (символ *D*) дії *mouth*. Крім цього, схема репрезентує спосіб здійснення дії (символ *I*) – рухом частини тіла (MOVE), тобто руки, у напрямку рота (*mouth*). Таким чином, дія *eat* інтерпретується як поєднання примітивів-дій *ingest* та *move*.

У концептуальних схемах використовуються концепти на позначення станів об'єктів: HEALTH, FEAR, MENTAL STATE, CONSCIOUSNESS. Стани можуть набувати числових значень заданого діапазону. У таблиці 12.7 наведено приклад числових значень.

Таблиця 12.8. Діапазон числових значень стану MENTAL STATE.

MENTAL STATE	(діапазон -10 ...+10)
catatonic	-9
depressed	-5
upset	-3
sad	-2
ok	0
pleased	+2
happy	+5
ecstatic	+10

Нижче побудовано концептуальну схему висловлення «John annoyed Mary» (див. рис. 12.8). Як можна побачити, схема відображає зміну емоційного стану *Mary*, що на початку складав > -2 , тобто *Mary* могла відчувати себе добре (0), задоволеною (+2), щасливою (+5) або найщасливішою (+10). Але в результаті дії агента *John* емоційний стан *Mary* набув значення -2 (засмученою). Крім цього, додатковий параметр «anger», що складає -2, означає переживання негативних емоцій.

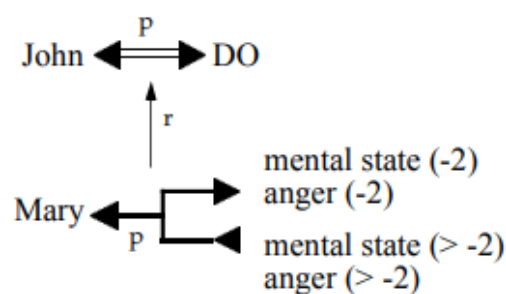


Рис. 12.9. Концептуальна схема висловлення «John annoyed Mary».

Теорія концептуальних залежностей, як наголошує Роберт Шенк, надає основу для відображення сенсу висловлювань та уможливлює:

– розкладання слів на примітиви, за рахунок чого опрацювання мови фокусується на загальних поняттях, а не на окремих словах;

– репрезентацію у вигляді концептуальних схем і фіксує спільність між різними словами та структурами (пор., наприклад: «John went to San Francisco by BART» та «John took BART to San Francisco»);

– відображення семантичної структури речення незалежно від конкретної мови, що полегшує машинний переклад;

– визначення концептуальної ролі та розв'язання проблеми семантичної неоднозначності.

Проте сам процес перетворення виразів у канонічну форму (у вигляді концептуальних схем) важко формалізувати, тоді як обчислювальна складність перетворення на набір примітивів низького рівня є досить високою. Крім того, самі примітиви не дають змоги відобразити деякі нюанси природної мови, наприклад, невизначеність деяких понять.

Контрольні запитання:

1. Порівняйте особливості відображення сенсу засобами моделі преференційної семантики та моделі концептуальної залежності.

2. За рахунок чого відбувається передача сенсу висловлювань у моделі концептуальної залежності?

3. Схарактеризуйте роль примітивів-дій у відтворенні сенсу речення або висловлювання за допомогою моделі концептуальної залежності.

Практичні завдання

1. Визначте, яким концептуальним схемам відповідають наведені нижче висловлювання:

1. Mary cried.
2. Mary took a book from John.
3. John killed Mary.
4. John made Mary happy.

(a) ... $\xleftrightarrow{P}$ DO ...

$\uparrow$
r

... $\xleftrightarrow{P}$ physical state (...)

... $\xleftrightarrow{P}$ physical state (...)

(b) ... $\xleftrightarrow{P}$ DO

$\uparrow$
r

... $\xleftrightarrow{P}$ mental state (...)

... $\xleftrightarrow{P}$ mental state (...)

(c)

... $\xleftrightarrow{P}$ EXPEL $\xleftarrow{O}$... $\xleftrightarrow{D}$ Y

(d)

... $\xleftrightarrow{P}$ ATRANS $\xleftarrow{O}$... $\xleftrightarrow{D}$...

2. Замість трьох крапок вставте необхідні елементи (наприклад, суб'єкт дії, об'єкт дії та інші параметри, що характеризують дію).

ПІСЛЯМОВА

Структурну лінгвістику можна схарактеризувати як один із важливих етапів розвитку мовознавчої науки, що був започаткований у 1920-х рр., а після 1950-х рр. поступово трансформувався у різні напрямки, серед яких генеративна, математична, комп'ютерна, когнітивна, функціональна лінгвістика тощо. Виникнення лінгвістичного структуралізму, з одного боку, пов'язано з кризою лінгвістики наприкінці XIX ст., а точніше із неефективністю елементарного підходу до вивчення мови (вивчення окремої словоформи, окремої фонемі), домінуванням історичного аспекту опису мов (діахронії) із нехтуванням їх сучасного стану, а також хибним ототожненням понять «мова» і «мовлення». З іншого боку, наукові досягнення в галузі природничих наук дали можливість лінгвістам змінити не тільки сприйняття феномену мови, а й також методики її дослідження.

Вихідним положенням лінгвістичного структуралізму є те, що мова являє собою не стихійно сформовану сукупність її одиниць, а структурну організацію, у якій елементи пов'язані між собою певним чином, причому сукупність таких зв'язків характеризує мову як систему. Дослідження мови як системно-структурного явища відбувалося в різних парадигмах, унаслідок чого виникли кілька шкіл структуралізму: празька, копенгагенська та американська. Методичний апарат дослідження мови, розроблений представниками цих шкіл, досі широко використовуються сучасними лінгвістами (наприклад, метод опозицій, метод дистрибутивного аналізу тощо).

