

ДО ПИТАННЯ ПРО ФОРМАЛЬНІ ГРАМАТИКИ

У статті проаналізовано різні методи репрезентації формально-синтаксичних структур, обґрунтовано необхідність удосконалення систем формальних граматики для моделювання трансформації синтаксичних структур.

Формальні методи є тією науковою мовою, яка завдяки своїй структурі дозволяє відкрити у спостережуваних фактах нові аспекти та зв'язки. Прогрес лінгвістики пов'язаний насамперед із підвищенням точності наукової мови, на якій формулюються опис матеріалу і наукові теорії. На думку Рея Джекендоффа, мова завжди інтегрована у когнітивні структури і “допомагає людині мислити” [19, 14:851]. З цього погляду точна наукова мова у лінгвістичному дослідженні відкриває новий простір для мислення, визначає нові способи пізнання мовознавчих об'єктів у створеній нею пізнавальній галузі. Створення формальних лінгвістичних моделей є актуальним завданням як при розробці комп'ютерних інформаційних систем, так і при теоретичному осмисленні процесів мовлення і розуміння, когнітивних аспектів мови, її функціонування в структурі комунікації та людського пізнання.

Лінгвістичній науці відомі різні методи представлення формально-синтаксичних структур (структури речень, словосполучень та елементарних синтаксичних одиниць, заснованих на синтагматичних прогнозуваннях граматичних категорій і значень) – метод представлення синтаксичної структури як системи безпосередніх складників, як системи рангових складників [16], як дерева залежностей [8] та як системи синтаксичних груп [5].

Основним одиницями в складі речення виступають *словоформи*, *групи словоформ* і *синтаксичні відношення* між ними. Словоформи в реченні розташовані в лінійній послідовності. Водночас вони організовані також на структурному рівні, утворюючи між собою синтаксичні відношення та функціональні групи, які визначають характер можливих змін порядку словоформ у реченні під час трансформацій різного типу (пасивізації, питальної трансформації тощо): за термінологією генеративної лінгвістики ця особливість речень має назву структурна залежність (*structure-dependency*) між словоформами. “Структурна залежність означає, що знання мови засноване на структурних відношеннях у реченні, а не на послідовності слів. ... Структурна залежність – принцип, загальний для усіх мов. Визначення: дії з реченням, такі як рух окремих елементів, вимагають знання про структурні відношення між словами, а не про їхню лінійну послідовність” [15: 3-4, 13].

Синтаксичний компонент (СК) – це словоформа, словосполучення або відношення синтаксичної трансформації між словоформами і словосполученнями, які є елементами (складниками) іншого словосполучення або речення. Синтаксична структура речення – це система його синтаксичних компонентів.

Між синтаксичними компонентами, що знаходяться на одному рівні членування, існують синтаксичні відношення. Загальна семантика речення впливає як із семантики синтаксичних компонентів, так і з семантики синтаксичних відношень між ними в структурі речення.

Спосіб об'єднання словоформ у групи репрезентує синтаксичну організацію речення на певному семіотичному рівні. На кожному рівні організації речення синтаксична структура може моделюватися подібними формальними репрезентаціями, утворюючи системи синтаксичних компонентів. Наведені визначення синтаксичного компонента та структури речення близькі до визначень складника та системи складників у генеративно-трансформаційній моделі синтаксичних структур, але відмінність полягає в тому, що синтаксичні компоненти репрезентують також відношення синтаксичної трансформації, а не лише статичні синтаксичні одиниці. У статті представлена структура репрезентації узагальнює формалізм систем безпосередніх складників (БС). Розширені системи БС містять інформацію про трансформаційні зв'язки певного речення з іншими елементами його парадигми, а отже, одночасно є і “дериваційним”, і “таксономічним”, “репрезентаційним” формалізмом [21:32-46]. Цей формалізм репрезентує “творчий аспект мови, правила, за якими відбувається така творча діяльність у мові” [6:25]. Також цей аспект “у сучасній формі втілений у трансформаційній моделі граматики, створеній Н.Хомським [10,11,12], а в традиційній формі – в теорії універсальної граматики сімнадцятого сторіччя, створеній, головним чином, філософами і граматистами Пор-Рояля. Фундаментальним положенням, на якому базувалася *Grammaire générale et raisonnée*, було те, що значення, або логічна форма речення, не має прямого або систематичного відношення до синтаксичної структури, виявленої у формах висловлювань речення, але має пряме та систематичне відношення до глибинної (underlying) синтаксичної структури, від якої експліцитна структура речення сама є функцією” [21: 46-47].

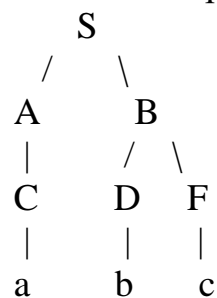
Метою формалізму є експліцитне представлення структурної організації речення та інваріанту його синтаксичної деривації, який дозволяє визначити множину трансформів для заданого елементарного речення (див. [4: 31-32]).

ГраMATика безпосередніх складників характеризує синтаксичну структуру речення як ієрархічну систему множин словоформ та їхніх груп. Система безпосередніх складників, утворена граматиною складників, містить інформацію про морфологічні категорії словоформ у вузлах синтаксичної структури [5: 42-45].

Суть граматики безпосередніх складників (БС) полягає в наборі символів і правил перекодування ланцюжків символів. Правила

перекодування мають форму $X \rightarrow Y$, де X і Y позначають ланцюги символів. Одні символи використовуються для позначення слів і морфем, інші – для позначення синтаксичних груп і речень. Перші утворюють так званий термінальний словник, другі – нетермінальний словник. Крім того, використовується особливий знак #, який позначає початок і кінець речення. На вході цієї структури знаходиться ланцюг # S #, який позначає речення як глобальний неподільний елемент. Застосувавши перше правило перекодування $\# S \# \rightarrow \# Y_1 \#$, одержуємо ланцюг $\# Y_1 \#$. Послідовність застосування правил $\# Y_1 \#, \# Y_2 \#, \# Y_3 \# \rightarrow \# Y_3 \#$ і т. ін. дає можливість ланцюжків $\# Y_2 \#, \# Y_3 \#$ і т. ін., доки не дійдемо до термінального ланцюга $\# Y_n \#$. Цей ланцюг називається термінальним, оскільки символи, які входять до його складу, не можуть замінюватися новими символами. Одержана в такий спосіб сукупність ланцюгів називається *деривацією ланцюга $\# Y_n \#$* .

Такий процес виникнення граматично правильних речень прийнято зображувати як генеалогічні дерева. Припустимо, дані правила: 1) $S \rightarrow A+B$, 2) $A \rightarrow C$, 3) $B \rightarrow D+F$, 4) $C \rightarrow a$, 5) $D \rightarrow b$, 6) $F \rightarrow c$, на основі яких одержимо таку деривацію: # S #, # $A+B$ #, # $C+B$ #, # $C+D+F$ #, # $a+D+F$ #, # $a+b+c$ #. Цю деривацію можна зобразити як генеалогічне дерево :



Якщо пояснювати символ A через термін “іменна група”, символ B – через термін “дієслівна група”, символ C – через термін “іменник”, символ D – через термін “дієслово”, символ F – через термін “додаток”, символ a – через слово “учень”, символ b – через слово “читає”, символ c – через слово “книжку”, то можна прийняти вищенаведену деривацію за структурний аналіз речення “Учень читає книжку”.

Застосування правил породження (виникнення) пов’язане з такими умовами: а) кожне окреме правило допускає перекодування тільки одного символу; не можна перекодувати одночасно два чи кілька символів; б) у процесі перекодування не допускається перестановка символів; в) при формулюванні правила породження ланцюгів приймається до уваги тільки стан кожного певного ланцюга; що ж стосується їх попередніх ланцюгів, то їхній стан не береться до уваги. Інакше кажучи, історія деривації кожного певного ланцюга не впливає на породження (виникнення) наступних ланцюгів.

