Богдан Бабич

Київ, Український мовно-інформаційний фонд НАН України Лексична семантика у синтаксичній структурі тексту: формальне представлення та інтерпретація

Словоформи в реченні утворюють лінійну послідовність. Крім цього, між ними існує система імпліцитних синтаксичних зв'язків (поверхнева синтаксична структура) та семантико-синтаксичних відношень (глибинна синтаксична структура), що має нелінійний характер¹. Цю систему можна графічно представити як мережу зв'язків між лексемами та їх групами; зв'язки можуть бути позначені назвами семантиковідношення і формально-синтаксичного зв'язку. Якщо конструкція представлена різними синтаксичними трансформаціями (напр., перевіряти завдання → перевірка завдань; перевірені завдання), то позначення типів синтаксичного зв'язку буде змінюватись залежно від виду трансформації (об'єктний зв'язок трансформувався в композитивний та в атрибутивний), але незмінним залишатиметься тип семантико-синтаксичного відношення (у даному прикладі – відношення "дія – об'єкт"). У більшості випадків поверхнева та глибинна синтаксичні структури розрізняються лише позначенням дуг на синтаксичних графах, але графи цих структур мають тотожну форму². Назви синтаксичних зв'язків мають формальний характер, а позначення семантико-синтаксичних відношень (ССВ) семантичний, тобто їх можна співвіднести з деякими позамовними реаліями. На рівні семантико-синтаксичних відношень відбувається "взаємодія лексики і граматики"3. Конкретніше, кожне семантико-синтаксичне відношення задає контекст, у якому функціонують поєднані елементи, отримуючи додаткові, зумовлені цим контекстом значення (напр., значення суб'єкта, об'єкта, інструмента дії, місця дії і т. ін.) – т. зв. структурно-синтаксична номінація, що з відрізняється з одного боку від лексичної, а з іншого – від лексико-синтаксичної номінації, яка притаманна ідіоматичним фразеологізмам типу брати участь, гострий на слово, дати перцю, де синтаксична конструкція функціонально тотожна лексемі⁴.

При структурно-синтаксичній номінації лексеми та їх групи у відповідних позиціях стають синтаксемами, виражаючи значення, не властиві їм до цього у лексичній системі мови. На графах синтаксичної структури речення вказується назва та напрямок ССВ (як правило, від залежного елемента до головного⁵). Синтаксеми є функціональними значеннями аргументів даного ССВ, а отже, вони є похідними від внутрішньої структури відношення. Наприклад, відношення дія — об'єкт; дія — інструмент; предикативне відношення передбачають процесуальну і відповідно об'єктну, інструментальну та суб'єктну синтаксеми. Атрибутивне та посесивне відношення передбачають атрибутивну та посесивну синтаксеми у поєднанні з абстрактною синтаксемою, тип якої визначається функціонуванням групи, що сформована за допомогою даного ССВ, на вищих щаблях структури речення.

Традиційно наголошується на такому аспекті синтаксичної семантики як (1) здатність синтаксем виражати функціональні, синтаксично зумовлені значення. Проте, у складі ССВ отримують вираження інші значення, які (так само як і синтаксемні) можуть бути співвіднесені з лексичними значеннями. Це зокрема:

- (2) "зовнішні", контекстуально незалежні граматичні значення елементів, поєднаних даним відношенням та лексико-граматичні і лексико-семантичні розряди цих елементів (т. зв. категоріальні значення). У випадку, якщо елементом є пексема це її граматичні значення та відповідні розряди. У випадку, якщо елементом є група із внутрішньою синтаксичною структурою, то її зовнішні категоріальні значення будуть або (2а) наслідуватись від одного з елементів цієї групи (напр., у посесивному ССВ {вступне слово} письменника група вступне слово наслідує зовнішні граматичні значення та лексичні розряди від іменника слово), або (2б) утворюватись за певними принципами, характерними для групи даного типу (напр., зв'язок узгодження при атрибутивному відношенні недільні {ранок і день} стає можливим завдяки тому, що хоча граматичне значення множини не притаманне жодному з елементів групи ранок і день, яка є аргументом цього відношення, проте єднальний сполучник формує групу зі значенням множини не залежно від граматичного значення числа елементів, поєднаних за його допомогою.
- (3) зовнішнє категоріальне значення структури, поєднаної даним ССВ, в цілому. Воно утворюється за аналогічним принципом, що і значення типу (2), і в загальній синтаксичній структурі тексту визначає спосіб функціонування групи. Якщо група пізніше ввійде до складу іншого ССВ на вищому щаблі синтаксичної структури речення, то значення типів (2) і (3) функціонують ідентично.