Перший етап розвитку структурної лінгвістики (1920-50-ті рр.) – класичний структуралізм – характеризувався підвищеною увагою до суто формального опису мови та ігноруванням її змістової частини, а також спробами відобразити систему мови як математично точну і правильну (універсальна лінгвістична теорія «Глосематика»). Починаючи з 1950-х рр., на зміну класичному структуралізму прийшов генеративізм, або генеративна лінгвістика, який мав розв'язувати актуальні на той момент завдання, що поступово виникали після другої світової війни у галузях машинного

перекладу, штучного інтелекту, автоматизованих системах управління тощо. Для успішного розв'язання цих завдань було розроблено математичний апарат, необхідний для моделювання інтелектуальної діяльності людини, що пов'язана з використанням мови. У той самий час активно розробляються також структурні методи опису значення / змісту, такі як компонентний аналіз, породжувальна семантика, інтерпретативна семантика.

У 1970-ті рр. структурна лінгвістика перестає існувати як окремий напрямок мовознавства, але її дослідницькі методи і надалі застосовуються в різних лінгвістичних дисциплінах, у тому числі у поєднанні з іншими науками. Наприклад, результатом поєднання з типологією стало виникнення структурної типології, що досліджує загальні закономірності будови окремих фрагментів системи мови загалом, зокрема лінгвістична типологія, мовні універсалії. Крім того, структурна лінгвістика сприяла широкому проникненню математичних методів дослідження у мовознавство, в результаті чого отримали значний розвиток інші сучасні напрями, такі як когнітивна, математична та комп'ютерна лінгвістика.

На сьогодні основні ідеї структурної лінгвістики також сприяють виникненню нових теоретичних схем, зокрема у вітчизняному мовознавстві, прикладом яких є теорія лексикографічних систем. На її основі вже створено віртуальні лексикографічні лабораторії: низку комп'ютерних систем для забезпечення мережевої професійної взаємодії лінгвістів. У цьому зв'язку слід зазначити, що структурна лінгвістика значно розширила методологічну базу мовознавства своїми методами і підходами, якими наразі фахівці користуються як для дослідження мовного матеріалу, так і для побудови формальних моделей, необхідних для створення комп'ютерних систем опрацювання природної мови.

ГЛОСАРІЙ

АБСТРАГУВАННЯ

Див. узагальнення.

АКТАНТ

Активний учасник дії (партиципant), «аналог» особи або предмета у висловлюванні.

АКТУАЛЬНЕ ЧЛЕНУВАННЯ РЕЧЕННЯ

Членування речення за змістом на дві змістові частини: перша називає предмет мовлення (тема) й те, що про нього говориться (рема), тобто якусь інформацію про цей об'єкт.

АЛОМОРФ

Варіанти морфеми, які можуть мати різні варіанти вимови та написання, без зміни їх значення. Приклади аломорфів: **друг** – **друзі**, **сад-ок** – **сад-к-а**, **учень** – **уч-н-я**.

ВАЛЕНТНІСТЬ

Здатність слова визначати кількість і якість залежних від нього словоформ, зумовлена його семантичними і граматичними властивостями.

ВІДМІНОК

Категорія іменних частин мови, служить для вираження граматичних і синтаксичних відношень між словами в реченні.

Семантичний відмінок: роль предметного імені (або партиципанта) у структурі речення.

ВІДНОШЕННЯ

Зв'язки, у яких перебувають одиниці в межах мови та які характеризують мову як систему.

Ієрархічні відношення: відношення входження простішої одиниці у складнішу (наприклад, входження фонему у морфему, морфеми – у слово).

Парадигматичні відношення: відношення у системі мови, за рахунок яких одиниці об'єднуються в категорії (групи, класи) на основі спільності та

протилежності їхніх характеристик. Спостерігаються на всіх рівнях мови: фонемному, морфемному, лексичному та синтаксичному рівнях.

Синтагматичні відношення: відношення у системі мови, що характеризують сполучуваність одиниць мови у більш складний лінгвістичний об'єкт (фонеми з фонемами, морфеми з морфемами, слова зі словами).

ГЕНЕРАТИВІЗМ

Напрям у структурній лінгвістиці, що виник у 50-ті рр. XX ст., предметом якого є питання створення формальних моделей процесів породження мовних конструкцій.

ГІПОТЕЗА

Твердження, що передбачає доказ, на відміну від аксіом та постулатів.

Див. також: *гіпотетико-дедуктивний метод*.

ГЛОСЕМАТИКА

Теорія празького структураліста Луї Єльмслева, що є спробою побудувати універсальну теорію мови, застосовуючи елементи математичної логіки. За нею, будь-яку мову, незалежно від її будь-яких фонетичних, граматичних та інших особливостей, можна описати як систему – мережу функцій або залежностей між елементами мови. Такими залежностями є інтердепенденція, детермінація та констеляція.

ГРАМАТИКА

Підсистема мови, що містить сукупність правил для породження цілісних об'єктів із мовних одиниць, враховуючи інформацію про їхні лексико-семантичні властивості зі словника.

Граматика складників: граматика, у яких на кожному кроці виведення замінюється лише один символ, тому в них із кожним виведенням асоціюється так зване дерево виведення.

Домінаційна граматика: граматика, що породжує безліч ланцюжків, інтерпретованих як речення разом із їхніми синтаксичними структурами у вигляді дерев підпорядкування.

Контекстно-вільна граматика: див. *породжувальна граматика типу 2*.

Контекстно-залежна граматика: див. *породжувальна граматика типу 1*.

Породжувальна граматика: абстрактний лінгвістичний механізм, що може генерувати безкінечну множину правильних речень за допомогою кінцевого числа граматичних правил.

Породжувальна граматика типу 0: граматика, яка не накладає жодних обмежень на вигляд її правил.

Породжувальна граматика типу 1: граматика, у якій можливість заміни ланцюжка елементів може визначатися контекстом.