Така в загальних рисах формальна характеристика функціонування того типу граматики, яку прийнято називати моделлю виникнення (породження) БС [10,11,12].

Виведення речення – це впорядкована послідовність правил формальної граматики, які моделюють відношення між реченням та однією з його

синтаксичних структур. Виведення описує послідовність, “протокол” відображення синтаксичної структури в речення або речення в синтаксичну структуру. У виведенні вказано конкретний шлях утворення речення або синтаксичної структури, “абсолютний” порядок застосування правил формальної граматики. Варіювання послідовності правил у виведенні, яке не порушує пріоритетів для тієї чи іншої інтерпретації речення, не є релевантним для синтаксичної структури: всі виведення речення, де збережено пріоритети правил формальної граматики, утворюють ту ж саму синтаксичну структуру. Виведення речення звичайно має єдиний принцип – “стратегію”, або *порядок виведення*, що звичайно визначається способом опрацювання синтаксичної структури речення.

Формальні граматики складників моделюють відношення між реченням і його синтаксичною структурою через виведення речення – схема (1):



Синтаксична структура речення і послідовність його словоформ утворюються одночасно при породженні речення за правилами формальної граматики. Підбір правил формальної граматики, у які повинні увійти усі словоформи певного речення та їх групи, здійснюється виведенням речення при синтаксичному аналізі; паралельно виведення речення формує синтаксичну структуру речення (як послідовність дібраних правил). Отже, як при породженні, так і при синтаксичному аналізі виведення речення можна назвати структурою виведення (англ. *derivation structure*). Речення, тобто лінійну послідовність словоформ, для якої виведення здійснює синтаксичний аналіз і породження, можна назвати результатом застосування граматики складників.

Суть **генеративної-трансформаційної граматики** полягає передусім в тому, що вона є абстрактною дедуктивною системою, на основі якої за допомогою кінцевої кількості правил можна проаналізувати безкінечну кількість правильно побудованих висловлювань певної мови. Генеративна граMATика синтетична і дедуктивна: вона спрямована від абстрактних моделей і правил до конкретних висловлювань відповідної мови. Її аналіз відбувається від одиниць вищого ярусу мовної системи - речення - до нижчих - словосполучення та слова. Власне, “генеративна граMATика є синтезом понять традиційної граматики з формалізованими методами

структурної граматики”, - вважає Л.Бархударов [2]. Генеративна граMATика не дає одиницям і поняттям ніяких визначень: правило, за яким виникає певна одиниця, і є його визначенням. Так, визначення речення подається його першим (початковим) правилом: $S \rightarrow NP + VP$, тобто речення – це іменна група плюс дієслівна група. Проте це правило само по собі ще не забезпечує формування безкінечної кількості речень із обмеженого набору вихідних елементів, оскільки воно не передбачає можливості “самовключення” одних речень у інші. Тому необхідно розширити назване правило таким чином, щоб надати йому рекурсивний характер, а саме: $S \rightarrow NP + S + VP$. З останнього правила випливає, що $S \rightarrow NP + (NP + S + VP) + VP$, і далі, що $S \rightarrow NP + [NP + (NP + S + VP) + VP] + VP$ і т. ін. Таким шляхом формуються речення будь-якої складності.

Уперше трансформаційні правила з’явилися у працях З.Гарріса [17], але заслуга чіткого обґрунтування необхідності цих правил у граматиці та побудови системи трансформаційної граматики належить Н.Хомському [11].

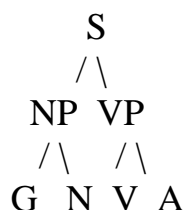
Трансформаційні правила репрезентують природне продовження правил виникнення речень за БС. За допомогою трансформаційних правил речення, які виникли звичайним способом за моделлю безпосередніх складників, повинні трансформуватися в нові речення. Для трансформаційних правил суттєво зняті обмеження, які існують при застосуванні моделі БС. Якщо за правилами моделі безпосередніх складників забороняється перекодування одночасно більше одного символу, то в трансформаційній моделі дозволяється одночасне перекодування декількох символів; якщо модель БС забороняє перестановку символів, то трансформаційна модель дозволяє переставляти символи; якщо модель БС забороняє звертатися до історії дериваційних ланцюгів, то трансформаційна модель, навпаки, вимагає звертатися до історії дериваційних ланцюгів.

Варто зазначити, що для трансформаційних правил є особливо суттєвою остання обставина – необхідність звертатися до історії деривації ланцюгів. Трансформаційні правила існують не тільки для виникнення одного речення з іншого, а саме для того, щоб шляхом перестановки символів і різного характеру змін у складі символів утворювати одні типи дериваційних дерев від інших типів дериваційних дерев.

У будь-якому трансформаційному правилі слід розрізняти три суттєвих аспекти, які нерозривно пов’язані між собою: а) вихідне дериваційне дерево; б) ланцюг, до якого безпосередньо застосовується трансформація; в) набір елементарних трансформаційних операцій.

Для ілюстрації дії трансформаційних правил звернемося до прикладу виникнення речення з однорідними членами *Листи батька і брата лежать тут* із речень *Листи батька лежать тут* і *Листи брата лежать тут*. Можна зробити висновок, що виникнення речень з однорідними членами виходить за рамки моделі БС, оскільки виникає необхідність звертання до історії деривації ланцюгів, які містяться в основі останніх двох речень. Тому правило виникнення речень з однорідними членами повинне трактуватися не як правило моделі безпосередніх складників, а як трансформаційне правило.

В аналізованих реченнях маємо вихідне дериваційне дерево, яке виникло на основі моделі БС:



Існують також два ланцюги, до яких застосовується трансформаційне правило:

$$\begin{array}{l}
 N + \text{батька} + V + Ad \\
 N + \text{брата} + V + Ad
 \end{array}$$

Крім цього, є набір елементарних трансформаційних операцій: а) операція зчіплювання слів *батька* і *брата* за допомогою сполучника *і*: *батька + і + брата*, б) операція підставлення групи *батька + і + брата* замість слова *батька* в першому реченні. У результаті цього утворився новий ланцюг *N + батька + і + брата + V + Ad*, який міститься в основі нового дериваційного дерева, що розгортається за правилами моделі БС.

Із цього випливає, що трансформаційні правила не скасовують модель безпосередніх складників, а зараховують цю модель у рамки більш ширшої граматичної системи, яку Н.Хомський називає трансформаційною граматику [12].

Трансформаційна граMATика репрезентує систему, яка складається з правил трьох типів: 1) правил моделі безпосередніх складників; 2) правил трансформаційної моделі; 3) морфофонемних правил [9:400].

За трансформаційною граматику виникнення речень відбувається так. Являючи собою кібернетичне впорядкування певного типу, трансформаційна граMATика має на виході символ *S*, який позначає речення як нерозчленований, глобальний елемент; шляхом низки послідовних дій цей символ перебудовується за правилами моделі БС в термінальний ланцюг цієї моделі. Набір подібного типу термінальних ланцюгів утворює ядро мови. До ядра мови належать прості розповідні активні речення – так звані ядерні речення. За допомогою використання трансформаційних правил до ядерних речень долучаються нові ланцюги, які водночас розгортаються за правилами моделі безпосередніх складників. Ланцюги, одержані в результаті правил моделі безпосередніх складників і трансформаційної моделі, перекодовуються на виході в ланцюги морфофонем згідно з спеціальними морфофонемними правилами [див.:12:238].

