

Катерина ПАРНУС,
orcid.org/0000-0003-1811-0980
викладач кафедри іноземних мов
Державного університету «Житомирська політехніка»
(Житомир, Україна) katemarchuk38@gmail.com

СТВОРЕННЯ ДЕРЕВ СИНТАКСИЧНОГО АНАЛІЗУ: ПРАКТИЧНЕ ЗАСТОСУВАННЯ

Обробка природної мови у парадигмі сучасних тенденцій прикладної лінгвістики є актуальним науковим питанням. Наявний стан розвитку комп'ютерних технологій дає змогу автоматизувати процес обробки природної мови. Це економить ресурс часу на здійснення аналізу мовних конструкцій природної мови, водночас надаючи потенціал до опрацювання таким способом великих обсягів емпіричного матеріалу. З огляду на перспективність означеного напрямку досліджень у статті розглянуто теоретичні та практичні засади обробки природної мови, зокрема засоби автоматичного синтаксичного аналізу на прикладі англomовного текстового матеріалу. Актуальність дослідження зумовлена необхідністю вивчення сполучуваності лексичних одиниць і мотивується нерозробленістю широкого кола як теоретичних, так і прикладних проблем у парадигмі дослідження механізмів природної мови.

У дослідженні подано можливі приклади синтаксичного розбору різних типів речень за допомогою синтаксичних дерев. Перевагою автоматичної обробки іноземної мови є отримання синтаксичного розбору за попередньо заданими параметрами. Недоліком автоматизованих систем обробки природної мови є вузьке функціональне коло. Вони не можуть описувати детально структуру речення та ґрунтовно узагальнювати дані про аналізовану синтаксичну одиницю. Вони можуть описувати синтаксичні відомості або ж семантичну інформацію. Стенфордська програма CoreNLP є однією з найбільш використовуваних у сфері аналізу природної мови. Числові результати обробки текстів надають парсери. Розрізняють висхідний і спадний синтаксичний аналіз, що відрізняються послідовністю побудови дерев розбору. Аналітичні дані узагальнюються в абстрактних синтаксичних деревах. Абстрактне синтаксичне дерево відрізняється від дерева розбору тим, що відсутні вузли й ребра для тих синтаксичних правил, які не впливають на семантику програми.

Синтаксичні дерева використовують в парсерах для проміжкового розуміння програми між деревом розбору (конкретним синтаксичним деревом) і структурою даних, яка потім використовується як внутрішнє уявлення у компіляторі або як інтерпретатор комп'ютерної програми для оптимізації та генерації коду. Можливі варіанти таких структур описуються абстрактним синтаксисом.

Ключові слова: парсер, обробка природної мови, абстрактне синтаксичне дерево, операнди.

Kateryna PARNUS,
orcid.org/0000-0003-1811-0980
Lecturer at the Department of Foreign Language
Zhytomyr Polytechnic State University
(Zhytomyr, Ukraine) katemarchuk38@gmail.com

CREATING OF SYNTAX ANALYSIS TREES: PRACTICAL APPLICATION

Processing of natural language in the paradigm of modern trends in applied linguistics is a topical scientific issue. The current state of development of computer technology makes it possible to automate the process of natural language processing. This saves time on the analysis of language structures of natural language, while providing the potential to process in this way large amounts of empirical material. Given the prospects of this area of research, the article considers the theoretical and practical principles of natural language processing, in particular the means of automatic parsing on the example of English text material. The relevance of the study is due to the need to study the compatibility of lexical units and is motivated by the lack of development of a wide range of both theoretical and applied problems in the paradigm of the study of the mechanisms of natural language.

The study presents possible examples of different types of sentences parsing by using syntactic trees. The advantage of automatic foreign language processing is to obtain parsing according to predefined parameters. Currently, the disadvantage of automated natural language processing systems is a narrow functional range. They cannot describe in detail the structure of the sentence and thoroughly generalize the data about the analyzed syntactic unit. They can describe syntactic information or semantic information. Stanford's CoreNLP program is one of the most widely used in natural language analysis. Numerical results of word processing are provided by parsers. There are ascending and descending parsing, which differ in the sequence of construction of parsing trees. Analytical data are summarized in

abstract syntax trees. An abstract syntax tree differs from a parsing tree in that there are no nodes and edges for those syntactic rules that do not affect the semantics of the program.

Syntax trees are used in parsers to intermediately understand a program between a parsing tree (a particular syntax tree) and a data structure, which is then used as an internal representation in the compiler or as an interpreter of a computer program to optimize and generate code. Possible variants of such structures are described by abstract syntax.

Key words: *parser, natural language processing, abstract syntax tree, operands.*

Постановка проблеми. Сучасний світ прикладної лінгвістики активно модернізується. Неможливо уявити сучасні дослідження без комп'ютерних технологій. Обробка природної мови є одним із основних напрямів роботи з іноземними мовами. Вченими запропонований унікальний алгоритм обробки природної мови. Цей алгоритм запроваджується за допомогою інтернет-мережі та допомагає розбивати речення на синтаксичні структури, маркувати й тагувати, розбивати на іменні фрази. Система працює на основі немаркованих даних і дає можливість працювати, тагуючи речення з мінімальною кількістю розрахунків.