Породжувальна граматика типу 2: граматика, правила якої дозволяють замінити нетермінальний елемент у лівій частині на будь-яку комбінацію термінальних та нетермінальних елементів без контекстних обмежень.

Породжувальна граматика типу 3: граматика, у якій ліва частина правил містить один нетермінальний елемент, а права частина – або термінальний, або термінальний, або термінальний та нетермінальний елемент, або порожній рядок.

Регулярна граматика: див. *породжувальна граматика типу 3*.

Розпізнавальна граматика: абстрактний лінгвістичний механізм, що імітує здатність відрізняти правильні речення від неправильних. На вході одержує деякий відрізок тексту природною мовою або його абстрактну репрезентацію штучною мовою і дає на «виході» відповідь, чи є цей відрізок граматично правильним або неправильним.

Універсальна граматика: 1. Набір правил чи принципів, властивих кожній людській мові. 2. Деяка система правил і принципів, за допомогою якої будуються осмислені речення будь-якої мови у світі.

Формальна граматика: будь-які «автоматичні пристрої» у вигляді числень або алгоритмів, що дають змогу задавати мову.

Див. також: *Задання мови*.

Δ-граматика: система правил перетворення дерев, інтерпретованих як «чисті» дерева підпорядкування речень, тобто дерева підпорядкування без лінійного порядку слів.

ГРАФ

Сукупність об'єктів зі зв'язками між ними. Наприклад, у вигляді графа можна репрезентувати дерево безпосередніх складників.

Див. також: *дерево безпосередніх складників*.

ДЕДУКЦІЯ

Спосіб міркування, за яким нове положення виводиться суто логічним шляхом від загальних положень до конкретних висновків.

Див. також: *індукція*.

ДЕРЕВО

Дерево безпосередніх складників: спосіб репрезентації синтаксичної структури речення у вигляді схеми або графа. Це дає змогу відобразити залежність елементів речення між собою.

ДЕТЕРМІНАЦІЯ

У глосематиці: одностороння залежність – має місце між двома елементами, один з яких передбачає існування іншого, але протилежне невірно (умовне позначення $x \rightarrow y$). Тобто елемент x наявний тоді, коли є елемент y , а не навпаки.

ДИНАМІКА СИНХРОНІЇ

Мережа глибинних відношень, пов'язаних із законами породження того чи іншого лінгвістичного об'єкта із найпростіших первинних елементів мови.

Див. також: *структурна лінгвістика (2)*.

ДИСТРИБУЦІЯ

Множина усіх можливих позицій одного елемента відносно інших елементів у тексті.

Див. також: *дистрибутивний клас*.

ДІАХРОНІЯ

Часовий зріз мови, при якому об'єктом лінгвістичного аналізу стає історичний розвиток мови.

ЕЛПС

Вид трансформації, що полягає у вилученні елемента зі складу речення, тобто прийом «згортання речення».

ЗАДАННЯ МОВИ

Встановлення вихідних синтаксичних конструкцій та правил, за якими породжують граматично правильні ланцюжки у вигляді речень або виразів.

ІНДУКЦІЯ

Висновок, у якому думка розвивається від знання одиничного до знання загального.

ІНТЕРДЕПЕНДЕНЦІЯ

У глосематиці: двостороння залежність між елементами в тексті або системі ($x \leftrightarrow y$). Тобто елемент x наявний тоді й тільки тоді, коли є елемент y .

КАТЕГОРІЯ

Граматична категорія: Низка протиставлених один одному однорідних граматичних значень, систематично виражених формальними показниками.

КВАЗІСТРУКТУРА

Мережа лінгвістичних відношень, що належать до статички синхронії.

Див. також: *статика синхронії*.

КЛАС

Дистрибутивний клас: клас елементів, що виникають в одному й тому самому оточенні.

Див. також: *дистрибуція*.

КОМПЕТЕНЦІЯ

Мовна компетенція: система уявлень про граматику мови, що притаманні людині.

КОНВЕРСІЯ

Спосіб вираження лексико-семантичних відношень зворотних за своїм значенням мовних одиниць у вихідному і похідному висловлюваннях.

КОНСТЕЛЯЦІЯ

У глосематиці: вільна залежність, що має місце між двома елементами, кожен з яких може існувати без іншого та не обумовлює іншого ($x - y$).

КОРЕЛЯТ

Лексичні кореляти: всі слова, які пов'язані з ключовим словом парадигматичними та синтагматичними відношеннями.

ЛІНГВІСТИКА

Діахронічна лінгвістика: лінгвістика, що вивчає відношення між елементами мови, які послідовно змінюють один одного в часі й не утворюють у своїй сукупності систему.

Синхронічна лінгвістика: лінгвістика, предметом якої є логічні та психологічні відношення, що пов'язують співіснуючі елементи у спільну систему. У межах СЛ ці відношення вивчаються так, як вони сприймаються однією й тією самою колективною свідомістю, тобто як одночасно існуючі.

Структурна лінгвістика: 1. Важливий розділ мовознавчої науки, предметом якої є репрезентація системи мови та її елементів у вигляді певних структур. 2. Наука про динамічний аспект синхронії мови.

Таксономічна лінгвістика: наука про статичний аспект синхронії мови.

Див. також: *статика синхронії*.

МЕТАМОДЕЛЬ

Модель, що описує структуру та принципи роботи іншої моделі.

МЕТОД

Гіпотетико-дедуктивний метод: метод науково-дослідної роботи, що полягає у виробленні припущень (гіпотез) та їх перевірка емпіричним шляхом, тобто шляхом спостережень, перевірок і експериментів.

Метод аналізу безпосередніх складників: прийом подання словотвірної структури слова і синтаксичної структури словосполучення та речення у вигляді ієрархії складових елементів.

Метод дистрибутивного аналізу: метод, що має на меті виявити порядок розташування (аранжування) і розподілу (дистрибуції) окремих частин висловлювання одна відносно одної.