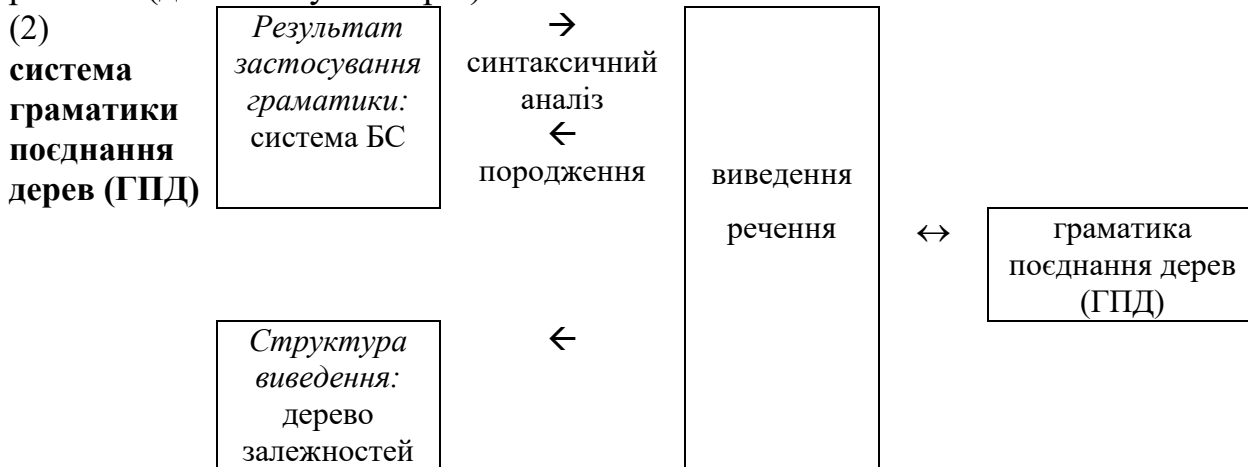
Клас формальних граMATик, що дістав назву **граматики поєднання дерев (ГПД)**, розглядається в роботі американських дослідників Аравінда Джоні та Івеса Шабеса [20]. Репрезентація синтаксичної структури речення, пов'язана з ГПД, здійснюється одночасно засобами дерева залежностей та ієрархізованих систем безпосередніх складників. Очевидно, ГПД задає співвідношення між ДЗ (дерево залежності) та узгодженою з нею системою БС подібно до того, як контекстно-вільна граMATика складників задає

співвідношення між системою безпосередніх складників і лінійною послідовністю словоформ речення. Специфікою граматики поєднання дерев є принцип лексикалізації формальної граматики.

Для контекстно-вільних граматик (як різновидом граматик безпосередніх складників) результатом застосування граматики (тобто об'єктом, який породжується при генеративній інтерпретації формальної граматики і розпізнається при рецептивній інтерпретації) є одновимірний лінійний об'єкт – послідовність словоформ речення. Нелінійна синтаксична структура речення у форматі систем БС утворюється паралельно із породженням або розпізнанням цього лінійного об'єкта у виведенні речення. І при породженні, і при розпізнанні послідовності словоформ будується така нелінійна структура. На позначення цієї структури у формальних граматиках різних видів існує термін - структура виведення (що протиставляється результатів застосування граматики). Будь-яку формальну граматику можна розглядати як механізм відображення результату застосування граматики (у випадку граматик складників, граматик залежностей - відображення лінійних послідовностей словоформ речень) у структуру виведення (в цьому випадку – у синтаксичну структуру речень у форматі відповідно систем БС, ДЗ).

Відмінність підходу, запропонованого в ГПД, полягає в тому, що результатом застосування граматики поєднання дерев є дерево фразових маркерів (тобто система БС певного речення) – нелінійний структурований об'єкт. Отже, при генеративній інтерпретації ГПД утворюється синтаксична структура речення у форматі системи БС, а не саме речення як лінійна послідовність словоформ. Аналогічно, при рецептивній інтерпретації ГПД на вході є система БС – структурований, а не одновимірний лінійний об'єкт [1].

У процесі породження або розпізнання системи БС певного речення виведення ГПД будує дерево залежностей, узгоджене із такою системою БС. ДЗ є структурою виведення ГПД. Отже, ГПД задає відображення результату застосування граматики – системи БС – у структуру виведення – ДЗ певного речення (див. схему 2 і пор.1):



ГПД моделює відношення між двома нелінійними структурованими об'єктами – системою БС і ДЗ речення. Відношення кожного із цих об'єктів до лінійної послідовності словоформ відповідного речення не розглядається в системі ГПД.

Значення ГПД для лінгвістики полягає насамперед у можливості з'ясувати сутність і формально обґрунтувати поняття синтаксичного зв'язку, яке є основним поняттям при побудові ДЗ. Необхідність такого обґрунтування зумовлена тим, що концепт синтаксичного зв'язку є теоретичним конструктом (синтаксичний зв'язок, на відміну, напр., від словоформи чи групи словоформ, які є основою поняття складника, не доступний для безпосереднього спостереження). Очевидно також, що синтаксичний зв'язок не є елементарним лінгвістичним об'єктом (напр., таким, як значення граматичної категорії): в його структурі мають бути принаймні два місця для поєднаних СК, а також можливість розмітки СК і розмітки самого зв'язку. Водночас у сучасній синтаксичній теорії поняття синтаксичного зв'язку використовується як вихідний концепт і не виводиться із елементарніших лінгвістичних понять.

У ГПД поняття синтаксичного зв'язку засноване на понятті складника – множини словоформ або інших складників, яка повинна мати референт у пов'язаній із реченням ситуації. Структура ГПД впливає зі структури контекстно-вільної граматики складників. Це дозволяє ефективно використовувати ГПД для моделювання співвідношення між різними видами репрезентації синтаксичної структури речень.

ГраMATика поєднання дерев (ГПД) – це система вихідних і допоміжних дерев фразових маркерів та правил комбінування цих дерев, внаслідок якого утворюються дерева фразових маркерів граматично правильних речень.

У вихідних деревах зовнішні вузли або містять словоформи речення (є лексикалізованими), або позначені спеціальним символом (“↓”) для заміни цього фразового маркера іншим вихідним деревом, вершина якого може бути уніфікована з цим маркером. Заміна фразових маркерів на початкові дерева може здійснюватися лише у зовнішніх вузлах.

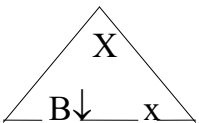
У допоміжних деревах зовнішні вузли або лексикалізовані, або позначені символом заміни (“↓”) для заміни – усі крім одного вузла (“основи” допоміжного дерева), який завжди містить маркер, що тотожний вершині свого дерева. Фразовий маркер в основі допоміжного дерева позначається символом (“*”) – символом поєднання цього дерева з іншим вихідним або допоміжним деревом. Поєднання допоміжного дерева з іншим деревом фразових маркерів може здійснюватися в будь-якому вузлі останнього, такому, який уніфікується із вершиною та основою цього допоміжного дерева. При поєднанні допоміжне дерево “включається” в інше дерево, замінюючи його будь-який внутрішній вузол, який тотожний вершині та основі цього допоміжного дерева.

Операція заміни зовнішніх вузлів вихідними деревами використовується для завершення лексикалізації гілок у виведених деревах. Операція поєднання дерев дозволяє включати синтаксичну структуру допоміжного дерева до складу іншого дерева, замість будь-якого внутрішнього вузла.

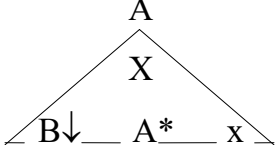
Якщо ГПД вимагає, щоб у зовнішніх фразових маркерах вихідних або допоміжних дерев знаходилася принаймні одна словоформа (“анкер” дерева), ці дерева і ГПД називаються лексикалізованими¹ [20: 3].

Лексикалізовані дерева містять лексичні анкери – словоформи речення. Процес комбінування вихідних та допоміжних дерев фразових маркерів (тобто виведення загального дерева фразових маркерів речення) формує структуру виведення – дерево синтаксичних залежностей між словоформами, які є лексичними анкерами дерев у ГПД.