Семантико-синтаксичні відношення виконують функцію формування позицій для синтаксем та конструктивну функцію при побудові структури речення. Виконання останньої функції передбачає врахування зовнішніх категоріальних значень лексем та їх груп і способів наслідування цих значень синтаксичними структурами вищих рівнів, оскільки аргументом-синтаксемою певного ССВ можуть бути лише лексеми та їх групи із заданим набором зовнішніх не залежних від контексту категоріальних характеристик. Такі характеристики та позиція елемента в реченні є критеріями для формування мережі ССВ та пошуку їхніх аргументів. Для синтаксичного аналізу важливими є синтаксично зумовлені значення мовних конструкцій та їх контекстуально незалежні значення. Бачимо, що в традиційному розумінні поняття синтаксема поєднано два різні аспекти: синтаксичні функціональні значення (напр., об'єкт або суб'єкт дії) та категоріальні значення деяких синтаксичних структур, що не залежать від контексту (напр., локатив, темпоратив, тощо). У синтаксичних репрезентаціях їх доцільно розрізняти. Наприклад, речення: Коли якийсь з час з тому 4 B_5 Києво-Могилянській академії, ∂e_8 читалися курси $_{10}$ хорватської мови $_{12}$, виникли $_{13}$ проблеми $_{14}$ $_{315}$ викладачем $_{16}$, посол $_{17}$ Хорватії $_{18}$ Джуро $_{19}$ Відмарович $_{20}$ виконував $_{21}$ ще $_{22}$ й $_{23}$ такі $_{24}$ функції $_{25}$ — має таку синтаксично-синтаксемну структуру

(суфікс $_{.B}$ позначає назви ССВ, $_{-.B.}$ — відсутність семантико-синтаксичного відношення внаслідок неповнозначності одного з елементів, суфікс $_{.3}$ — назву синтаксичного зв'язку, якщо поверхнева структура відрізняється від глибинної, суфікс $_{.K}$ — категоріальне контекстуально незалежне значення утвореної структури, $_{-.K}$ позначає відсутність категоріального значення у випадку приєднання вставних структур):

```
A_{(2-4)} - [meмпоративн. B. \{(ampuбутивн. B. ЯКИЙСЬ ЧАС)_{oб'єкт. K}\} тому] темпоратив. K.; E_{(5-7)} - [-.B. 6] \{(ampuбутивн. B. KUЄво-Могилянській академіі)_{oб'єкт. K.}\} локативн. K.; B_{(9-12)} - [nped-cyō'єкт. 3./дія-об'єкт. B. Читалися <math>\{\kappa_{OMno3um.3.../ampuбутивн. B.} курси (ampuбутивн. B. авикладачем) комітатив. <math>K. об'єкт. K. лодія. K.; I_{(13-16)} - [nped-cyō'єкт. 3./дія-об'єкт. B. виникли <math>\{meматив. B. npoблеми (-.B. 3 викладачем) комітатив. <math>K. об'єкт. K. лодія. K.; I_{(17-20)} - [nom. B. \{(\kappaomno3um.3.../ampuбутив. B. nocon Xopbamii)_{oб'єкт. K.}\} (nom. B. Джуро Bidmapobuu)_{oб'єкт. K.} об'єкт. K. об'єкт. K.; I_{(17-25)} - I_{(1
```

Використаний формалізм систем синтаксемних груп також імпліцитно репрезентує синтаксемні значення, оскільки кожному типу ССВ відповідає впорядкований набір аргументів з їх синтаксичними функціями, які можна експлікувати. Запропоновані репрезентації дозволяють моделювати взаємодію синтаксичної та лексичної семантики у реченні.