Перевагою автоматичної обробки іноземної мови є отримання синтаксичного розбору за заданими параметрами. На жаль, такі системи не дають детального аналізу речення. Вони не можуть описувати детально структуру речення, що б узагальнило дані про аналізоване речення. Вони можуть описувати синтаксичні відомості (наприклад, додавання тегів до мови, фрагментування й розбір) або семантичну інформацію (напр., слово-сенс, семантичне маркування, ім'я видобування сутностей, а також дозвіл анапора). Тексти-корпори були вручну анотовані з такими структурами даних для того, щоб порівняти продуктивність різних систем. Наявність стандартних критеріїв стимулювала дослідження в обробці природної мови (зокрема йдеться про нейролінгвістичне програмування – НЛП).

Актуальність дослідження полягає в необхідності вивчення сполучуваності лексичних одиниць і зумовлена нерозробленістю широкого кола як теоретичних, так і прикладних проблем. Об'єкти дослідження: власне природні мови; використання комп'ютерних програм для аналізу англійських текстових матеріалів.

Аналіз досліджень. Питанням обробки природної мови присвячена низка робіт у зарубіжній лінгвістиці: Д. Йурафську, Й. Р. М. Реесе, Й. Голдберг, Й. Л. Флеїсс, Б. Левін, Ч. Р. Мені (2013), Ч. Холлінгсворт, В. Ковар. На теренах українських досліджень: Н. Г. Чейлитко (2009: 268), І. Р. Вихованець, А. П. Загнітко (2004), В. Ю. Таранухи, Н. П. Дарчук (2013: 11) та ін. Обробка природної мови вимагає сучасних підходів і досліджень. Існує низка проблем, які є не вивченими та потре-

бують наукової уваги. Зокрема, доцільно виокремити теоретичні аспекти проблеми – вивчення синонімії, граматичних одиниць, сполучуваності граматичних одиниць, законів комбінаторики словосполучень різних типів і розрядів, лексичну валентність як критерій розмежування вільних і фразеологічних словосполучень, взаємодію стійкості й ідіоматичності тощо. Прикладними аспектами проблеми є автоматичний аналіз речень, автоматична обробка природної мови, автоматичне реферування й анотування тексту на основі сполучувальних критеріїв тощо.

Метою дослідження є аналіз проблем при обробці природної мови за допомогою мережевих програм і наступне пояснення їх на основі комп'ютерних і мережевих програм.

Виклад основного матеріалу. Хоча є кілька хороших інструментальних засобів аналізу природної мови, стенфордська програма CoreNLP є однією з найбільш використовуваних. Стенфордські вчені рекомендують використовувати саме CoreNLP як базовий інструмент обробки природної мови (Гирич, 2017: 26).

Парсери зазвичай надають числові результати обробки текстів. Для отримання структурних даних необхідно створити дерево синтаксичного аналізу (Борисова, 2016: 204).

Розрізняють два типи алгоритмів синтаксичного аналізу: спадний синтаксичний аналіз – завдання побудови дерева розбору для вхідного рядка, починаючи з кореня та створюючи вузли дерева розбору у прямому порядку обходу або, що те ж саме, спадний синтаксичний аналіз можна розглядати як пошук лівого породження вхідного рядка. Висхідний синтаксичний аналіз будує дерево розбору для вхідного рядка, починаючи з листя (знизу) і йдучи в напрямку до кореня (угору) (Гирич, 2017: 27).

Абстрактне синтаксичне дерево (АСД) – це в інформатиці кінцеве помічене орієнтовне дерево, у якому внутрішні вершини зіставлені з операторами мови програмування, а листя – з відповідними операндами. Отже, листки вважаються порожніми операторами та становлять лише змінні й константи.

Синтаксичні дерева використовують у парсерах для проміжкового розуміння програми між деревом розбору (конкретним синтаксичним

деревом) і структурою даних, яка потім використовується як внутрішнє уявлення в компіляторі або як інтерпретатор комп'ютерної програми для оптимізації й генерації коду. Можливі варіанти подібних структур описуються абстрактним синтаксисом.

Абстрактне синтаксичне дерево відрізняється від дерева розбору тим, що відсутні вузли й ребра для тих синтаксичних правил, які не впливають на семантику програми. Класичним прикладом такої відсутності є групувальні дужки, оскільки в АСД групування операндів явно задається структурою дерева.

Для мови, що описується контекстно-вільною граматику, якими є майже всі мови програмування, створення абстрактного дерева в синтаксичному аналізаторі є тривіальним завданням. Більшість правил у граматиці створюють нову вершину, а символи у правилі стають ребрами. Правила, які нічого не привносять в АСД (наприклад, групувальні правила), просто замінюються у вершині одним зі своїх символів. Окрім того, аналізатор може створити повне дерево розбору й потім пройти ним, видаляючи вузли й ребра, які не використовуються в абстрактному синтаксисі, для отримання АСД (Chomsky, 1971: 183).