Початкові дерева фразових маркерів у ГПД мають структуру (3):

- (3)
- | | | |
|---|---|---|
| A | – | A – вершина початкового дерева |
| X | – | X – внутрішні вузли |
|  | – | зовнішні вузли: B↓ – нелексикалізовані, позначені для заміни символом “↓”; x – лексикалізований вузол |

Допоміжні дерева фразових маркерів у ГПД мають структуру (4):

- (4)
- | | | |
|---|---|--|
| A | – | A – вершина допоміжного дерева |
| X | – | X – внутрішні вузли |
|  | – | зовнішні вузли: B↓ – нелексикалізовані, позначені для заміни символом “↓”; x – лексикалізований вузол; A* – нелексикалізований вузол, тотожний вершині, позначений для поєднання символом “*”
– основа допоміжного дерева |

Малі літери на схемах дерев фразових маркерів позначають лексикалізовані зовнішні вузли, в яких знаходяться словоформи – анкери відповідних дерев. Напр., в деревах (3) і (4) анкером є словоформа, позначена

¹ Повне формальне визначення ГПД наведене в зазначеній праці: “Грамматика поєднання дерев складається із п’ятірки об’єктів (Σ , NT, I, A, S), де

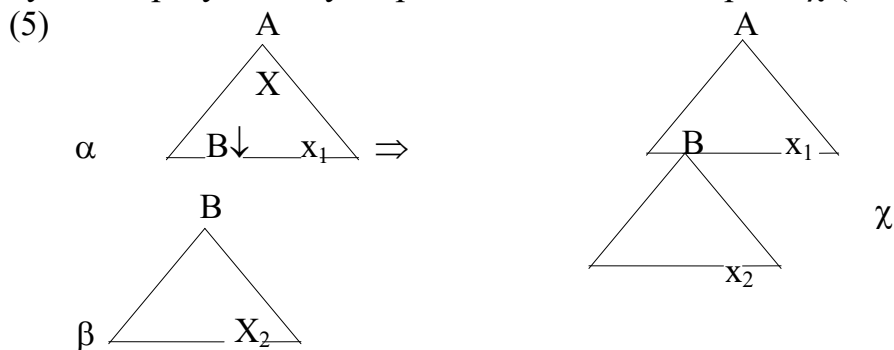
- (i) Σ – це скінченна множина термінальних символів;
- (ii) NT – це скінченна множина не-термінальних символів: $\Sigma \cap NT = \emptyset$
- (iii) S – це виділений не-термінальний символ: $S \in NT$;
- (iv) I – це скінченна множина скінчених дерев фразових маркерів, що називаються *початковими деревами* (англ. *initial trees*)
- внутрішні вузли дерева позначені не-термінальними символами
 - вузли на границі дерева (зовнішні вузли) позначені термінальними або не-термінальними символами; не-термінальні символи на границі початкових дерев у множині I марковані “для заміни” (що позначається знаком “↓”)
- (v) A – це скінченна множина дерев фразових маркерів, що називаються *допоміжними деревами* (англ. *auxiliary trees*), які характеризуються таким чином:
- внутрішні вузли позначені не-термінальними символами;
 - вузли на границі допоміжного дерева позначені термінальними або не-термінальними символами. Не-термінальні символи на границі дерев у множині A позначені “для заміни”, за винятком одного вузла, який називається *основою* (англ. *foot node*); він умовно позначається символом (“*”). Фразовий маркер основи повинен бути ідентичним фразовому маркеру у вершині (англ. *root node*) допоміжного дерева.

У лексикалізованих ГПД, принаймні один термінальний символ (анкер – англ. *anchor*) повинен позначати вузол на границі усіх початкових або допоміжних дерев.

Множина початкових і допоміжних дерев $I \cup A$ називається множиною *елементарних дерев* (англ. *elementary trees*) [...] Дерево, побудоване поєднанням двох інших дерев називається *виведеним деревом* (англ. *derived tree*).” [20: 2-3].

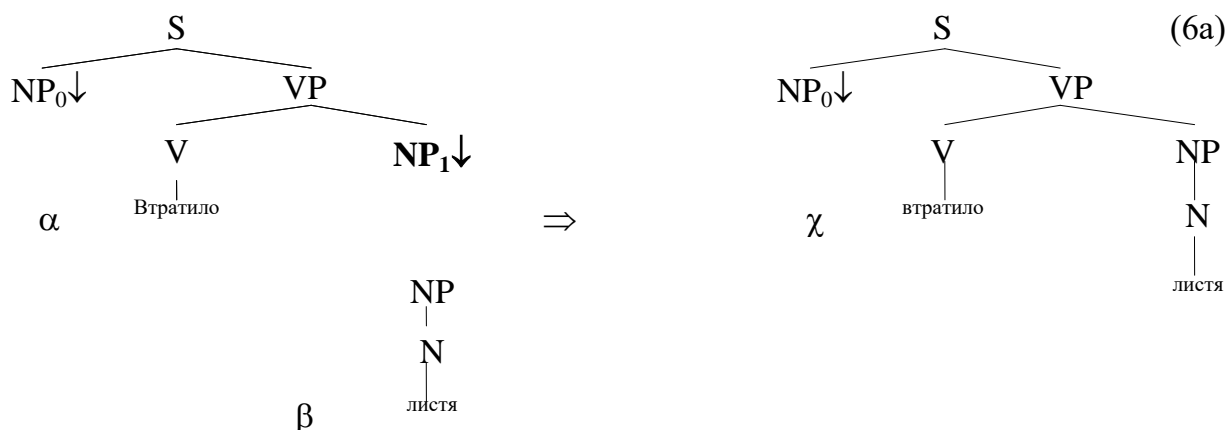
“ χ ”. Якщо початкове або допоміжне дерево фразових маркерів у ГПД має декілька анкерів, то вони оформлюються як складник у ДЗ (структурі виведення ГПД).

Початкові дерева фразових маркерів приєднуються у зовнішніх вузлах, позначених для заміни (символом “ \downarrow ”): вершина B початкового дерева β , яке приєднується до зовнішнього вузла $B\downarrow$ у дереві α , замінює цей зовнішній вузол. У результаті утворюється виведене дерево χ (схема 5):

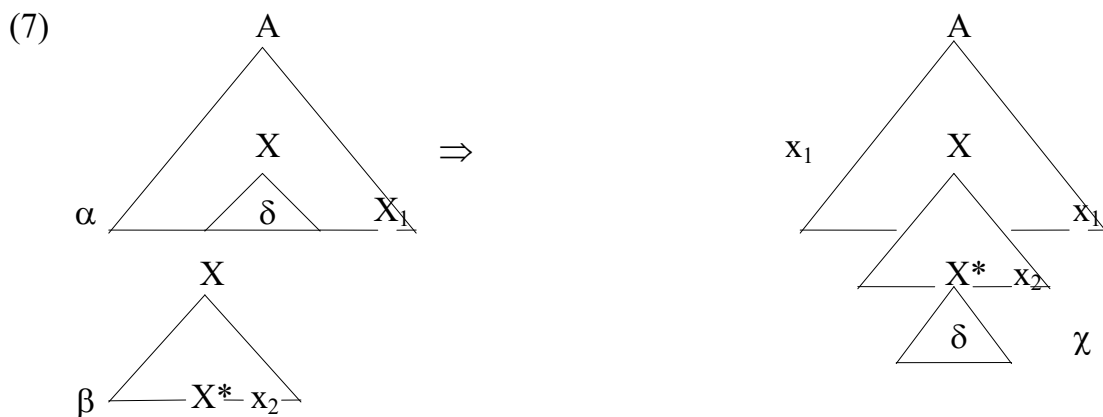


Напр., у системі БС речення (6) початкове дерево β , лексикалізоване іменником, замінює зовнішній вузол у дереві α , лексикалізованому дієсловом (на схемі дерева α виділено вузол, який замінюється початковим деревом β):