Лінгвістичні моделі семантичної структури речення можна віднести до двох типів. Перший тип становлять т.зв. *атомістичні семантичні моделі*, що ставлять за мету експлікувати "атоми смислу" — семи, семантичні примітиви та їх структури, що імпліцитно присутні в лексемах і актуалізуються в певних контекстах 6 . Атомістичні семантичні моделі передбачають співвіднесення мовних структур з представленням смислу, що є інваріантним для усіх ситуацій вжитку.

Другий тип становлять *процедурно-інтерпретаційні моделі* лексичної семантики, які передбачають, що текст описує *процес* побудови деякої ментальної репрезентації екстралінгвістичного знання⁷. Згідно з цією точкою зору мовні структури не описують зміст, що утворюється в результаті розуміння тексту, а контролюють хід процесу розуміння, здійснюючи в певній послідовності доступ до різних ділянок позамовних знань, вже наявних у ментальній моделі на момент розпізнання змісту деякої мовної структури, доповнюючи та коригуючи ці знання.

Процедурно-інтерпретаційні моделі семантики пояснюють ширше коло мовних явищ, ніж атомістичні. По-перше, той самий текст може бути використаний у різних комунікативних ситуаціях, отримуючи різне осмислення. Згідно з процедурно-

інтерпретаційною теорією можна пояснити, чому це є можливим: у різний час текст взаємодіє з різними структурами екстралінгвістичного знання, і, відповідно, кожного разу спричиняє різний процедурний ефект, по-іншому будуючи, доповнюючи або змінюючи ментальні моделі. По-друге, в різних комунікативних ситуаціях текст часто потребує різних способів перекладу на іншу мову та передбачає різні способи зняття синтаксичної і лексико-семантичної омонімії, що означає, що конкретне осмислення тексту утворюється кожен раз знову в процесі розуміння. Кількість можливих осмислень потенційно безмежна, а їх збіг випливає лише зі спільності знань, які було використано під час розуміння.

Термін ситуаційний тезаурус (СТ) вживається в даній роботі для позначення формалізованої моделі позамовних знань про деяку ситуацію, у якій вжито текст. Компонентами ситуаційного тезауруса є об'єкти, відношення, та процеси. Об'єкти мають назву, набір атрибутів, набір методів для взаємодії з іншими структурами та життєвий цикл, що складається зі станів даного об'єкта та умов переходу з одного стану в інший, в т.ч. умов утворення та знищення даного об'єкта. Процесами є генератори подій, процеси-конструктори та деструктори об'єктів, тощо. Реагуючи на події, об'єкти СТ можуть змінювати стани свого життєвого циклу, змінювати значення атрибутів, утворювати певні відношення з іншими об'єктами. Ситуаційний тезаурує сам по собі не має лінгвістичного значення, але оскільки він використовується в процесі розуміння тексту, лінгвістичне значення має проблема їхньої взаємолії.

Можна показати, що формально-синтаксична структура тексту задає послідовність і характер (тобто здійснює вибір аргументів для процедур обробки СТ серед синтаксичних структур) доступу мовних конструкцій до знань із ситуаційного тезауруса. Семантико-синтаксичне та лексико-семантичне наповнення синтаксичних репрезентацій специфікує посилання на назви, методи конкретних об'єктів, на конкретні відношення та процеси в активному фрагменті СТ, тощо. Ті самі мовні конструкції можуть взаємодіяти не з фіксованим, а з потенційно необмеженим набором об'єктів, тому кожного разу лексеми і синтаксеми оформлюють доступ до об'єктів різного типу, для яких можливий різний набір атрибутів та методів, що зумовлює появу контекстуальних значень таких мовних конструкцій.

Лексико-семантичні варіанти повнозначних слів можуть бути представлені на графі синтаксичної структури речення аналогічно до семантико-синтаксичних структур, тобто як сукупність семантичних примітивів, що поєднані між собою мережею семантико-синтаксичних відношень. Лексична та синтаксична семантика, а також — лексична багатозначність та синтаксична омонімія в цьому випадку будуть репрезентовані тотожними засобами. Мовні значення, виражені синтаксично або лексично отримують однорідну інтерпретацію в ситуаційному тезаурусі.