Синтаксичні дерева можуть бути побудовані компілятором для будь-якої частини вхідної програми. Не завжди синтаксичному дереву повинен відповідати фрагмент коду результатної програми. Наприклад, можлива побудова синтаксичних дерев для декларативної частини мови. У цьому разі операції, наявні в дереві, не вимагають породження об'єктного коду, але несуть інформацію про дії, які повинен виконати сам компілятор над відповідними елементами. Коли синтаксичному дереву відповідає деяка послідовність операцій, що тягне породження фрагмента об'єктного коду, говорять про дерево операцій (Полховська, 2013: 32).

Дерево операцій можна безпосередньо побудувати з дерева виведення, породженого синтаксичним аналізатором. Для цього достатньо вилучити з дерева виведення ланцюжок нетермінальних символів, а також вузли, що не мають семантичного навантаження при генерації коду. Прикладом таких вузлів можуть бути різні дужки, які змінюють порядок виконання операції й операторів, але після побудови дерева ніякого смислового навантаження не несуть. Їм не відповідає ніякий об'єктний код. Те, який вузол у дереві є операцією, а який операндом, ніяк неможливо визначити із граматики, котра описує синтаксис вхідної мови. Також нізвідки не впливає, яким опера-

ціям повинен відповідати об'єктний код внаслідок застосування програми, а яким ні. Усе це визначається тільки із семантики, «сенсу», мови вхідної програми. Тому тільки розробник компілятора може чітко визначити, як при побудові дерева операції повинні відрізнятися операнди та самі операції, а також те, які операції є семантично незначущими для породження об'єктного коду (Гирич, 2017: 26).

Алгоритм перетворення дерева семантичного розбору й дерева операції можна уявити так:

Крок 1. Якщо у дереві більше не міститься вузлів, помічених нетермінованими символами, то виконання алгоритму завершено; інакше – перейти до кроку 2

Крок 2. Вибрати крайній лівий вузол дерева, позначений нетермінальним символом граматики, і зробити його поточним. Перейти до кроку 3.

Крок 3. Якщо поточний вузол має тільки один нижчий вузол, то поточний вузол необхідно видалити з дерева, а пов'язаний із ним вузол приєднати до вузла вищого рівня (виключити з дерева ланцюжок) і повернутися до кроку 1; інакше – перейти до кроку 4.

Крок 4. Якщо поточний вузол має нижчий вузол (листя дерева), позначений термінальним символом, який не несе семантичного навантаження, тоді цей листя потрібно видалити з дерева та повернутися до кроку 3, інакше – перейти до кроку 5.

Крок 5. Якщо поточний вузол має один нижчий вузол (листя дерева), позначений термінальним символом, що позначає знак операції, а інші вузли позначені як операнди, то листя, позначений знаком операції, треба видалити з дерева, поточний вузол позначити цим знаком операції та перейти до кроку 1; інакше – перейти до кроку 6.

Метою синтаксичного аналізу є визначення структури вхідного тексту. Ця структура складається з ієрархії фраз, найменші з яких є основними символами та найбільшою з яких є речення. Структуру можна описати деревом з одним вузлом для кожної фрази. Основні символи представлені значеннями, що зберігаються у вузлах. Корінь дерева становить речення.

За теорією Чомського (Chomsky, 1971: 183), побудова синтаксичного дерева має такий вигляд:

1. S NP VP
S – Sentence, NP – Noun Phrase, VP – Verb Phrase
2. NP (Art) (AP) N (PP)
Art – Article, AP – Adjective Phrase, N – Noun, PP – Prepositional Phrase
3. VP (Aux) V (NP)
Aux – Helping verb, V – Verb, NP – Noun Phrase

- 4. PP P (NP)
- P – Preposition
- 5. CP C S

Отже, керуючись описаним правилом, ми можемо побудувати синтаксичний аналіз різних типів речень в англійській мові: простого та складного речення, складнопідрядного речення, складносурядного речення, еліптичних речень, безособових речень і т. д. Більш детально автоматичний синтаксичний аналіз кожного з речень розглянемо далі.

Розповідне речення (*declarative sentence*) – в англійській мові викладає факти у формі ствердження чи заперечення, передає якусь особливу інформацію чи факти реципієнту. У таких реченнях підмет обов'язково стоїть перед присудком. Зазвичай вони вимовляються зі спадною інтонацією. Має часові рамки (Полховська, 2013: 32). Розповідне речення в теперішньому часі може бути простим або складним. Просте розповідне речення складається з підмета та присудка.

Приклади ручного та парсерного гілкування простого речення *The mouse is white* (*Миша біла*) зображено на рис. 1 і рис. 2.