(6) *Дерево втратило листя*

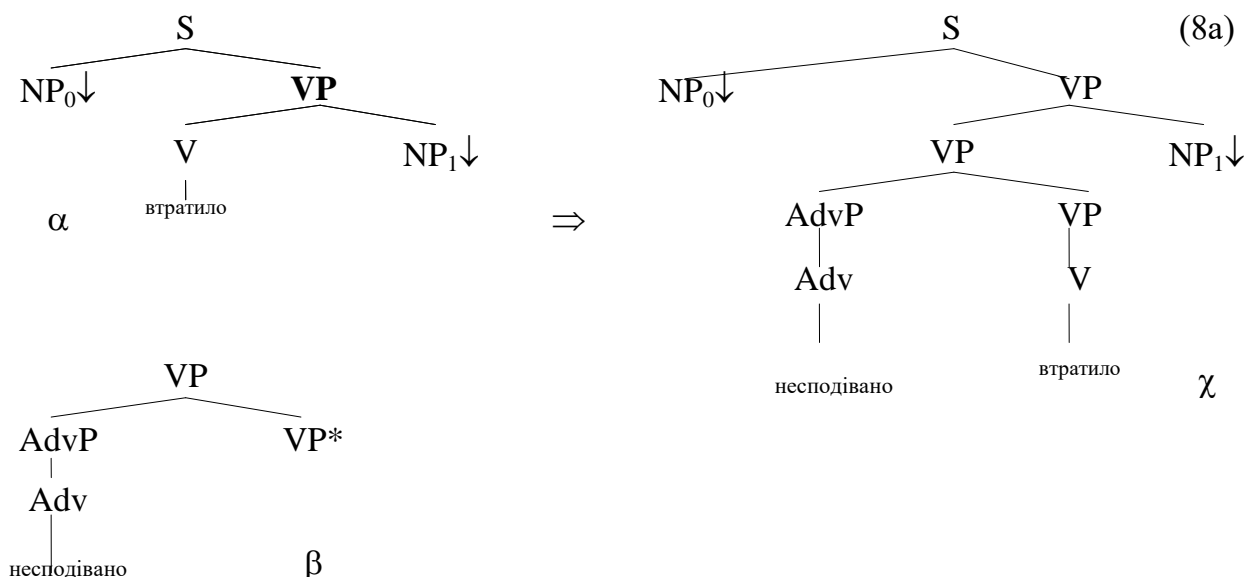


Допоміжні дерева приєднуються в будь-якому внутрішньому вузлі іншого початкового або допоміжного дерева. Цей вузол повинен бути тотожний вершині і основі допоміжного дерева (що позначена символом поєднання “*”). Виведене дерево в цьому випадку утворюється шляхом “розщеплення” синтаксичної структури допоміжним деревом. При поєднанні допоміжного дерева β до дерева α : (1) із синтаксичної структури дерева α вилучається піддерево δ , вершиною якого є вузол X , тотожний вершині допоміжного дерева β ; допоміжне дерево β замінює вилучену структуру δ ; (2) вилучене піддерево δ приєднується до утвореної синтаксичної структури χ , замінюючи основу допоміжного дерева – X^* (вузол, позначений для поєднання) – схема (7):



Напр., у системі БС речення (8) здійснюється поєднання допоміжного дерева β , лексикалізованого прислівником (на схемі виділено вузол дерева α , в якому відбувається поєднання допоміжного дерева β):

(8) *Дерево несподівано втратило листя.*



Виведення ГПД у процесі породження системи БС паралельно формує іншу репрезентацію синтаксичної структури речення – дерево залежностей (як структуру виведення ГПД). Побудова ДЗ речення у виведенні ГПД здійснюється так: процес породження системи БС розпочинається із деякого початкового дерева, яке містить складники найвищого рівня членування для певної синтаксичної структури – *вихідного дерева*. (1) Анкер вихідного дерева фразових маркерів – одна словоформа або група словоформ (як складник) – стає вершиною ДЗ. (2) Словоформа (або група словоформ), яка є анкером дерева фразових маркерів, *приєднаного* до вихідного дерева, у виведенні ГПД формує синтаксичний зв'язок залежності із вершиною утвореного ДЗ. Цей процес повторюється стільки разів, скільки (початкових та допоміжних) дерев приєднується до вихідного дерева фразових маркерів. У виведенні ГПД анкери усіх приєднаних дерев формують зв'язки залежності із вершиною ДЗ (анкером вихідного дерева фразових маркерів). (3) Деревя фразових маркерів, що були приєднані до вихідного дерева, у свою чергу, можуть змінюватися: до них (так само, як до

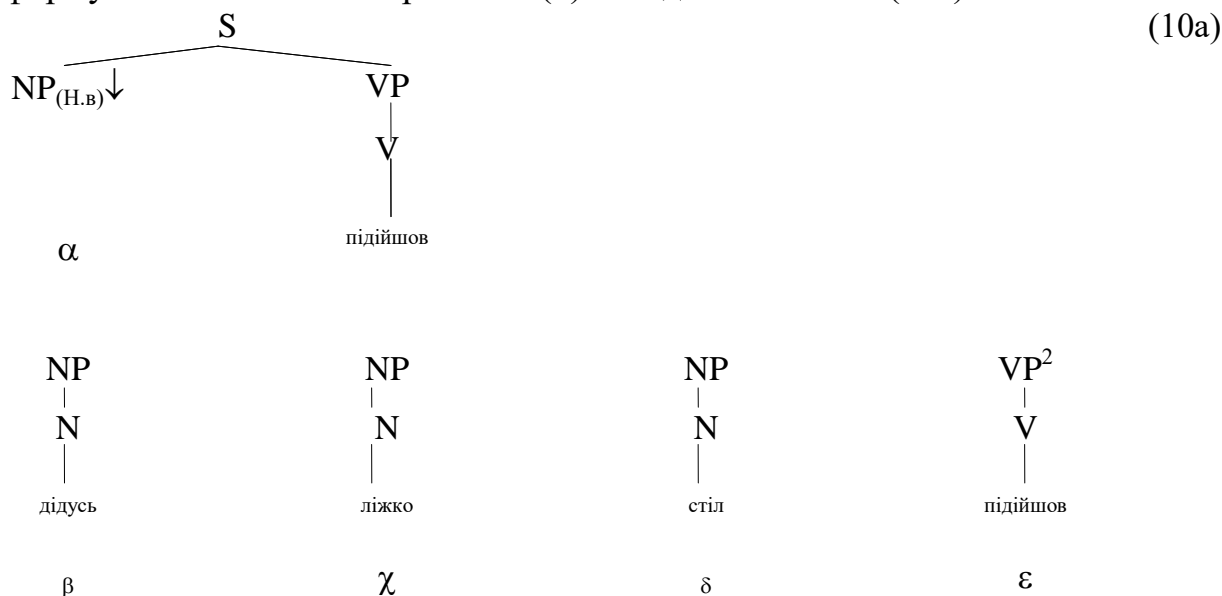
вихідного дерева) приєднуються інші початкові та допоміжні дерева. У виведенні ГПД лексичні анкери приєднаних дерев фразових маркерів формують зв'язки залежності із анкерами дерев, до яких вони приєднуються.

Результатом застосування ГПД стає система БС (виведене дерево фразових маркерів певного речення). Паралельно з цим у виведенні ГПД утворюється ДЗ певного речення, узгоджене із його системою БС. ДЗ формується в міру того як синтаксичні зв'язки залежності між лексичними анкерами поєднуваних дерев фразових маркерів утворюють зв'язний граф. Дуги дерева орієнтовані в напрямку синтаксичних залежностей між словоформами.