Метод компонентного аналізу: метод, метою якого є розкладання значення на мінімальні семантичні складники – семи.

Див. також: *сема*.

Метод опозицій: метод, що полягає у протиставлянні одних мовних одиниць іншим за їхніми властивостями або за виконуваними функціями.

Трансформаційний метод: метод відображення синтаксичної структури речення, що ґрунтується на виведенні складних синтаксичних структур із простіших за допомогою невеликого набору правил перетворення (трансформацій).

МОВА

Складноструктурована система пов'язаних між собою елементів, яка служить для комунікації, зберігання й передавання інформації.

Ергативна мова: мова, в якій відбувається протиставлення більш активних, менш активних та самостійних дій.

Номінативна мова: мова, в якій будова речення спрямована на розрізнення об'єкта та суб'єкта дії.

МОДЕЛЬ

Матеріальний або уявний заміник реального об'єкта, що містить у собі його найсуттєвіші характеристики або властивості, на основі якого отримують нові знання про цей реальний об'єкт і формулюють прогнози щодо його майбутньої поведінки.

Аналітична модель: модель певної мови, що отримує на «вході» деякий відрізок тексту (зазвичай не менший, ніж висловлювання) і дає на «виході» його змістовий запис.

Див. також: *синтетична модель*.

Вербальна модель: модель, що репрезентується у мисленнєвій або розмовній формі у вигляді словосполучення, висловлювання або тексту.

Знакова модель: модель, що репрезентована за допомогою спеціальних знаків, тобто засобами будь-якої формальної мови.

Дедуктивна модель: модель, що формується на основі певної гіпотези та перевіряє істинність останньої на відповідному мовному матеріалі.

Дослідницька модель: модель, що імітує діяльність науковця-лінгвіста.

Індуктивна модель: модель, що розробляється на основі аналізу певного мовного матеріалу для формування гіпотези щодо будови та функціонування відповідного лінгвістичного об'єкта або явища. Див.: *дедуктивна модель*.

Інформаційна модель: модель, що являє собою сукупність інформації щодо стану та властивості об'єкта, процесу або явища.

Класифікаційна модель: модель, що відображає різновиди елементів того чи іншого об'єкта.

Комп'ютерна модель: модель, що реалізується засобами програмного середовища.

Лінгвістична модель: певний абстрактний механізм, відтворюваний формальними засобами, що за своєю будовою та принципом функціонування імітує поведінку реального суб'єкта, який використовує природну мову.

Лінгвістична глобальна модель: модель, що охоплює мову в цілому.

Лінгвістична специфічна модель: модель, що призначена для конкретної мови або групи мов.

Лінгвістична універсальна модель: модель, що охоплює або прагне охопити всі мови світу.

Лінгвістична часткова модель: модель, що описує лише окремий рівень мови.

Матеріальна модель: модель, що відтворює геометричні та фізичні властивості оригіналу та завжди має реальне втілення.

Модель аналізу: модель, що містить кінцеву кількість правил, за якими можна проаналізувати безліч речень певної мови.

Модель граматичної правильності: модель, що відображає уміння відрізняти правильне від неправильного в мові.

Модель лінгвістичних знань: концепція щодо будови та функціонування мови, притаманна тим чи іншим лінгвістичним школам (Женевська лінгвістична школа, Копенгагенська школа, Празький лінгвістичний осередок та інші) або окремим ученим.

Модель мови: модель, що використовується для опису мови та окремих її аспектів.

Модель наукової діяльності: певна система поглядів, що визначає підходи, методи та технічні засоби, які мають використовуватися для отримання нових знань.

Модель синтезу: модель, що надає скінчену кількість правил, за якими можна побудувати велику кількість правильних речень.

Описова модель: будь-який опис об'єкта.

Породжувальна модель: див. *породжувальна граматика*.

Прикладна модель: модель, що призначена для розв'язання практичних завдань.

Розпізнавальна модель: див. *розпізнавальна граматика*.

Синтетична модель: модель будь-якої мови, що на «вході» отримує семантичний запис (тобто уявлення деякого фрагмента смислу), а на «виході» формує множину синонімічних текстів цією мовою.

Див. також: *аналітична модель*.

Структурна модель: модель, що відображає складники об'єкта.

Теоретична модель: теоретичне та схематичне уявлення про реальність за допомогою ідеалізованих абстракцій або об'єктів.

Формальна модель: модель, побудована з використанням математичного апарату.

Функціональна модель: модель, що розкриває уміння співвідносити зміст мовлення (план змісту) з його формою (планом вираження).

МОДЕЛЮВАННЯ

Метод опосередкованого пізнання, за якого досліджуваний об'єкт-оригінал перебуває в певній відповідності з іншим об'єктом-моделлю, причому модель здатна в тому чи іншому відношенні заміщати оригінал на деяких стадіях пізнавального процесу.

Знакове моделювання: вид моделювання, що використовує у вигляді моделі певні знакові перетворення: схеми, графіки, малюнки.

Ідеальне моделювання: вид моделювання, заснований не на матеріальній аналогії об'єкта, а на аналогії ідеальної, уявної.

Математичне моделювання: вид моделювання, під час якого дослідження об'єкта відбувається за допомогою моделі, побудованої мовою математики (наприклад, дослідження законів механіки Ньютоном за допомогою математичних формул).

Матеріальне моделювання: моделювання, під час якого реальному об'єкту протиставляється його збільшена або зменшена копія, яка дає змогу виконувати дослідження в лабораторних умовах з подальшим перерахуванням характеристик процесів або явищ з моделі на об'єкт.

МОЛОДОГРАМАТИКИ

Німецька мовознавча школа (1870-і рр.), що запропонувала природно-наукові принципи для верифікації фактів мови.