[S [NP [mouse]] [VP [V is] [^AP white]]]

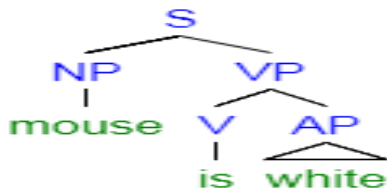


Рис. 1. Ручне гілкування

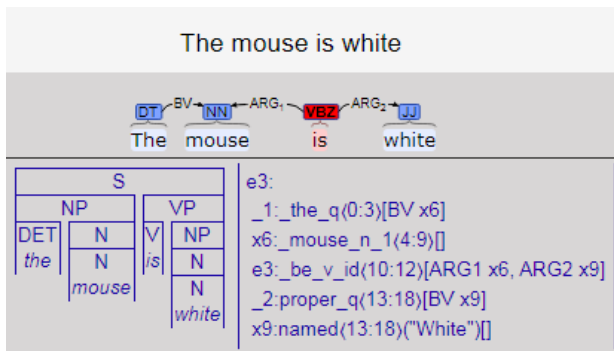


Рис. 2. Гілкування за допомогою парсера

Вручну побудувавши лінійний розбір речення, ми вводимо наш приклад в автоматичний синтаксичний аналізатор (Карпіловська, 2006). Як результат, ми отримуємо синтаксичне дерево, де S – sentence, NP – Noun Phrase *mouse* (іменна

фраза), VP – Verb Phrase (дієслівна фраза), яка членується на V- Verb (дієслово) *is*, та AP – Adjective Phrase (прикметник) *white*.

Складне розповідне речення (*compound declarative sentence*) об'єднує дві взаємопов'язані фрази. Фрази об'єднуються комою або сполучниками, такими як *and*, *yet* or *but*. Сполучниками також можуть виступати крапка з комою або транзитивні сполучники *however*, *besides* або *therefore* (Fleiss et al., 2013: 125; Jackendoff, 1995: 67).

Приклади ручного та парсерного гілкування речення *He wanted to play football, but she wanted to play basketball* (*Він хотів грати у футбол, а вона – у баскетбол*) подано на рис. 3–4.

[S[NP[N he]][VP [V wanted]][VP[COMP to]][VP[V play]][NP[N football]]][S[CONJ but][NP[N she][VP[V wanted]][VP[V play]][NP[N basketball]]]]]

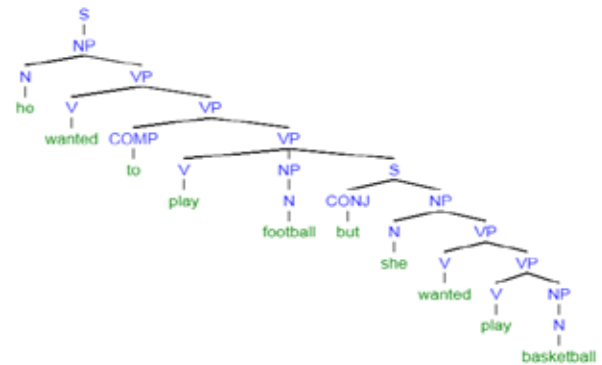


Рис. 3. Зразок ручного гілкування складного розповідного речення

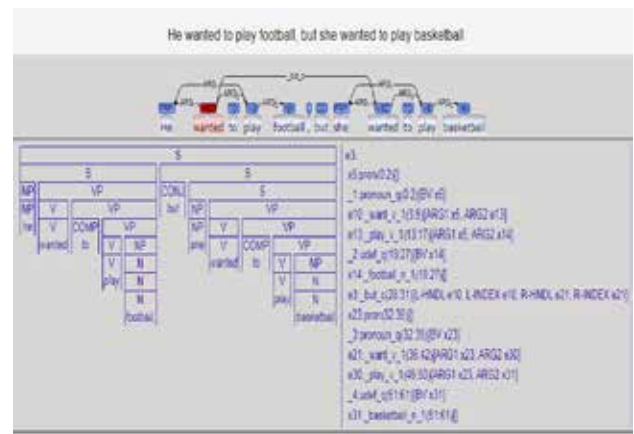


Рис. 4. Зразок гілкування парсером складного розповідного речення

Результатом є набагато ширше синтаксичне дерево, яке відображає синтаксичні зв'язки в цьому реченні. Складне розповідне речення гілкується іншим чином, якщо є дві VP (verb phrase).

Синтаксичні зв'язки у наведеному реченні мають такий вигляд: S(sentence) – NP – VP; NP, де N(noun) – he; VP(verb phrase), де V(verb) – want, COMP – to; VP(verb phrase), де V(verb) – play; NP(noun phrase), де N(noun) – football, CONJ – but; NP(noun phrase) де N(noun) – she, VP(verb phrase), де V(verb) – play і NP(noun phrase), де N(noun) – basketball.

Розповідні речення мають певний період часу та містять певну інформацію. Вони також виражають чиюсь думку чи позицію.

Наведемо приклад речення, що містить інформацію про певну позицію та думку, а також проаналізуємо його за допомогою парсера й ручного гілкування (рис. 5–6).

Наприклад: *I love my cat* (Я люблю мого кота).
[S[NP[N I]][VP[V love]][NP[AP my][N[cat]]]]

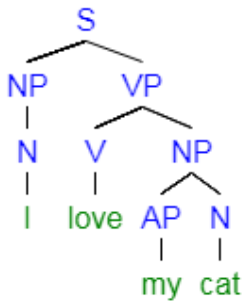


Рис. 5. Зразок гілкування вручну розповідного речення

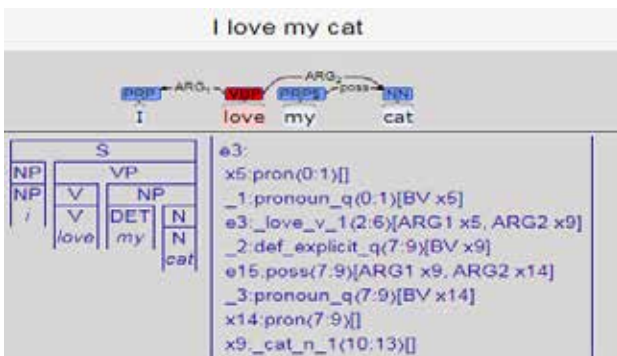


Рис. 6. Зразок гілкування розповідного речення за допомогою парсера

Наведений приклад виражає думку певної особи. Окрім того, ми простежуємо такі семантичні зв'язки: S NP VP, де S(sentence) NP(noun phrase), де N(noun) – I, VP(verb phrase), де V(verb) – love, NP(noun phrase), де DET – my, N(noun) – cat.