У статті розглядаємо лексикалізовані ГПД, у яких анкери початкових та допоміжних дерев фразових маркерів складаються з однієї словоформи. Такі ГПД у виведенні формують ДЗ без складників. Але в загальному випадку у вузлах ДЗ можуть знаходитися як словоформи, так і складники – останні утворюються в тому випадку, якщо анкером деякого початкового або допоміжного дерева є декілька словоформ. Утворення ДЗ у виведенні ГПД паралельно зі створенням системи БС може бути проілюстрована на прикладі речення (9):

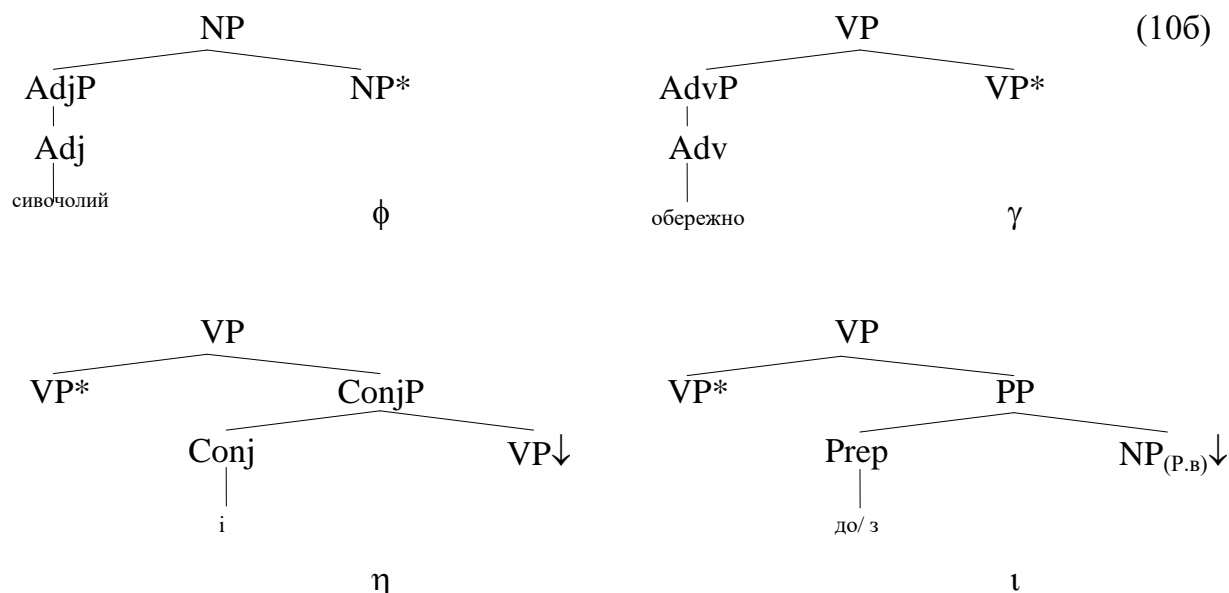
(9) *Сивочолий дідусь обережно спустився з ліжка і підійшов до столу.*

Набір початкових дерев фразових маркерів ГПД, необхідний для формування системи БС речення (9) наведено на схемі (10а):



Множину допоміжних дерев ГПД, необхідних для побудови системи БС речення (9) наведено на схемі (10б):

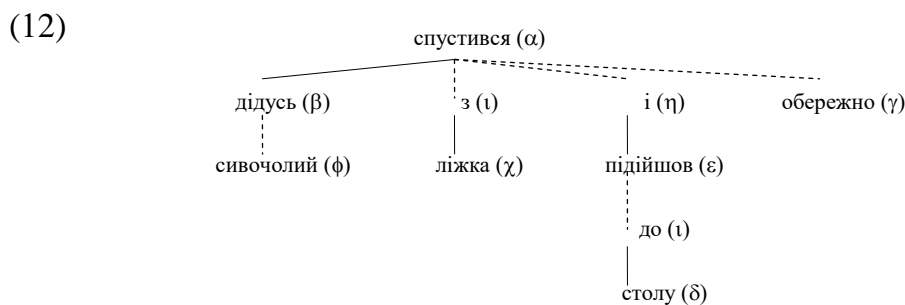
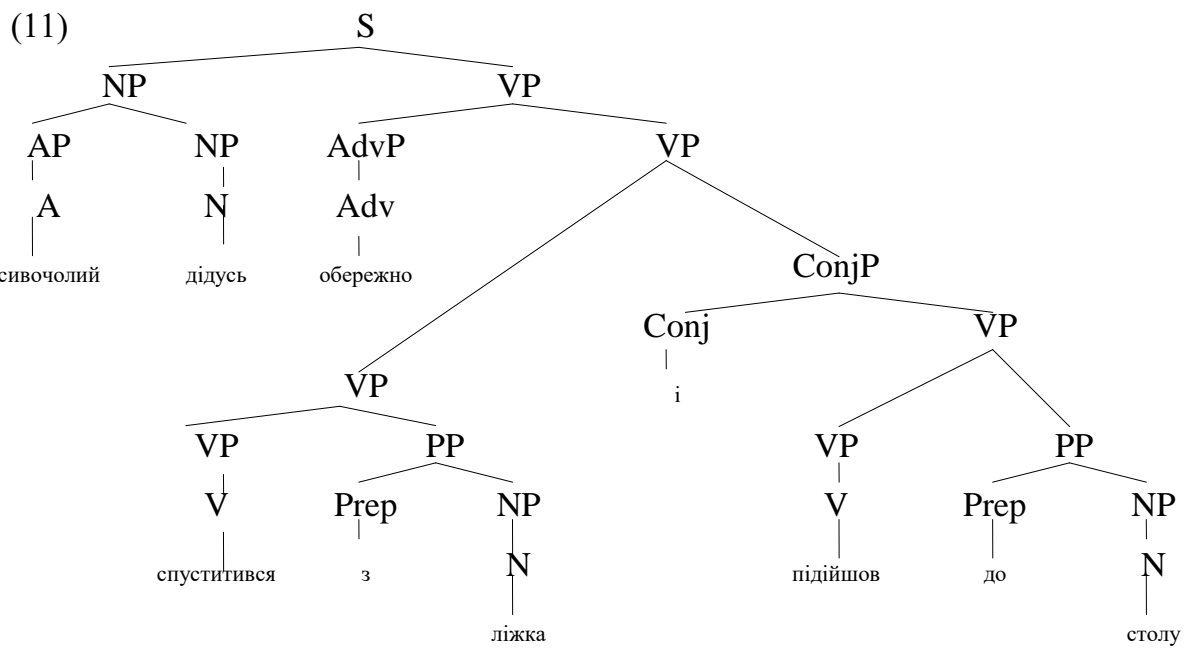
² Відмінність у репрезентації першого і другого дієслів зумовлена тим, що в цьому реченні обидва дієслова пов'язані із тим самим підметом а також – між собою сурядним зв'язком. Отже, граф синтаксичних зв'язків речення (9) реально має циклічну структуру. В той же час, у традиційній формі ГПД не передбачено засобів, які б оперували із циклічними синтаксичними структурами. Можливою альтернативою наведеної репрезентації є лексикалізація групи підмета і дієслівних груп сполучником. Початкове дерево такого виду у виведенні ГПД формує ДЗ, в якому підмет і обидва дієслова підпорядковані сурядному сполучникові. Ця структура синтаксичних зв'язків не має задовільної лінгвістичної інтерпретації.



Виведення ГПД у процесі приєднання дерев фразових маркерів паралельно формує ДЗ цього речення: початкове дерево α є вихідним для породження системи БС, тому його лексичний анкер *підійти* стає вершиною ДЗ. До вихідного дерева шляхом заміни зовнішнього фразового маркера приєднується дерево β , лексикалізоване анкером *дідусь*. У результаті утворюється виведене дерево, яке лексикалізоване підметом *і* присудком. Внаслідок такого поєднання дерев фразових маркерів виникає підрядний синтаксичний зв'язок між їхніми анкерами: іменник підпорядковується дієслову. До дерева β , лексикалізованого іменником *дідусь* у вузлі NP приєднується (шляхом поєднання дерев) допоміжне дерево ϕ , лексикалізоване прикметником *сивочолий*. У виведенні ГПД це створює синтаксичний зв'язок, спрямований від іменника до прикметника. Отже, процес поєднання дерев фразових маркерів зумовлює утворення зв'язного дерева залежностей між словоформами речення (9), які є анкерами у початкових і допоміжних деревах ГПД.

На схемі (11) представлено виведене дерево фразових маркерів (система БС) речення (9), утворене як результат застосування ГПД (9).