МОРФОНЕМА

Ряд фонем, що чергуються в аломорфах однієї морфеми. Приклади морфонеми: **е/і** (веселий – весілля), **е/0** (палець – пальця), **е/о** (женити – жонатий, шести – шостий), **о/и** (кров – кривавий), **к/ч/ц** (мука – мучний – муці), **д/дж** (радити – раджу), **з/ж** (возити – вожу), **х/ш** (рухати – рушу).

НЕГАЦІЯ

Вид трансформації, під час якої речення перетворюється у заперечувальне.

ОДИНИЦЯ МОВИ

Постійний елемент, що виконує свою певну функцію і має своє особливе призначення.

Знакова одиниця мови: одиниця, що має план вираження (матеріальне вираження) та план змісту (має певне значення). Такою одиницею є слово.

Конститутивна одиниця мови: базова одиниця у системі мови: фонема, морфема, словоформа, слово, словосполучення, речення.

Незнакова одиниця мови: одиниця, у якій є план вираження (матеріальне вираження), але немає плану змісту (певного значення). До таких одиниць належать фонема, морфема та синтаксема.

Неконститутивна одиниця мови: одиниця, яка виконує внутрішньо- та міжривневі функції (наприклад, наголос, інтонація, фонетичне слово).

Комунікативна одиниця мови: відрізок мовлення, що може самостійно передавати повідомлення.

Номінативна одиниця мови: одиниця, що виконує функцію найменування та повідомлення про предмети, процеси і властивості.

Стройова одиниця мови: одиниця мови, що служить для побудови номінативних та комунікативних одиниць.

ОПОЗИЦІЯ

Граматична опозиція: протиставлення граматичних значень (наприклад, значення однини, протиставлене значенню множини).

Фонологічна опозиція: звукове протиставлення, за допомогою якого можна диференціювати значення слів.

ПАРАДИГМА

Наукова парадигма: модель організації наукової діяльності, прийнята науковим суспільством.

ПАРТИЦИПАНТ

Учасник комунікативної ситуації, що відображається у висловлюванні.

Див. також: *актант* і *сирконстант*.

ПАСИВІЗАЦІЯ

Вид трансформації, під час якої утворюються пасивні синтаксичні конструкції.

ПОЛЕ

Семантичне поле: сукупність мовних одиниць, об'єднаних спільністю лексичного значення.

ПОЛІПАРАДИГМАТИЗМ

Багатоманіття моделей або напрямків наукової діяльності, метою яких є отримання нових знань про мову (сукупності мов) у різних її аспектах із використанням відповідних методів та засобів.

РЕКОНСТРУКЦІЯ

Комплекс прийомів та процедур відновлення незасвідчених мовних станів, форм, явищ шляхом історичного порівняння відповідних одиниць окремої мови, групи та сім'ї мов. У структурній лінгвістиці реконструкція використовується для відтворення прадавньої системи мови.

РЕЧЕННЯ

Ядерне речення: просте активне стверджувальне речення, що описує елементарну ситуацію.

СЕМА

Мінімальна одиниця лексичного значення.

Див. також: *метод компонентного аналізу*.

Диференційна сема: сема, за якою певне слово змістовно відрізняється від інших слів.

Інтегральна сема: сема, за якою слова об'єднуються в семантичні поля.

Див. також: *семантичне поле*.

СЕМАНТИКА

Генеративна семантика: напрямок у структурній лінгвістиці, що характеризується побудовою моделі мови за схемою «від значення до тексту» і «від тексту до значення».

СИМВОЛ

Нетермінальний символ: об'єкт, що позначає певну сутність мови та може набувати те чи інше символічне значення. Наприклад: нетермінальний елемент *N*, що позначає іменник, може набувати таких значень: «собака», «людина», «дім» тощо.

Див. також: *термінальний символ*.

Термінальний символ: об'єкт, безпосередньо існуючий в мові (слово). Так само як і нетермінальний символ, використовується для задання правил породжувальної граматики.

Див. також: *нетермінальний символ*.

СИНКРЕТИЗМ

Явище, коли в одній формі одночасно об'єднано кілька значень або компонентів значень (наприклад, *бабах* може означати «удар» і звук «звук удару»).

СИНТАГМА

Частина фрази, що виділяється ритміко-мелодійними засобами та вимовляється на одному диханні.

СИНТАКСИС

Генеративний синтаксис: подальший напрямок розвитку структурної лінгвістики, що зосереджений на моделюванні синтаксичних структур.

СИНХРОНІЯ

Умовне виокремлення певного історичного етапу в розвитку мови, який обирається як об'єкт лінгвістичного дослідження.

СИРКОНСТАНТ

Додаткові учасники комунікативної ситуації, що відображаються у висловлюванні.

Див. також: *актант* і *партиципант*.

СИСТЕМА

Лексикографічна система: спосіб формальної репрезентації будь-якої інформаційної системи, принцип функціонування якої схожий на природну мову.

Система мови: 1. З погляду класичного структуралізму: сукупність таких відношень, завдяки яким різноманітні елементи утворюють мову як єдине ціле. 2. З погляду генеративної лінгвістики та інформаційної теорії мови: абстрактний механізм, що породжує лінгвістичні об'єкти будь-якого ступеня складності, маючи в основі три важливих компоненти: структуру, суб'єкт та субстанцію.

СЛОВНИК

Підсистема мови, що містить інформацію про семантичні та інші властивості, що впливають на утворення граматикою лінгвістичного об'єкта.

СЛОВО

Фонетичне слово: фрагмент мовленнєвого потоку, об'єднаний одним словесним наголосом.

СЛОВОЗМІНА

Сукупність граматичних форм тієї самої лексики, що утворюються шляхом додавання певного набору закінчень (квазіфлексій) до її незмінюваної частини (квазіоснови).

СТАТИКА СИНХРОНІЇ

Мережа таксономічних відношень у системі мови.

Див. динаміка синхронії.

СТРАТИФІКАЦІЯ

Поділ мови на рівні, що репрезентують сукупності конструктивних одиниць, таких як фонеми, морфеми, слововиди, слова, словосполучення, речення.

