

Наталя Коциба,
Інститут Славістики
Польської Академії Наук
Варшава, Польща

Семантична функція як засіб вираження категорії означеності в англійській та українській мовах

Щойно в останні роки в славістичних працях почали з'являтися системні описи семантичної категорії означеності. Більшість слов'янських мов (за винятком болгарської та македонської) не має формальних засобів вираження означеності таких, як артиклі в романо-германських мовах. Можливість перекладу текстів з безартиклевих мов на артиклеві свідчить однак про те, що значення означеності передається в перших іншим способом.

Намагаючись з'ясувати як саме це відбувається, ми проаналізували уривки української повісті та їх англійського перекладу, зробленого англо-українським білінгвом. Особливу увагу ми звертали на ті українські іменникові фрази (ІФ), англійські відповідники котрих супроводжувалися означеними артиклями, і водночас складниками котрих не були ані вказівні, ані присвійні займенники – чи не єдині формальні засоби української мови, в семантичну структуру котрих входить сема означеності; а також основа котрих вводиться до даного тексту вперше (щоб уникнути випадків анафори).

Згідно із статистичними даними, поданими в праці Поезіо і Віейри¹ біля 70% усіх ІФ, котрі вперше згадуються в англійських текстах,

¹ Poesio Massimo and Renata Vieira: *A Corpus-based Investigation of Definite Description Use* [in]: "Computational Linguistics", v. 24, n.2, 1998, стор. 183-216.

супроводжуються означеними артиклями без впровадження за допомогою неозначеного артикля. Тим часом більшість робіт, присвячених проблематиці означеності, концентрується в основному на питаннях анафори. У даній статті ми намагатимемось показати, що значна частина неанафоричних випадків означеності є зумовлена семантичними функціями, під якими ми розуміємо однозначні відношення між поняттями.

Ми також робимо спробу вступної класифікації семантичних функцій і пропонуємо автоматизований спосіб екстракції з тлумачних словників лексичних одиниць, в семантичну структуру котрих входить сема означеності.

Проаналізований фрагмент повісті Ольги Кобилянської та його переклад на англійську мову² дозволив виділити окрім власних назв (*гори, названої Чабаниця*), котрі перекладаються на англійську без артиклів, ще декілька груп означених ІФ, котрі передаються англійською з артиклем або іншим формальним вказівником означеності³, напр. означеним займенником:

- 1) ІФ із означувальними підрядними реченнями *місцевість, де склалася подія, про яку розказується (the events chronicled here)*;
- 2) ІФ із вказівними та присвійними займенниками *тих гір, наше село, до тієї-то ріки*;

² Kobylanska Olha, *On Sunday Morning She Gathered Herbs*, translated by Mary Skrypnuk, CIUS, Toronto, 2001, 180 с.

³ Типологічні дослідження вживання показують, що різні мови мають різні правила вживання артиклів, що мало би ставити під сумнів існування категорії означеності як семантичної. Тим не менш, детальний розгляд (наприклад Stanisław Karolak, *Gramatyka kontrastywna przedimka (rodzajnika) francuskiego i angielskiego*, 2002, Częstochowa) свідчить, що основна розбіжність у вживанні артиклевих форм проявляється не на рівні конкретно-предметних фраз (за винятком різних, конвенційно зумовлених вживань в стосунку до власних назв та фразеологічних сполучень), а на рівні абстрактних назв, назв подій, станів та процесів. Ці твердження дозволяють зробити припущення про існування певного інваріантного значення означеного артикля. Репродуктивний характер аналізованого тексту повісті, де ІФ відносяться в основному до одиничних предметів, а не одноіменних їх класів, дозволяє нам з певним спрощенням використовувати їх англійські відповідники з означеними артиклями в якості сигналізаторів означеності.

3) ІФ, котрі не мають формального вказівника означеності: *ліси тих гір; в ногах одної гори; з самого верху Чабаниці; в долині (the forests in these mountains; the foot of one of these mountains; from the heights, at the bottom right at the river's edge).*

Вважаємо, що сема означеності може входити у значеннєву структуру слова, через що додаткове вираження означеності за допомогою інших мовних засобів стає надлишкове. Частина ІФ, які вживаються у перекладі із означеним артиклем, позначає унікальні предмети, як то *сонце, небо, зірки* (унікальна множина). Інша частина тих ІФ позначає загальні предмети, котрі є унікальними відносно інших означених. Семантичну вартість означеності таких лексем можна підтвердити також певними обмеженнями на вживання їх із формальними засобами означеності. Наприклад, перефразовуючи відомі речення Рассела і Доннеллана, отримуємо неграматичні (чи принаймні ненатуральні) конструкції: **A (present) king of France is wise. *A murderer of Smith is insane*. Подібним чином аномальними будуть фрази: **a present husband of Julie; *a most important lesson*.

