

У цьому аспекті WordNet більше нагадує тезаурус, аніж словник. Однак, на відміну від традиційних тезаурусів, WordNet (хоча він і характеризується певною надлишковістю входжень, зумовленою потребою щоразу експлікувати ту чи іншу лінгвістичну інформацію) влаштовано у такий спосіб, що зазначені процедури, завдяки можливостям комп'ютерів, виконуються миттєво, а результат пошуку є явним і повним. Так, якщо в друкованому виданні алфавітного тезауруса X і Y є синонімами, то ця пара повинна бути наведеною двічі — на X і на Y, а в тематичному тезаурусі її пошук взагалі вимагатиме двох процедур — спочатку в алфавітному списку, а потім власне у тезаурусі — цих кроків користувач WordNet взагалі не помічає.

Проте WordNet — це не простий діалоговий тезаурус. Щоб оцінити його переваги, треба виходити з того, що лексична семантика починається із розуміння факту, що слово реалізує загальноприйнятту асоціацію між лексикалізованим поняттям та синтаксично зумовленим висловлюванням. Це визначення зумовлює появу комплексу з трьох аспектів дослідження: по-перше, які види висловлювань вступають у ці лексичні асоціації; по-друге, що являють собою і як організовані поняття, що виражаються словами; і по-третє, яку синтаксичну роль виконують різні слова. Не ігноруючи жодної з цих трьох проблем, теорія WordNet зосереджується на другому колі питань, що має безпосереднє відношення до семантичної структури англійського лексикону (та й лексикону будь-якої мови).

Оскільки концепт “слово”, як правило, вживається для позначення як висловлювання, так і його асоціативного поняття, виникає певна термінологічна двозначність. Задля її уникнення формою слова (чи словоформою) будемо називати позначення “фізичного” висловлювання чи напису, а значенням слова — позначення лексикалізованого поняття, для виразу якого використовується ця форма. Тому дослідження лексичної семантики у такій спрощеній постановці полягає у встановленні відповідності між формами та значеннями слів.

Якщо існують дві різні інтерпретації (у “правій” частині словникової статті) для однієї словоформи, то вона є полісемічною; якщо ж різні словоформи мають однакові інтерпретації то вони є синонімічними. З погляду психолінгвістики, полісемія і синонімія виявляють проблеми, зумовлені пошуком відповідного доступу до інформації в ментальному лексиконі.

У своїй “класичній” концепції WordNet має справу винятково зі встановленням семантичних відношень між лексикалізованими по-

няттями, тобто WordNet можна кваліфікувати як теорію поля значення слова. Проте не можна ігнорувати лексичні відношення і між формами слова, адже семантичні відношення організують WordNet. Оскільки семантичні відношення — це відношення між значеннями та оскільки синонімічні ряди репрезентують значення, природно вважати семантичні відношення покажчиками синонімічних рядів. Характерно, що семантичні відношення є симетричними: якщо існують семантичні відношення  $R$  між значенням  $\{x, x', \dots\}$  і значенням  $\{y, y', \dots\}$ , тоді є також відношення  $R'$  між  $\{y, y', \dots\}$  і  $\{x, x', \dots\}$ . У рамках обговорюваної теми імена семантичних відношень відіграватимуть подвійну роль: якщо  $R$  — відношення між значеннями  $\{x, x', \dots\}$  та  $\{y, y', \dots\}$ , тоді  $R$  буде також застосовуватися для позначення відношень між індивідуальними формами слова, які належать до тих самих синонімічних рядів.

У WordNet представлено такі типи відношень: синонімія; антонімія; гіпонімія; меронімія; морфологічні відношення.

Коротко зупинимося на їхній характеристиці.

В основу розуміння поняття “синонімія” у системі WordNet покладено наступне: два слова є синонімічними в контексті  $S$ , якщо заміна одного на інше в  $S$  не змінює його істинності. Отже, синонімічні ряди можуть утворювати лише слова однієї частини мови, а слова різних синтаксичних категорій не можуть бути синонімами. Іншими словами, використання синонімічних рядів для представлення значень слів пов'язується з психолінгвістичним поглядом на те, що іменники, дієслова і модифікатори є незалежно організованими в семантичній пам'яті. На користь цього аргументові може служити той факт, що деякі слова однієї і тієї ж синтаксичної категорії (в основному дієслова) виражають дуже подібні поняття, хоча не можуть бути взаємозамінними, оскільки речення стає граматично неправильним.

Антонімія — це відношення протилежності, що являє собою лексичне відношення між словоформами, а не семантичне відношення між значеннями слів. Це примушує розрізняти семантичні відношення між словоформами і семантичні відношення між значеннями слів. Антонімія у WordNet є центральним організуючим принципом для прикметників і прислівників.

На відміну від синонімії та антонімії, які є лексичними відношеннями між словоформами, гіпонімія і гіперонімія є семантичними відношеннями між значеннями (наприклад, “клен” є гіпонімом до “дерево”, “дерево” є гіпонімом до “рослина”). Оскільки звичайно поняття має тільки один гіперонім, він породжує ієрархічну се-

мантичну структуру, в якій вважається, що гіпонім належить гіпероніму. Гіпонім вбирає в себе усі риси загального поняття і додає риси, що відрізняють його від гіпероніма і від інших гіпонімів цього гіпероніма. Таким чином забезпечується центральний принцип організації іменників у системі WordNet.

Іншим типом відношень є "частина-ціле", що відомі як меронімія / голонімія. Поняття, представлене набором  $\{x, x', \dots\}$ , є меронімом поняття, представленого набором  $\{y, y', \dots\}$  якщо носій мови сприймає речення, сконструйоване з такої основи, як: *Y має X (як частину)* або *ж X є частиною Y*. Відношення меронімії можуть використовуватися для побудови ієрархії частин (з деякими обмеженнями, оскільки меронім може мати багато голонімів).

Морфологічні відношення між словоформами є важливою частиною лексичних відношень. Проте в системі WordNet спочатку інтерес обмежувався лише семантичними відношеннями; не було проєктів, які б враховували морфологічні відношення. Проте з часом програми удосконалювалися і стало очевидним, що WordNet повинен слугувати будь-яким запитам, у тому числі пов'язаним із морфологічними характеристиками (а саме, із розвиненою словозмінною системою).

На сьогодні в системі WordNet для англійської мови розроблено програмне забезпечення щодо морфологічної парадигми. Проте питання щодо дериваційної морфології, наскільки нам відомо, поки що не реалізовано.

Розглянемо детальніше принцип подання інформації у WordNet на прикладі іменників. Дефініції іменників звичайно подаються через узагальнюючий термін та характеристичні властивості; ця інформація забезпечує основу для організації іменникових файлів у WordNet. Відношення узагальнення формує ієрархічну семантичну організацію, яка дублюється в іменникових файлах шляхом використання маркованих покажчиків між синонімічними рядами (синсетами). Ієрархія має обмежену глибину, яка рідко перевищує дванадцять рівнів. Характеристичні властивості вводяться у такий спосіб, щоб створити систему лексичного наслідування — систему, в якій кожне слово успадковує характеристичні властивості всіх своїх узагальнюючих понять. Розглядаються три типи характеристичних властивостей: атрибути (модифікація), частини (меронімія) та функції (предикація), але в іменникових файлах реалізовано лише меронімію. Антонімія також відбувається між іменниками, але вона не є для них фундаментальним організуючим принципом.