СТРУКТУРА

Глибинна структура: 1. Теоретична конструкція, що прагне пов'язати низку речень, спільних за своїм змістом, але з різними синтаксичними конструкціями. 2. Спосіб абстрактного опису семантичної структури речення; абстрактна формула, утворена найбільш загальними, універсальними елементами змісту.

Дискретна структура: матеріальна структура, яку складають окремі елементи.

Поверхнева структура: 1. Реалізація глибинної структури у вигляді речення з певною синтаксичною будовою. 2. Засіб вираження глибинної синтаксичної семантики граматичними класами слів, тобто частинами мови.

Семантична структура: конструкція речення, що передбачає замикання дії предиката на суб'єкті.

СТРУКТУРАЛІЗМ

Напрямок у мовознавстві, що розглядає мову як чітко структуровану систему, елементи якої перебувають у певних відношеннях між собою та набувають свого місця у системі завдяки цим відношенням.

СТРУКТУРА

Структура мови: 1. З погляду класичного структуралізму: Сукупність елементів і підсистем мови, які організовують єдине ціле за допомогою різних відношень. 2. З погляду генеративної лінгвістики та інформаційної теорії мови: головний елемент системи, який в абстрактному вигляді задає будову породжуваного лінгвістичного об'єкта, наприклад: речення, словозмінної парадигми, словоформи тощо. 3. Мережа лінгвістичних відношень, що належать до динаміки синхронії.

Див. також: *динаміка синхронії*.

СУБ'ЄКТ

В інформаційній теорії мови: елемент системи, здатний ідентифікувати, інтерпретувати структурно-субстанційні властивості й елементи системи та здійснювати з ними операції за певними правилами та алгоритмами. Суб'єкт може бути як людина (або спільнота загалом), так і комп'ютерна система.

СУБСТАНЦІЯ

В інформаційній теорії мови: елемент системи, тобто набір мовних елементів, необхідних для побудови лінгвістичного об'єкта (наприклад, граматичної форми, речення).

СУБСТИТУЦІЯ

Вид трансформації, під час якої один елемент речення замінюється на інший.

СУПЛЕТИВІЗМ

Утворення граматичних форм одного й того ж слова від різних основ або коренів.

ТИПОЛОГІЯ

Мовна типологія: дисципліна, що має своїм завданням систематизацію структурних типів конкретних мов світу.

Мовна морфологічна типологія: спосіб класифікації мов світу, який групує мови відповідно до їхніх спільних морфологічних структур. За цією

класифікацією мови світу поділяються на флективні, аглютинативні, ізольовані та полісинтетичні.

Мовна синтаксична типологія: спосіб класифікації мов світу, який групує мови залежно від типу будови речень. За нею мови світу поділяються на мови активного строю, мови ергативного строю та мови номінативного строю.

Мовна фонологічна типологія: спосіб класифікації мов світу, який групує мови відповідно до їх фонологічних особливостей.

ТРАНСФОРМА

Просте нерозповсюджене стверджувальне речення активного стану, на основі якого можливо побудувати різні речення.

ТРАНСФОРМАЦІЯ

У теорії породжувальних граматик: правило, за яким із ядерного речення утворюється похідне речення.

УЗАГАЛЬНЕННЯ

Результат розумової діяльності, який відображає загальні ознаки або якості тих чи інших об'єктів або явищ дійсності.

ФУНКЦІЯ

Термін, що не має однозначного тлумачення серед представників тієї чи іншої школи структуралізму. Так, празькі структуралісти використовують його тоді, коли говорять про значення (функція слова, функція речення) або про структуру смислових одиниць (функція фонем). У розумінні копенгагенських структуралістів: математичний вираз залежності між одиницями мови.

ЗМІСТ

ВСТУП	3
МОДУЛЬ 1 ІСТОРІЯ СТРУКТУРНОЇ ЛІНГВІСТИКИ	5
ТЕМА 1 ПЕРЕДУМОВИ ВИНИКНЕННЯ СТРУКТУРАЛІЗМУ	5
1.1. Сучасні наукові парадигми у лінгвістиці	5
1.2. Структурний підхід у лінгвістиці	8
1.3. Основні принципи лінгвістичного структуралізму	12
ТЕМА 2 ШКОЛИ СТРУКТУРАЛІЗМУ ТА ЇХ РОЛЬ У СУЧАСНІЙ ЛІНГВІСТИЦІ	19
2.1. Празька лінгвістична школа	19
2.2. Копенгагенський структуралізм (гloseматика)	26
2.3. Американський дескриптивізм	30
ТЕМА 3 ДОСЛІДЖЕННЯ МОВИ З ПОГЛЯДУ СИСТЕМИ ТА СТРУКТУРИ	35
3.1. Мова як системно-структурне утворення	35
3.2. Конститутивні та неконститутивні одиниці мови	42
3.3. Основні напрямки вивчення системи мови	48
МОДУЛЬ 2 СТРУКТУРНА ЛІНГВІСТИКА ЯК ДИСЦИПЛІНА	52
ТЕМА 4 СТРУКТУРНА ЛІНГВІСТИКА ТА ПРЕДМЕТ ЇЇ ВИВЧЕННЯ	52
4.1. Структурна лінгвістика: її предмет, завдання і методи	52
4.2. Структура і система: співвідношення понять	62
4.3. Граматика і словник як дві підсистеми мови	68
ТЕМА 5 СТРУКТУРНА І ТАКСОНОМІЧНА ЛІНГВІСТИКА	74
5.1. Синхронія та діяхронія. Статика і динаміка	74
5.2. Специфіка структурної і таксономічної лінгвістики	80
5.3. Системна тріада: структура, суб'єкт і субстанція	85
ТЕМА 6 СТРУКТУРНА ЛІНГВІСТИКА СЕРЕД ІНШИХ ГАЛУЗЕЙ МОВОЗНАВСТВА	91
6.1. Структурна лінгвістика і порівняльне мовознавство	91
6.2. Структурна лінгвістика і навчання мові	101
6.3. Структурна лінгвістика і процес перекладу	109
МОДУЛЬ 3 ЛІНГВІСТИЧНІ МОДЕЛІ	112
ТЕМА 7 МОДЕЛІ В ЛІНГВІСТИЧНИХ ДОСЛІДЖЕННЯХ	112
7.1. Поняття моделювання і моделі	112