Заперечні розповідні речення – речення, що несуть певну інформацію з використанням заперечного слова. Такими словами є: заперечна частка

not; заперечні займенники no (ніякий, не); nobody, no one (ніхто); nothing (ніщо, нічого); заперечні прислівники never (ніколи); nowhere (ніде); заперечний сполучник neither...nor (ні...ні) (Stern and Dagan, 2013: 123).

Зразки ручного та парсерного гілкування заперечних речень подано на рис. 7–8.

Наприклад: *My cat is not black* (Мій кіт не чорний).

[S[NP[AP my][N[cat]]][VP[V[is]][NP[NegP not][AP[black]]]]]

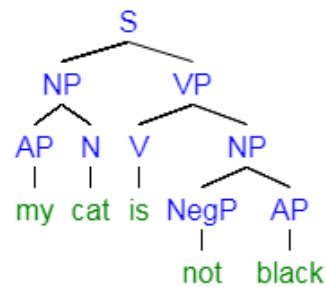


Рис. 7. Зразок гілкування заперечного речення

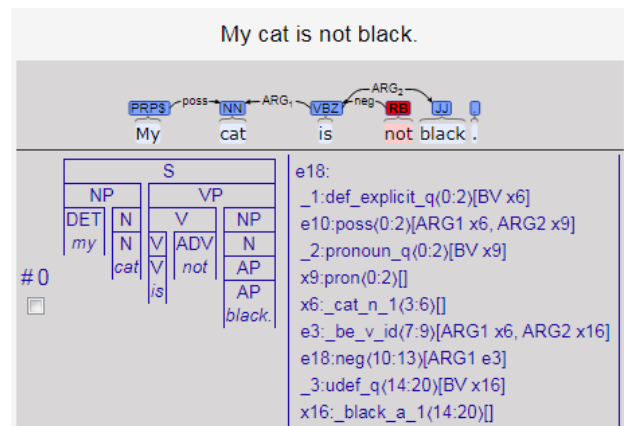


Рис. 8. Зразок гілкування заперечного речення парсером

Інверсія в розповідному реченні трапляється, до прикладу, коли прислівники ставляться на початку речення, а дієслово ставиться перед підметом (Stern and Dagan, 2013: 123). Порівняймо автоматичний синтаксичний аналіз речення без інверсії та з інверсією (рис. 9–10).

Наприклад:

The game had scarcely started when the rain came pouring down (Гра ледве почалася, як пішов проливний дощ).

[S[NP[Art the][N game]][VP[V had][VP[ADV scarcely]][VP[V started]][PP[P when]][S[NP[Art the][N rain]][VP[V came]][PP[VP[V powering][P down]]]]]]]

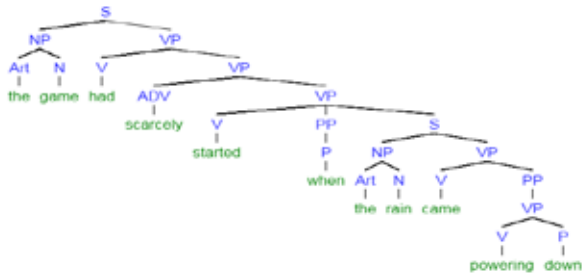


Рис. 9. Зразок гілкування речення вручну



Рис. 12. Зразок синтаксичного розбору речення з інверсією за допомогою парсера



Рис. 10. Зразок розбору речення за допомогою парсера

[S[PP[NP where]][PP[V do][NP we]][VP[Vmake]][NP[DET our][N[AP next][N turn]]]

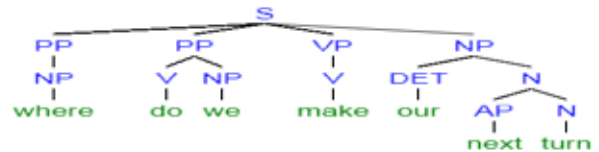


Рис. 13. Зразок синтаксичного розбору питального речення вручну

У цьому зразку простежуємо прямий порядок слів, а наступним прикладом є інверсія вказаного речення рис. 11–12.

Наприклад: *Scarcely had the game started when the rain came pouring down* (Ледве почалася гра, як пішов проливний дощ).