До визначення необхідної та достатньої сукупності семантичних примітивів для представлення лексичних значень існують різні підходи. З одного боку, було

здійснено спробу розробити семантичну метамову з мінімальною достатньою кількістю семантичних примітивів, що б дозволило їх комбінуванням репрезентувати усі можливі лексичні значення 10. Проте, в інтерпретаційній моделі лексичної і синтаксичної семантики більш продуктивним є інший підхід, згідно з яким сукупність "елементарних лексем" ϵ відкритою 11 і пов'язаною зі способом формалізації предметної галузі (зі структурою ситуаційного тезауруса). Взаємодія тексту з об'єктами, відношеннями і процесами в СТ здійснюється через деякі ключові лексеми, які можуть бути тотожними до назв об'єктів, їхніх методів, відношень, до значень атрибутів, тощо. Та сама ключова лексема при взаємодії з різними об'єктами СТ, іноді вживаючись у нетипових контекстах, може спричиняти різний ефект, наприклад, приписувати різним об'єктам можливі для них атрибути різного типу, що пов'язані зі специфікою цих об'єктів. Отже, функціонування семантичної метамови в цьому аспекті тотожне до функціонування реальних лексем, оскільки для ключових лексем також можливий контекстуальний та метафоричний вжиток. Це дозволяє приписувати лексико-семантичним варіантам слова у словнику фіксовану структуру семантико-синтаксичних відношень між ключовими лексемами метамови, динамічно утворюючи контекстуальні значення в процесі інтерпретації тексту в CT^{12} .

Тлумачення тих слів, що не пов'язані безпосередньо з назвами об'єктів, атрибутів, методів, відношень і процесів у ситуаційному тезаурусі (лексем, що не ϵ ключовими, і тому не входять до семантичної метамови) можна ефективно здійснювати, використовуючи структури семантико-синтаксичних відношень між ключовими лексемами. У наведеному вище прикладі семантико-синтаксичної структури речення розглянемо зразки репрезентації лексичної семантики для фрази $A_{(2-4)}$ - [якийсь час тому/. Згідно з синтаксичною структурою цієї фрази лексема час оброблятиметься першою. У ситуаційному тезаурусі, очевидно, її доцільно зробити ключовою лексемою, пов'язаною з активізацією абстрактного складного об'єкта ЧАС, що містить інший об'єкт ЛІНІЙНА ШКАЛА, ділянки якої можуть бути позначені атрибутами МИНУЛЕ, СУЧАСНЕ, МАЙБУТНЄ. Життєвим циклом цього об'єкта є зміщення ділянки з атрибутом СУЧАСНЕ з відповідною зміною атрибутів пов'язаних подій. В об'єкта ЧАС не існує деструктора, тому вираз вбивати час не може означати "анігіляції просторово-часового зразок континууму", інтерпретуватись метафорично стосовно продуктивності використання "часових ресурсів" конкретними об'єктами. Точки на часовій шкалі можуть бути пов'язані з певними системами виміру і ставати атрибутами інших об'єктів та подій, а також утворювати відношення "А раніше / пізніше ніж Б". Процедурною семантикою лексеми uac у даному випадку ε "виклик" такого об'єкта, але в іншому випадку, якщо ця лексема сама не "створюватиме" контекст, а реагуватиме на контексти інших лексем, її інтерпретація може передбачати, напр., актуалізацію "часових атрибутів" об'єктів іншого типу.

Займенниковий прикметник якийсь не доцільно робити ключовою лексемою; його слід тлумачити як єднальне семантико-синтаксичне відношення із запереченням, що поєднує дві ключові лексеми: $[\epsilon_{\partial нальн.B}$ конкретний і $\{(\epsilon_{3anepeu.B}, e)$ визначений) $\epsilon_{ampuoym.K}\}$ атриб К. В об'єкті ЧАС ці атрибути специфікують певну ділянку ("період") на шкалі, яка поки що не пов'язується з жодними подіями.

Слово *тому* омонімічне, проте оскільки під час аналізу фрази активним ϵ об'єкт YAC, "спрацьовує" лише один із омонімів, який має "хронологічне тлумачення": (- BVминулому) темпоратив.К. Ця лексико-семантична структура пов'язує створену на попередньому етапі ділянку на часовій шкалі із "константною" ділянкою шкали об'єкта ЧАС – МИНУЛЕ.