На схемі (12) представлено ДЗ, утворене як структура виведення ГПД (10). На схемі (12) суцільні лінії спрямовані до словоформ, приєднаних шляхом заміни вузлів; пунктирні лінії – до словоформ, приєднаних шляхом поєднання дерев фразових маркерів [20: 7].



ГПД утворюється шляхом лексикалізації контекстно-вільних граматики складників [20:18-21]. Це означає, що початкові та допоміжні дерева фразових маркерів утворюються в процесі застосування правил граматики складників до того моменту, доки не буде досягнуто одного або кількох термінальних символів, що містять або словоформу, або символ її граматичної категорії. Граматика складників може бути лексикалізована різними способами, напр., символ дієслова (V) було обрано як анкер дерева α в граматиці (10), але замість нього може бути обрано інший анкер – символ іменника (N) (в результаті чого дієслово підпорядковується іменникові на ДЗ, утвореному при виведенні дерева фразових маркерів 11). Також кілька словоформ можуть бути анкером дерева фразових маркерів у ГПД, напр., якщо анкер початкового дерева α в граматиці (10) – дві словоформи – іменник та дієслово, то це дерево більше не містить вузлів, позначених для заміни.

У ГПД формально необхідними є лише вузли, позначені для поєднання – початкові і допоміжні дерева можуть не містити вузлів, позначених для заміни. Останні дозволяють створити більш компактну ГПД, але не є обов'язковими для структури такої граматики [20: 19, 21]. При лексикалізації

граматики складників множина початкових дерев утворюється із множини нерекурсивних правил граматики (правил, у яких символи із правої частини не повторюються в лівій частині); множина допоміжних дерев – із множини рекурсивних правил виду ‘ $B \rightarrow a B c$ ’, де компоненти ‘ a ’ і ‘ c ’ можуть поширюватися лише нерекурсивними правилами [20:20].

Лінгвістична інтерпретація формального ядра ГПД пов’язана із теоретичними проблемами синтаксису речення і словосполучення. По-перше, структура виведення ГПД – дерево залежностей (напр., ДЗ 12) – розрізняє два класи синтаксичного зв’язку між словоформами: зв’язок, утворений операцією заміни вузлів (позначений на схемі (6) суцільною лінією) та зв’язок, утворений операцією поєднання дерев фразових маркерів (позначений пунктирною лінією).

Операція заміни вузлів формує формально-синтаксичний мінімум речення або групи службового слова, а операція поєднання дерев – поширює таку мінімальну структуру за рахунок підрядних словосполучень. Напр., формально-синтаксичний мінімум для речення – структура іменника і дієслова (N V), для прийменникової групи – прийменник і іменник (Prep N), для групи сурядного сполучника, що поєднує однорідні присудки, – сполучник та дієслово (Conj V).

У формально-синтаксичному мінімумі мовних конструкцій напрямок синтаксичного зв’язку між компонентами не визначається, оскільки можливі декілька способів лексикалізації нерекурсивних правил у граматиці складників, і всі лексичні компоненти у такому дереві фразових маркерів фактично рівноправні. Операція заміни вузлів у ГПД, яка оформлює певний вид синтаксичного зв’язку, формально не є необхідною. ГПД дозволяє трактувати зв’язок між словоформами у формально-синтаксичному мінімумі речень і в групі службових слів як зв’язок з можливою зміною напрямку. Наявність цього виду зв’язку є обов’язковою в синтаксичній структурі речення або групи службового слова. За цими ознаками такий вид синтаксичного зв’язку протиставляється односторонньому підрядному зв’язку, який виникає між СК і структурами, які їх поширюють.

Для підрядних словосполучень – конструкцій, які поширюють формально-синтаксичний мінімум речення або групи службового слова лексикалізація рекурсивних правил граматики складників однозначно задає напрямок синтаксичного зв’язку. В ГПД такі конструкції моделюються як допоміжні дерева фразових маркерів. Зв’язок конструкцій цього виду із компонентами формально-синтаксичного мінімуму має факультативний характер. Такий вид зв’язку завжди здійснюється в межах підрядних словосполучень і є одностороннім підрядним зв’язком.

По-друге, ГПД можна інтерпретувати як модель синтаксичних прогнозувань, властивих граматичним значенням словоформ, тобто як модель взаємодії граматичних значень і формально-синтаксичної структури речення. Лексикалізовані синтаксичні структури – це фрагменти загальної структури речення, що прогнозуються граматичними значеннями словоформ – анкерів певної структури. Ті словоформи, граматичні значення

яких прогнозують відповідне дерево, можуть бути узагальнено представлені у ГПД саме як набір цих граматичних значень. Така узагальнена репрезентація дозволяє абстрагуватися від конкретного лексичного наповнення повнозначних словоформ і від тих граматичних значень, які не пов'язані із прогнозуванням відповідного фрагменту синтаксичної структури. Відокремлення граматичних значень, які прогнозують синтаксичну структуру, від супровідних значень здійснюється за трансформаційним критерієм: якщо зміна значення граматичної категорії або самої категорії призводить до перебудови синтаксичної структури мовного конструкта, таке значення або категорія прогнозує відповідну синтаксичну структуру, а отже, здійснює її лексикалізацію; і навпаки, якщо трансформація граматичного значення не впливає на синтаксичну структуру, таке значення не є релевантним для лексикалізації структури. Напр., дієслівна граматична категорія, категорія особовості, а також збіг граматичних значень категорії числа іменника та дієслова – прогнозують формально-синтаксичний мінімум речення (початкове дерево ГПД, яке включає підмет і присудок речення); прийменникова та іменникова граматичні категорії, лексична реалізація прийменника та граматичне значення відмінка іменника прогнозують формально-синтаксичний мінімум прийменникової групи (структуру допоміжного дерева ГПД, що включає прийменник та відмінкову форму іменника). Ці значення формують структуру анкерів відповідних дерев у ГПД.

Граматичні категорії дієслівного часу і виду є супровідними у формально-синтаксичному мінімумі речення, оскільки зміна їх значень не призводить до перебудови синтаксичної структури, і впливає лише на семантику речення. Ці граматичні категорії не пов'язані із синтаксичними прогнозуваннями у поверхневій структурі, тому ці категорії не репрезентуються у ГПД (в структурі анкерів відповідних дерев). Так само, в структуру анкерів ГПД не входить лексична реалізація повнозначних словоформ: поверхневу структуру речення прогнозують лише набори їхніх граматичних значень.

Синтаксичні прогнозування пов'язують фрагменти формально-синтаксичної структури речення, аналогічно до того, як лексико-семантичні валентності лексем пов'язують фрагменти семантико-синтаксичної структури. У формально-синтаксичному мінімумі мовних конструктів синтаксичні прогнозування є двосторонніми (так само, як і синтаксичний зв'язок між компонентами цієї конструкції). У підрядних словосполученнях, що розширюють формально-синтаксичний мінімум, напрямок синтаксичних прогнозувань завжди протилежний напрямкові синтаксичного зв'язку між компонентами: прогнозування здійснюється залежним компонентом і спрямоване від залежного до опорного компонента. Напр., прислівник і прийменникова група прогнозують дієслово, прикметник прогнозує іменник тощо. Односторонні прогнозування репрезентовані в структурі допоміжних дерев ГПД – прогнозування допоміжного дерева – це його вершина (і вузол, який позначений для поєднання, що тотожний вершині): напр., вершина в

допоміжному дереві, лексикалізованому прислівником або прийменниково-відмінковою конструкцією – це дієслівна група, отже прислівник і прийменниково-відмінковий комплекс прогноують дієслово; вершина в допоміжному дереві, лексикалізованому прикметником – прикметникова група, а отже, прикметник прогноує іменник у формально-синтаксичній структурі речення.

Комплекс граматичних значень словоформ прогноує граматичні значення інших словоформ і фрагмент синтаксичної структури речення – систему синтаксичних зв'язків і складників, у які входять словоформи із заданими граматичними значеннями. Прогнозована структура – дерево фразових маркерів ГПД, лексикалізоване комплексом граматичних значень словоформ є одиницею лексикографічного моделювання.