[S[VP[ADV scarcely]][VP[V had][Art the][NP[N game]][S[VP[V started]][PP[P when][S[NP[Art the][N rain][VP[V came]][PP[VP[V powering][P down]]]]]]]]]

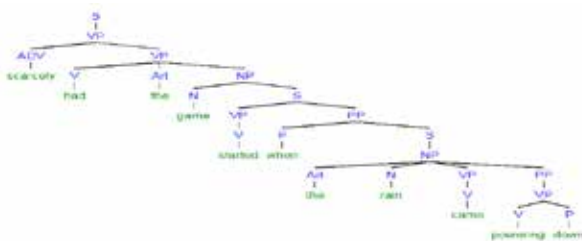


Рис. 11. Зразок синтаксичного розбору речення з інверсією вручну

Питальні речення (*an interrogative sentence*) – речення, представлене питальним словом на початку та знаком питання в кінці (Fleiss and Myunghee, 2013; Кобрина, 1986: 38). Зазвичай підмет стоїть після питального слова. Таке речення висловлюють із метою отримання певної інформації. Зразки розбору питального речення продемонстровано на рис. 13–14.

Наприклад: *Where do we make our next turn?*
Де ми повертаємо наступного разу?



Рис. 14. Зразок автоматичного синтаксичного розбору речення за допомогою парсера

Складне речення, що складається із двох або більше предикативних одиниць, поєднаних у єдине семантичне та граматичне ціле підрядним зв'язком за допомогою сполучників підрядності чи сполучних слів та інших граматичних засобів, називається складнопідрядним (Гузеєва, Костыгіна, 2000: 28; Кубрякова та ін., 1996: 74–77).

Предикативні частини складнопідрядного речення синтаксично нерівноправні: одна з них підпорядкована іншій (Карамішева, 2005: 69). Отже, частина складнопідрядного речення, що перебуває в синтаксичній залежності від іншої, головної частини, називається підрядною. Частина складнопідрядного речення, якій підпорядкована підрядна частина, називається головною (Загнітко, 2004: 87–89).

Атрибутивні речення – речення з означенням. Означення (*the Attribute*) – це другорядний член

речення, який показує ознаку предмета. Будь-який член речення, виражений іменником, може мати одне або кілька означень (Fleiss and Myunghee, 2013: 127).

В англійській мові є два види означень, котрі розрізняються за їх місцем відносно означуваного слова: препозитивні означення, які стоять перед означуваним іменником, і постпозитивні означення, що йдуть за означуваним, іменником (Долгополова, 2010: 43–51).

Атрибутивні (предикатно-атрибутивні) семантико-синтаксичні відношення виформовуються між предметом і його потенційно предикативною ознакою (Stern and Dagan, 2013: 142; Борисова, 2016: 204). Зразки аналізу речень за вказаною парадигмою графічно наведено на рис. 15–16.

Наприклад: *The Kyiv Metro is the most beautiful underground railway in the whole world* (Київське метро – найкрасивіше метро в усьому світі).

[S [NP[Atr the][N[N Kyiv][N metro]]][VP[V is][NP[AP[Art the][AP beautiful]]][N[AP underground][NP[N railway][PP[P in][NP[Art the][N[AP whole][N world]]]]]]]

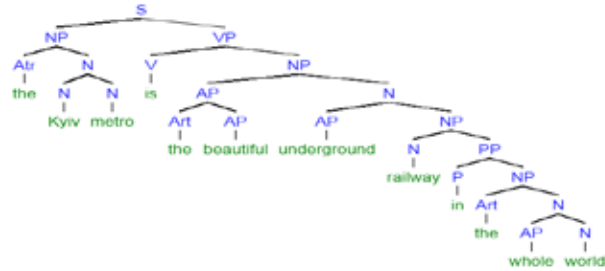


Рис. 15. Зразок синтаксичного розбору речення вручну

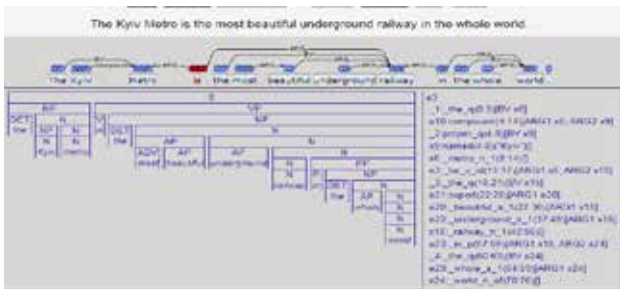


Рис. 16. Зразок розбору речення за допомогою парсера

Постпозитивне означення ставиться безпосередньо після означуваного іменника. Постпозитивне означення може бути виражене: іменником або герундієм із прийменником; означальні сполучення з іменниками утворюють такі прийменники: *about, for, from, in, on, to, with, without, to*.

Розглянемо речення з постпозитивним означенням *about*.

Наприклад:

about: I like books about great travellers.

[S[NP I][VP[NP[N books][PP[P about][NP[AP great][N travellers]]]]]



Рис. 17. Зразок розбору речення із постпозитивним означенням вручну



Рис. 18. Зразок розбору речення із постпозитивним означенням за допомогою парсера

Розглянемо на прикладі речення із постпозитивним означенням *to* (рис. 19–20). Наприклад, *to: Here is the key to the door.* – Ось ключ від дверей.