Аналогічно відбувається подальший процес інтерпретації лексико-синтаксичної структури речення - структури, де представлені семантико-синтаксичні відношення між лексемами та семантико-синтаксичні репрезентації тлумачень неелементарних лексем. Структурний запис тлумачень має також і лінійне прочитання, так само, як і синтаксична структура будь-якої іншої фрази у реченні та у тексті. Якщо під час створення формул тлумачення не допускати виникнення синтаксичної омонімії і використовувати ключові лексеми, які мають лише одну процедурну інтерпретацію у кожному конкретному контексті вживання, то між лінійною і структурною формою запису тлумачень існуватиме однозначна відповідність, що дозволяє використовувати звичайний словниковий формат для опису лексико-семантичних варіантів слова.

Запропонований метод може бути використано при комп'ютерному моделюванні процесу розпізнання змісту тексту на рівнях синтаксичної та лексичної семантики. Загальну семантику речення можна описати як процедурний ефект тексту, потенціал зміни ситуаційного тезауруса, спричинений інтерпретацією тексту в конкретний час у середовищі знань про стан справ у конкретній ситуації і зумовлений взаємодією мовних конструкцій та ситуаційного тезауруса. Це дозволяє моделювати нескінченну кількість різних інтерпретацій, можливих для кожного речення та для тексту в цілому.

Вихованець І.Р. Граматика української мови. Синтаксис. – К., Либідь, 1993. – С. 6, 18; Вихованець І.Р., Городенська К.Г., Русанівський В.М. Семантико-синтаксична структура речення. – К., Наук. думка, 1983. - С. 7. Для репрезентації мережі синтаксичних зв'язків та семантикосинтаксичних відношень було запропоновано як структури ієрархічного типу (т. зв. синтаксичні дерева безпосередніх складників, дерева залежностей, системи синтаксичних груп – Гладкий А.В. Синтаксические структуры естественного языка в автоматизированных системах общения. - М., Наука, 1985. – 144 с.), так і графи, що можуть мати цикли та ділянки із метаоператорами, напр., диз'юнктивним оператором "АБО" для представлення синтаксичної омонімії різного типу. Такими є системи синтаксемних груп, які використано у даній роботі, див. також Бабич Б.В. Представлення та інтерпретація омонімічних семантико-синтаксичних структур // Українське мовознавство. – 1997. - Випуск 21. - C. 89-100.

Пор. думку А.М. Мухіна про те, що одиниці глибинної структури – синтаксеми – "виділяються в позиціях [...] елементарних синтаксичних одиниць поверхневої структури речення" – Мухин А.М. Синтаксемный анализ и проблема уровней языка. – Л., Наука, ЛО, 1980. – С. 3-4. Проте, при такому аналізі певні труднощі становлять т. зв. складні синтаксеми типу він прийшов втомлений (втомленим), я побачив його сумного (сумним), які можна вважати трансформаціями двох простих

синтаксем, що відбулась або при перетворенні глибинної структури в поверхневу, або при перетворенні семантичної структури в глибинну синтаксичну. В україністиці прийнята друга точка зору: вважається, що між формально-синтаксичним і семантико-синтаксичним планами речення "нема тотожності, але ε паралелізм" – Buxoванець I.P. Зазнач. праця. – C. 43, прим.