7.2. Гіпотетико-дедуктивний метод	118
7.3. Моделювання в лінгвістиці	124
ТЕМА 8 ПОБУДОВА ЛІНГВІСТИЧНИХ МОДЕЛЕЙ	132
8.1. Лінгвістична модель та вимоги до неї	132
8.2. Лінгвістична модель і «чорна скринька»	138
8.3. Моделі у наукових дослідженнях	144
ТЕМА 9 КЛАСИФІКАЦІЯ ЛІНГВІСТИЧНИХ МОДЕЛЕЙ	151
9.1. Базова класифікація	151
9.2. Класифікація за об'єктом моделювання	157
9.3. Класифікації за способами побудови або аналізу мовних об'єктів	164
МОДУЛЬ 4 ФОРМАЛЬНІ ГРАМАТИКИ І СЕМАНТИКИ	167
ТЕМА 10 ФОРМАЛЬНІ МОВИ І ФОРМАЛЬНІ ГРАМАТИКИ	167
10.1. Поняття формальної мови та формальної граматики	167
10.2. Породжувальна граMATика як один із типів формальних граматик	175
10.3. Історичний погляд на формування універсальної граматики	183
ТЕМА 11 ПОБУДОВА МОДЕЛЕЙ ПОРОДЖУВАЛЬНОЇ ГРАМАТИКИ	188
11.1. Елементи моделі породжувальної граматики	188
11.2. Типологія породжувальних граматик	193
11.3. Трансформації в межах генеративної лінгвістики	200
ТЕМА 12 МОДЕЛІ У ГЕНЕРАТИВНІЙ СЕМАНТИЦІ	206
12.1. Модель «Смисл ↔ Текст»	206
12.2. Модель преференційної семантики	215
12.3. Модель концептуальної залежності	226
ПІСЛЯМОВА	233
ГЛОСАРІЙ	235
БІБЛІОГРАФІЯ	253

БІБЛІОГРАФІЯ

1. «Словники України» on-line / Український мовно-інформаційний фонд НАНУ. URL: <http://lcorp.ulif.org.ua/dictua/>.
2. Бацевич Ф. Філософсько-методологічні засади сучасної лінгвістики: спроба обґрунтування. *Мовознавство*. 2006. № 6. С. 33–40.
3. Васильєв О., Чалий О., Васильєва І. Математичні методи й моделі в лінгвістиці. *Україна Модерна*. 2019. №27. С. 9–28.
4. Великий електронний словник української мови (ВЕСУМ). URL: <https://r2u.org.ua/vesum/>.
5. Вихованець І. Р., Городенська К. Г., Русанівський В. М. Семантико-синтаксична структура речення .К.: Наук. думка, 1983. 219 с.
6. Габідулліна А. Р., Колесніченко О. Л. Методологія сучасних лінгвістичних досліджень: навчальний посібник. Слов'янськ: Вид-во Б. І. Маторіна, 2019. 322 с.
7. Дарчук Н. П. Традиційна лінгвістика – структурна лінгвістика – комп'ютерна лінгвістика – триєдина сутність. Вісник київського національного університету імені Тараса Шевченка, серія «літературознавство, мовознавство, фольклористика», № 1(26), 2016. С. 24 – 27.
8. Дудок В. І. Системність і структурність мови. *Мовознавство*. 2003. № 5. С. 26–33.
- Загнітко А.П. Сучасний лінгвістичний словник. Вінниця: «Твори», 2020. 920 с.
9. Йосипенко О. Структуралізм // Філософський енциклопедичний словник / В. І. Шинкарук (гол. редкол.) та ін. Київ: Інститут філософії імені Григорія Сковороди НАН України: Абрис, 2002. С. 611.
10. Карпіловська Є. А. Здобутки академічної структурної та математичної лінгвістики у моделюванні українського слова. *Українська мова: наук.-теор. журн.*, 2019 № 1. С. 18–35.
11. Кочерган М. П. Загальне мовознавство: підр. для студ. філол. спец. ВНЗ. 2. вид., випр. і доп. К.: Вид. центр «Академія», 2006. 463 с.

12. Купріянов Є. В. Лексикографічна система іспанської мови. Феноменологія інтегрального опису. К.: УМІФ НАНУ, 2018. 254 с.

13. Купріянов Є. В., Юрченко О. М., Гулієва Д. О. Методичні вказівки з виконання і захисту наукових рефератів з дисципліни «Структурна лінгвістика» для студентів спеціальності 035 «Філологія». Х: Вид-во НТУ «ХП», 2019. 27 с.

14. Лінгвістично-інформаційні студії: праці Українського мовно-інформаційного фонду НАН України: у 5 т. / В.А. Широков та ін. Т. 1: Наукова парадигма та основні мовно-інформаційні структури. Київ. Український мовно-інформаційний фонд НАН України. 2018. 271 с.

15. Лінгвістично-інформаційні студії: праці Українського мовно-інформаційного фонду НАН України: у 5 т. / В. А. Широков та ін. Т. 2: Граматичні системи. Київ. Український мовно-інформаційний фонд НАН України. 2018. 300 с.

16. Лінгвістично-інформаційні студії: праці Українського мовно-інформаційного фонду НАН України: у 5 т. / В. А. Широков та ін. Т. 3: Тлумачна лексикографія, Кн. 2: Системна семантика тлумачних словників. Київ. Український мовно-інформаційний фонд НАН України. 2018. 250 с.