[S[NP[N here]]][VP[V is][NP[N[N key][PP[P to][NP[Art the][N door]]]]]]]

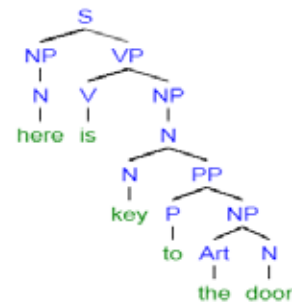


Рис. 19. Зразок синтаксичного розбору речення із постпозитивним означенням *to*

Усі речення англійської мови можуть бути простими (*simple sentences*), складносурядними (*compound sentences*) і складнопідрядними (*complex sentences*). Прості речення можна роз-

бити на дві групи – поширені (*extended*) і непоширені (*unextended*). В останніх присутні тільки головні члени речення (підмет і присудок), а ось другий доповнені ще і другорядними (означеннями, обставинами, доповненнями).

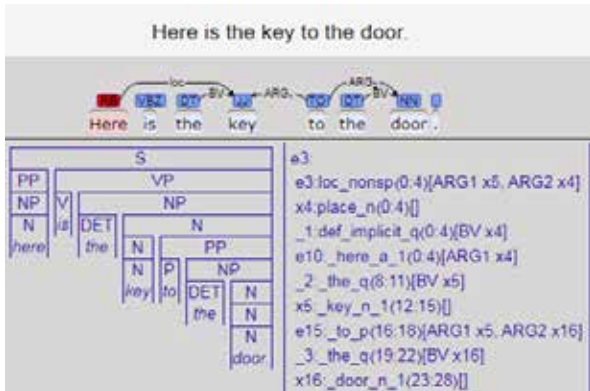


Рис. 20. Зразок автоматичного синтаксичного розбору речення із постпозитивним означенням «to» за допомогою парсера

Окрім того, варто пам'ятати, що всі речення в англійській мові мають різні семантичні зв'язки. Семантичні зв'язки – це глибинний зміст речення, який ми можемо зрозуміти за допомогою синтаксичних аналізаторів, парсерів, граберів і т. д., що значно полегшує роботу науковця й студента в процесі вивчення мови на її глибинних рівнях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Борисова Н. В., Кочуєва З. А., Оліфенко І. В. Метод автоматизованої лематизації дієслів німецької мови. *Вісник національного університету «Львівська політехніка». Серія: «Інформаційні системи та мережі»*. 2016. № 854. С. 204–210. URL: http://vlp.com.ua/files/1_2_zmist_854.pdf (дата звернення: 11.04.2021).
2. Гирин О. В. Автоматичний синтаксичний аналіз англійської мови: застосування та перспективи. *Вісник Житомирського державного університету імені Івана Франка. Філологічні науки*. 2017. Вип. 1. С. 26–30. URL: http://nbuv.gov.ua/UJRN/VZhDUF_2017_1_7 (дата звернення: 11.04.2021).
3. Гладкий А. В. Синтаксические структуры естественного языка в автоматизированных системах общения. *Проблемы искусственного интеллекта*. Москва : Наука, 1985. 144 с.
4. Гузеева К. А., Костыгина С. И. Инфинитив. *Грамматика английского языка*. Санкт-Петербург : Союз, 2000. 219 с.
5. Долгополова Л. А. Инфинитив в истории немецкой мови (IX–XX ст.) : монографія. Київ : Вид. центр КНЛУ, 2010. 207 с.
6. Дарчук Н. П. Автоматичний синтаксичний аналіз текстів корпусу української мови. *Українське мовознавство*. 2013. № 43. С. 11–19.
7. Загнітко А. П. Основи українського теоретичного синтаксису. Ч. 1. Горлівка, 2004. 367 с.
8. Карпіловська Є. А. Комп'ютерна лінгвістика. *Енциклопедія Сучасної України* : електронна версія / гол. редкол. : І. М. Дзюба, А. І. Жуковський, М. Г. Железняк та ін. ; НАН України, НТШ. Київ : Інститут енциклопедичних досліджень НАН України, 2006. URL: http://esu.com.ua/search_articles.php?id=4396 (дата звернення: 11.04.2021).
9. Кармишева І. Д. Структурні та функціональні особливості вторинної предикації в сучасній англійській мові (досвід формально-граматичного моделювання) : дис. ... канд. філол. наук : 10.02.04. Київ, 2005. 320 с.
10. Кобрин Н. А., Корнеева Е. А., Оссовская М. И. *Грамматика английского языка. Синтаксис : учебное пособие*. Москва : Просвещение, 1986. 158 с.
11. Кубрякова Е. С., Демьянков В. З., Панкрац Ю. Г., Лузина Л. Г. *Краткий словарь когнитивных терминов*. Москва, 1996.
12. Чейлитко Н. Г. Корпусне дослідження зон зв'язків словоформ в українському реченні. *Лінгвістичні студії*. 2009. Вип. 18. С. 268–275.