- 3 Золотова Г.А. Синтаксический словарь: Репертуар элементарных единиц русского синтаксиса. М., Наука, 1988. С. 4.
- 4 Золотова $\Gamma.A$. Коммуникативные аспекты русского синтаксиса. М., Наука, 1982. С. 82-83.
- ⁵ Вихованець І.Р. Зазнач. праця. С. 43.
- 6 Назва для моделей цього типу пов'язана з роботою Л.Вітгенштайна "Логіко-філософський трактат" (Вітгенштайн Л. Тгастати logico-philosophicus; Філософські дослідження. К., Основи, 1995. С.8-86), яка започаткувала філософський напрямок логічного атомізму, одним з аспектів якого є теорія семантики мовних виразів. Атомістичної точки зору на лексичну семантику дотримуються автори лінгвістичної моделі "Смисл Текст" (Мельчук И.А. Опыт теории лингвистических моделей "Смысл Текст". Семантика, Синтаксис. М., Наука, 1974. 315 с.; Апресян Ю.Д. Типы информации для поверхностно-семантического компонента модели "Смысл ⇔ Текст" // Апресян Ю.Д. Избранные труды. Т. 2. Интегральное описание языка и системная лексикография. М., Школа "Языки русской культуры", 1995. С. 8-101), Концепція Л.Вітгенштайна покладена в основу теорії семантичної структури речення в роботі Вихованець І.Р., Городенська К.Г., Русанівський В.М. Зазнач. праця. С. 30. Аналогічний підхід використовується у Pustejovsky J. Linguistic constraints on type coercion. In: Computational lexical semantics. Cambridge, Cambridge University Press, 1995. Р. 71-97.
- ⁷ Процедурно-інтерпретаційні моделі лексичної семантики хронологічно пізніші за атомістичні моделі. Їх поява пов'язана з критикою логічного атомізму, зокрема в роботі Л.Вітгенштайна "Філософські дослідження" (Вітгенштайн Л. Зазнач. праця. С. 87-310), з лінгвістичною філософією (Austin J.L. How to do things with words. Cambridge, MA, Cambridge University Press, 1994. 169 р.). Основні положення процедурної семантики викладено в роботах Джонсон-Лерд Ф. Процедурная семантика и психология значения // Новое в зарубежной лингвистике. Вып. 23. М., 1988. С. 234-257.; Виноград Т., Флоренс Ф. О понимании компьютеров и познания // Язык и интеллект. М., Прогресс, 1995. С. 185-229.
- ⁸ *Бабич Б.* Інтерпретаційний семантичний аналіз тексту // Мова у соціальному і культурному контексті. Збірник наукових праць. К., Київський університет ім. Тараса Шевченка, 1997. С. 173-183.
- ⁹ Подібний формалізм використовується у комп'ютерному моделюванні виробничих, деяких природних та соціальних процесів і ситуацій: *Шлеер С., Меллор С.* Объектно-ориентированный анализ: моделирование мира в состояниях. К., Диалектика, 1993. 240 с.
- ¹⁰ Wierzbicka A. Semantics: primes and universals. Oxford, New York, Oxford University Press, 1996. 500 р.; Апресян Ю.Д. О языке толкований и семантических примитивах // Апресян Ю.Д. Избранные труды. Т. 2. Интегральное описание языка и системная лексикография. М., Школа "Языки русской культуры", 1995. С. 466-484.; Шенк Р. Обработка концептуальной информации. М., Энергия, 1980. 361 с.
- 11 Такий підхід використовує Рей Джекендофф у його теорії концептуальної семантики (Jackendoff R.. Conceptual semantics. In: Meaning and mental representations. Bloomington, Indianapolis, Indiana University Press, 1988. P. 81-97.; Saeed J.I. Semantics. Cambridge, MA, Blackwell Publishers, 1997. P. 249-259). Так, наприклад, у випадку, коли "фразу сконцентровано на семантиці руху, сутність "дім" ("the house") подається як неаналізований атом значення" Saeed J.I. Зазнач. праця. С. 250. Це дозволяє не дублювати функції об'єктів ситуаційного тезауруса у лексичних репрезентаціях, подаючи лише компактні посилання на ці об'єкти та описуючи послідовність взаємодії лексем речення із різними структурами СТ.
- ¹² Подібну вимогу не помножувати мовні значення було висунуто в роботі *Grice H.P.* Logic and conversation. In: Pragmatics: a reader. Ed. by Davis S. New York, Oxford, Oxford University Press, 1991. P. 305-315. Можна стверджувати, що значення комунікативної імплікатури (conversational implicature) мовного виразу утворюються подібно до контекстуальних значень окремих слів в процесі інтерпретації тексту в ситуаційному тезаурусі.

Бабич Богдан Віталійович аспірант Українського мовно-інформаційного фонду НАН України домашня адреса: просп. Героїв Сталінграда, 49, кв. 121,

Київ, 254 213

телефон: (044) 413-6449; тел. УМІФ (044) 267-4859

e-mail: B_Babych@ffd.univ.kiev.ua