Взаємодія між синтаксичною структурою речення і генеративними моделями синтаксичних структур відбувається у межах системи формальної граматики, яка включає в себе формальну граматику, виведення, результат застосування граматики і структуру виведення. Система формальної граматики розглядається не лише як модель породження речень (див. [11]), а як спосіб визначення відношення між (а) елементарними синтаксичними одиницями – конструкціями, (б) неелементарними синтаксичними одиницями – конструкторами і/ або (в) синтаксичною структурою неелементарних одиниць.

Модель формально-синтаксичних структур розглядається в контексті більш загальної семіотичної моделі мовного синтаксису, запропонованої у функціональній лінгвістиці [3,7], зокрема розглянуто використання генеративно-трансформаційного підходу для моделювання формально-синтаксичної структури речення. Об'єктом подальшого дослідження є моделювання за допомогою розроблених методів інших двох семіотичних рівнів організації речення – семантико-синтаксичної та комунікативно-синтаксичної структури. Формальні генеративно-трансформаційні методи не лише не суперечать функціональному підходові до синтаксису, але доповнюють його, природно інтегруючись у функціональну семіотичну модель.

Така модель синтаксичних структур дозволяє більш чітко сформулювати такі поняття функціональної лінгвістики, як синтаксичний зв'язок, одиниця синтаксичної системи мови і синтаксична структура речення. Окремі словоформи поєднуються в реченні, по-перше, як елементи лінійного ланцюжка словоформ, а по-друге, як елементи нелінійної організації речення – його синтаксичної структури, яка складається із системи складників (груп, або множин словоформ, у яких визначено центральний елемент).

Словоформи можуть прогноувати появу інших словоформ або їхніх груп, входження їх до складу однієї групи та їхню взаємну нелінійну організацію у цій групі. Елементарна одиниця формально-синтаксичної системи мови – це граматична конструкція, тобто система формально-синтаксичних прогноувань окремої словоформи або окремої граматичної категорії та її значення у повнозначній словоформі. Елементарні синтаксичні

одиниці є фрагментами загальної формально-синтаксичної структури речення. Шляхом поєднання слівформ у лінійну послідовність одночасно із поєднанням формально-синтаксичних прогнозувань цих слівформ у єдину зв'язну синтаксичну структуру утворюються речення і словосполучення – структурно та лінійно організовані синтаксичні конструкти. Відношення між системою елементарних формально-синтаксичних одиниць (граматичних конструкцій) та формально-синтаксичними структурами речень і словосполучень (мовних конструктів) моделюється формальною граматиною поєднання дерев.

Синтаксична структура речення може бути побудована не лише як система складників, але і як система синтаксичних зв'язків між слівформами. Поняття синтаксичного зв'язку є більш складним і походить від поняття складника. Синтаксичний зв'язок – це структурне формально-синтаксичне відношення між слівформами, яке утворюється в процесі поєднання систем формально-синтаксичних прогнозувань цих слівформ. Поєднання синтаксичних зв'язків окремих слівформ утворює зв'язний граф (як правило – дерево) синтаксичних залежностей. Кожна система складників речення відповідає лише одному узгодженому графові синтаксичних залежностей, але граф синтаксичних залежностей може відповідати одночасно декільком узгодженим системам складників. Репрезентація синтаксичної структури у вигляді системи складників є більш простою та інформативнішою порівняно із її репрезентацією у вигляді графа синтаксичних залежностей.

Література:

1. Бабич Б.В. Лексична семантика в синтаксичній структурі тексту: формальне представлення та інтерпретація // Матеріали Всеукраїнської наукової конференції “Семантика, синтактика і прагматика мовленнєвої діяльності.— Львів: “Літопис”, 1999. — С. 101-108.
2. Бархударов Л.С. К вопросу о поверхностной и глубинной структуре предложения // Вопросы языкознания – 1973. – №3. – С.50-61.
3. Вихованець І.Р. Граматика української мови. Синтаксис. — К.: Либідь, 1993.— 368 с.
4. Вихованець І.Р., Городенська К.Г., Русанівський В.М. Семантико-синтаксична структура речення. — К.: Наук. думка, 1983. — 220 с.
5. Гладкий А. В. Синтаксические структуры естественного языка в автоматизированных системах общения. — М.: Наука, 1985. — 144 с.
6. Звегинцев В. А. Теоретическая и прикладная лингвистика. — М.: Просвещение, 1967.— 338 с.
7. Современный русский язык. Под ред. В.А. Белошапковой. — М.: Высшая школа, 1989. — 800 с.
8. Теньер Л. Основы структурного синтаксиса. — М.: Прогресс, 1988.— 656 с.
9. Шаумян С.К. Теоретические основы трансформационной грамматики // Новое в лингвистике. — Вып. II. / Состав. В.А. Звегинцев. — М.: Изд-во иностр. Лит-ры, 1962. — С. 391-411.

10. Хомский Н. Три модели описания языка. / Пер с англ. // Кибернетический сборник. – Вып.2. — М.: Изд-во иностранной литературы, 1961. — С. 12-238.
11. Хомский Н. Аспекты теории синтаксиса. — М.: Изд-во Моск. ун-та, 1972. — 260 с.
12. Хомский Н. Синтаксические структуры // Новое в лингвистике. — Вып. II. / Состав. В.А. Звегинцев. — М.: Изд-во иностр. лит-ры, 1962. — С. 412-638.
13. Хэррис З.С. Совместная встречаемость и трансформация в языковой структуре // Новое в лингвистике. - II. — М., 1962. — С. 528-637.
14. Aitchison J. Review of “The architecture of language faculty. By Ray Jackendoff” // Language, vol. 74, 1998. — Number 4 — p. 850-851.
15. Cook, V. I., Newson M. Chomsky’s universal grammar. — Oxford, UK, Cambridge, MA: Blackwell, 1996.-369 p.
16. Halliday M.A.K. An introduction to functional grammar. — London: Edward Arnold, 1994.-434 p.
17. Harris Z. Methods in structural linguistics — Chicago, 1951. –158с.
18. Hornstein N. Logical form: From GB to minimalism. — Cambridge, MA: Blackwell, 1995.-ix, 267 p.
19. Jackendoff R. The architecture of the language faculty.— Cambridge, MA: MIT Press, 1997.—xvi, 262 p.
20. Joshi A., Schabes Y. Tree adjoining grammars // Salomaa and S.Rosenberg (eds.) Handbook of formal languages and automata. — Berlin: Springer-Verlag, 1997.-P. 1-27.
21. Katz J. J. The underlying reality of language and its philosophical import. — New York: Harper Torchbooks, Harper and Row, Publishers, Inc., 1971.-189 p.

Н.П.Гальона

ЛЕКСИКО-СЕМАНТИЧНІ ВЛАСТИВОСТІ СЛОВА ТА ЙОГО ВАЛЕНТНІСТЬ

Речення як основна одиниця синтаксису і комунікативна одиниця мови формується лексико-граматичними одиницями – словами. Виконуючи синтаксичні функції та певним чином пов’язуючись між собою, вони утворюють цілком завершену конструкцію, яка забезпечує інформативність, змістову достатність висловлення і втілює комунікативні настанови мовця. Компонентний склад синтаксичної побудови, її структуру та особливості семантико-синтаксичних відношень між словами пов’язують передусім з валентністю предиката, яка зумовлюються багатьма мовними та позамовними чинниками.

Аналіз будови речення у зв’язку зі структурою денотативної ситуації, здійснюваний протягом останніх десятиліть у вітчизняному та зарубіжному мовознавстві, сприяв з’ясуванню особливостей вияву валентнісних властивостей слів на синтаксичному рівні. Активно ведуться пошуки й закономірностей формування та виявлення валентної структури предиката у зв’язку з його лексичною семантикою, структурою словозначення та дериваційним потенціалом. „Останнім часом не лише посилилась увага до слова як елемента, який безпосередньо (у вигляді словоформи) входить у