13. Полховська М. В. Аналіз англійських медіальних конструкцій з позиції генеративної граматики. *Studia philologica*. 2013. Вип. 2. С. 32–36.
14. Chomsky, N. Deep structure, surface structure, and semantic interpretation. *Semantics*. 1971. P. 183–216.
15. Jackendoff, R. X. *Syntax: A study of Phrase structure*. Cambridge : MIT Press, 1995.
16. Fleiss, J. L., Levin, B., Myunghee, Ch. P. *Statistical methods for rates and proportions*. New Jersey : John Wiley & Sons, 2013. 800 p.
17. Stern A., Dagan, I. The BIUTEE research platform for transformation-based textual entailment recognition. *Linguistic Issues in Language Technology*. 2013. № 9. P. 120–146.

REFERENCES

1. Borisova N. V, Kochueva Z. A., Olifenko I. V. Metod avtomatyzovanoi lematyzatsii diiesliv nimetskoj movy [Method of automated lemmatization of German verbs]. *Bulletin of Lviv Polytechnic National University. Series: "Information systems and networks"*. 2016. № 854. P. 204–210. URL: http://vlp.com.ua/files/1_2_zmist_854.pdf (Date of access: 11.04.2021). [in Ukrainian]
2. Girin O. V. Avtomatychnyi syntaksychnyi analiz anhliiskoi movy: zastosuvannya ta perspektyvy [Automatic parsing of the English language: application and prospects]. *Bulletin of Zhytomyr State University named after Ivan Franko. Philological sciences*, 2017, № 1, pp. 26–30. URL: http://nbuv.gov.ua/UJRN/VZhDUF_2017_1_7 (Date of access: 11.04.2021) [in Russian]
3. Gladkiy A. V. Sintaksicheskiye struktury estestvennogo yazyka v avtomatizirovannykh sistemakh obshcheniya. Problemy iskusstvennogo intellekta [Syntactic structures of natural language in automated communication systems]. *Problems of artificial intelligence*. Moscow: Nauka, 1985. 144 p. [in Russian]
4. Guzeeva K. A. , Kostygina, S. I. Infinitiv. Grammatika angliyskogo yazyka [Infinitive. English grammar]. Saint Petersburg, 2000. [in Russian]
5. Dolgoplova LA Infinityv v istorii nimetskoj movy (IX–XX st.) : monohrafiia [Infinitive in the history of the German language (IX–XX centuries.): Monograph]. Kyiv: Ed. KNLU Center, 2010. [in Ukrainian]
6. Darchuk N. P. Automatic parsing of texts of the corpus of the Ukrainian language [Avtomatychnyi syntaksychnyi analiz tekstiv korpusu ukrainskoj movy]. *Ukrainian linguistics*, 2013, № 43, pp. 11–19. [in Ukrainian]
7. Zagnitko A. P. Osnovy ukrainskoho teoretychnoho syntaksysu [Fundamentals of Ukrainian theoretical syntax]. P. 1. Horlivka, 2004. [in Ukrainian]
8. Karpilovska E. A. Kompiuterna lingvistyka [Computational linguistics]. *Encyclopedia of Modern Ukraine: electronic version* / ch. editors: I. M. Dziuba, A. I. Zhukovsky, M. G. Zheleznyak, etc. ; NAS of Ukraine, NTSh. Kyiv: Institute of Encyclopedic Research of the National Academy of Sciences of Ukraine, 2006. Available at: http://esu.com.ua/search_articles.php?id=4396 [in Ukrainian]
9. Karamisheva I. D. Strukturni ta funktsionalni osoblyvosti vtorynnoi predykatsii v suchasni anhliiskii movi (dosvid formalno-hramatychnoho modeliuвання) [Structural and functional features of secondary predication in modern English (experience of formal-grammatical modeling)]: PhD dis. ... : specialty 10.02.04. Kyiv, 2005. [in Ukrainian]
10. Kobrina N. A., Korneeva E. A., Ossovskaya M. I. Grammatika angliyskogo yazyka. Sintaksis : ucheb. posob [Grammar of the English language. Syntax: textbook. manual]. Moscow: Education, 1986. [in Russian]
11. .Kubryakova E. S., Demyankov V. Z., Pankrats Yu. G., Luzina L. G. Kratkiy slovar kognitivnykh terminov [A short dictionary of cognitive terms]. Moscow, 1996. [in Russian]
12. Cheylytko N. G. Korpusne doslidzhennia zon zviazkiv slovoform v ukrainskomu rechenni [Corpus study of word zones in Ukrainian sentences]. *Linguistic studies: collection of scientific works*, 2009, Vol. 18, pp. 268–275. [in Ukrainian]
13. Polkhovskaya M. V. Analiz anhliiskykh medialnykh konstruksii z pozytsii heneratyvnoi hramatyky [Analysis of English medial constructions from the standpoint of generative grammar]. *Studia philologica*, 2013, Vol. 2, pp. 32–36. [in Ukrainian]
14. Chomsky N. Deep structure, surface structure, and semantic interpretation. *Semantics*, 1971, pp. 83–216.
15. Jackendoff, R. X. *Syntax: A study of Phrase structure*. Cambridge: MIT Press, 1995.
16. Fleiss J. L., Myunghee, Ch. P. *Statistical methods for rates and proportions*. John Wiley & Sons, 2013.
17. Stern A., Dagan, I. The Biutee research platform for transformation-based textual entailment recognition. *Linguistic Issues in Language Technology*, 2013, № 9, pp. 120–146.