17. Методи лінгвістичних досліджень. URL: <https://discourse.com.ua/lekcii/metodi-lingvistichnih-doslidzhen/>

18. Перебийніс В. С. Напрямки і школи в сучасній структурній лінгвістиці // Проблеми та методи структурної лінгвістики. К, 1969. Вип. 2. С 5–25.

19. Повторева С. М. Структурний підхід – структуралізм – постструктуралізм (еволюція методології та її поширення у гуманітарних студіях). Монографія. Львів: Видавництво Львівської політехніки, 2010. 336 с.

20. Рабулець О. Г., Сухарина Н. М., Широков В. А., Якименко К. С. Дієслово в лексикографічній системі. Київ: Довіра, 2004. 259 с.

21. Селіванова О.О. Сучасна лінгвістика: термінологічна енциклопедія / Олена Олександрівна Селіванова. Полтава: Довкілля, 2006. 716 с.

22. Соссюр Ф. де. Курс загальної лінгвістики / Ф. де Соссюр. К.: Основи, 1998. 324с.

23. Структурні методи в лінгвістиці. URL:
http://linguist.univ.kiev.ua/courses_methods.htm
24. Сухорольська С. М. Методи лінгвістичних досліджень: навч. посібник / С. М. Сухорольська, О. І. Федоренко. Львів: Інтелект-захід, 2009. 348 с.
25. Сучасні інформаційні технології в лінгвістиці: навч. посіб. / С. В. Петрасова, Н. Ф. Хайрова. Харків: ФОП Панов А.М., 2020. 124 с.
26. Український правопис, 2019. URL:
<https://mon.gov.ua/storage/app/media/zagalna%20serednya/%202019.pdf>
27. Широков В. А. Комп'ютерна лексикографія. Київ: Наук. думка, 2011. 352 с.
28. Широков В. А. Лінгвістика і системність мови. *Актуальні проблеми прикладної лінгвістики*: матеріали IV Міжнародної Інтернет-конференції. URL:
http://prlingv.at.ua/publ/statti_2014_rik/v_a_shirokov_lingvistika_i_sistemnist_movi/2-1-0-19
29. Широков В. А. Феноменологія лексикографічних систем. Київ : Наук. думка, 2004. 326 с.
30. Широков В. А., Рабулець О. Г. Формалізація в галузі лінгвістики. *Актуальні проблеми української лінгвістики*. Київ, 2002. Вип. V. С. 3–28.
31. Широков В. А., Шевченко І. В. Граматика у феноменологічному вимірі. *Мовознавство*, 2014. № 4. С. 3–27.
32. Anderson John M. Structuralism and autonomy: from Saussure to Chomsky. *Historiographia Linguistica*. 2005. 32 (1). P. 117–148.
33. Blevins James P. American descriptivism ('structuralism'). *The Oxford Handbook of the History of Linguistics* / P. Blevins James. Oxford University Press, 2013. P. 418–437.
34. Bosque, I. Las categorías gramaticales. Relaciones y diferencias: Madrid: Síntesis, 2015. 238p.
35. Brunot, F. La pensée et la langue. Méthodes, principes et plan d'une théorie nouvelle du langage appliquée au français: Paris: Masson, 1922. 1004p.

36. Chomsky N. Syntax Structures. Berlin – New York: Mouton de Gruyter, 2002. 117 p.
37. Chomsky N. Three models for the description of language. IRE Trans. Inform. Theory. 1956. Vol. 2, № 3. P. 113–124.
38. Chomsky, N. Aspects of the theory of syntax: Massachusetts, Cambridge: THE M.I.T. PRESS, 1965. 258p.
39. Chomsky, N. Remarks on Nominalization: Readings in English Transformational Grammar / eds. R. A. Jacobs, P. S. Rosenbaum. Boston, Ginn, 1970. P. 184-221.
40. Collins English Dictionary: / eds. J. Crozier. HarperCollins, 2006. 1014p.
41. Conceptual dependency theory. URL: <https://people.cs.pitt.edu/~wiebe/courses/CS3730/Spring2013/cd.pdf>.
- Coseriu, E. Semántica y gramática: Gramática, semántica, universales. Madrid, 1978. P. 128-147.
42. Dosse F. History of Structuralism, Vol.2: The Sign Sets. Minnessota: University of Minnessota Press, 1967. 543 p.
43. Morrill G. Categorical grammar. Logical syntax, semantics, and processing. Oxford Univ. Press, 2011. 92 p.
44. Oxford Living Dictionaries. URL: <https://en.oxforddictionaries.com>.
45. Saussure, F. De. Cours de Linguistique Générale: / eds.P. de Tullio de Mauro. Paris: 1916. 336p.
46. Shirokov V. Language. Information. System. К.: Академперіодика. 2021. 164 с.
47. Structural linguistics / The Concise Dictionary of Linguistics (3rd ed.) // P. H Matthews. Oxford: Oxford University Press, 2014. 443 p.
48. Webster's New World College Dictionary: Houghton Mifflin Harcourt, 2016. 1728 p.

Навчальне видання

КУПРІЯНОВ Євген Валерійович
УГОЛЬНІКОВА Наталія Сергіївна
ЮРЧЕНКО Олена Миколаївна

СТРУКТУРНА ЛІНГВІСТИКА В ТЕОРІЇ ТА ПРАКТИЦІ

Навчальний посібник
для бакалаврів напряму підготовки
035 «Філологія»
денної форми навчання

Відповідальний за випуск	проф. Шаронова Н. В.
Роботу до видання рекомендував	проф. Кіпенський А. В.

В авторській редакції

План 2024 р., поз. 47

Підп. до друку 2024р. Гарнітура Times New Roman.

Видавничий центр НТУ «ХП».

Свідоцтво про державну реєстрацію ДК № 5478 від 21.08.2017 р.

61002, Харків, вул. Кирпичова, 2

Електронне видання