Замінюючи *king of France* на *citizen*, а *murderer* на *mistress* чи *neighbour* отримаємо речення, котрі після перекладу на безартиклеву мову втрачають вказівку на означеність і повторний переклад на артиклеву мову вимагатиме додаткової інформації про цю граматичну категорію. Такої втрати інформації натомість не спостерігаємо у випадку оригінальних речень. Франція, як і будь-яка інша країна, може мати тільки одного короля в той самий час, а Сміт – тільки одного вбивцю, оскільки померти може тільки один раз.

Згадані однозначні семантичні відношення є істотні з точки зору категорії означеності. Їх можна віднести до поняття функції, котра у точних науках дефінюється як відношення, що кожному елементу області значень ставить

у відповідність лише один елемент області визначень, напр. функція *бути народженим кимось*⁴. Властивостями функцій є **комплетність** (кожному елементу області значень ставиться у відповідність певна вартість) та **однозначність** (кожному елементу області значень ставиться у відповідність один і лише один елемент області визначень)⁵.

Обернено однозначним є відношення, котре кожному елементу області визначень ставить у відповідність тільки один елемент області значень, наприклад *бути батьком*. Підмножиною функцій є обернені функції, тобто взаємно однозначні відношення, котрі є одночасно однозначні і обернено однозначні, наприклад *бути зарученим*, *бути столицею*, і т. д.

Функції можуть поєднуватися і результатом цих поєднань є т. зв. їх відносні добутки. Відносним добутком відношень R і S ($R \circ S$) є відношення між першим і другим предметом, умовою якого є існування третього предмета, до котрого перший перебуває й відношенні R , і котрий перебуває у відношенні S до другого⁶. Прикладом добутку відношень у природній мові є відношення *бути дідом*.

У працях, присвячених неанафоричному вживанню англійських артиклів (*first mention definites, discourse new definites* або *inferrables*), згадується про асоціативні групи, де слово-якір⁷ (*anchor* або *trigger*) викликає асоціацію слова-референта, напр. *We looked at a new house yesterday. The kitchen was tiny.*

Асоціативний зв'язок між поняттями є запорукою когерентності дискурсу – характеристики, яка виділяє текст (понадфразову єдність) з-поміж інших мовних одиниць, і без якої текст взагалі не може існувати. Тому логічно

⁴ Rasiowa, Helena: *Wstęp do matematyki współczesnej*, Warszawa, PWN, 1968, 302 с.

⁵ Cann, Ronnie: *Formal Semantics. An Introduction*, CUP, 1993, стор. 94.

⁶ Stanosz, Barbara: *Wprowadzenie do logiki formalnej. Podręcznik dla humanistów*, Warszawa, PWN, 2004, стор. 103.

⁷ Hawkins, John A.: *Definiteness and Indefiniteness. A Study in Reference and Grammaticality Prediction*, London, Groom Helm, 1978, 316 с.

припустити, що такі зв'язки є в мові досить численні, і що фактично за кожним поняттям є закріплений цілий ряд інших понять, які пов'язані з ним семантично. Оскільки поняття асоціації допускає досить різноманітні типи зв'язків, включаючи і ті, які на мовному рівні не можна передбачити через те, що вони виникають у конкретних контекстах, видається доцільним сконцентруватися саме на позаконтекстних зв'язках, закодованих на рівні мовної системи. Важливо також встановити вид семантичного пов'язання між поняттями для того, щоб можна було представити мовний матеріал більш упорядковано, у вигляді певної ієрархії.

Деякі типи семантичних функцій

Типи однозначних відношень, котрі представляємо нижче, є зумовлені онтологічно і відповідають загальноприйнятому світобаченню (це, звісно, обмежує нас до ситуацій з реально існуючого світу і виключає розгляд метафоричних і фантастичних представлень дійсності). Базовим онтологічним поділом за часовим критерієм є поділ на предмети і події⁸. В залежності від виду поняття області значень семантичні функції ділимо відповідно на предметні та ситуаційні. До області визначень можуть входити як предмети, так і події в обидвох випадках.

Предметні:

1) Відносно унікальні поняття типу *місяць, зорі, сонце* (до області значень у даному випадку відноситься більшість предметів, що оточують людину, і вона сама). Пробуючи звести до спільного знаменника „унікальні” одиничні предмети і предмети, унікальні тільки відносно інших, можемо

⁸ Див. Hawley, Katherine: *Temporal Parts*, “Stanford Encyclopedia of Philosophy”, 2004. <http://plato.stanford.edu/archives/spr2004/entries/temporal-parts/>. Casati, Roberto, Varzi, Achille: *Events*, “The Stanford Encyclopedia of Philosophy” (Fall 2002 Edition), Edward N. Zalta (ed.). <http://plato.stanford.edu/archives/fall2002/entries/events>.

взяти за основу критерій (відносної) часової співвіднесеності. При цьому діє загальне правило: предмет, котрий триває у часі довше, є однозначним відносно „коротшого” предмета, або вимагається принаймні їх рівночасність.

2) Поняття, пов’язані з виникненням x (напр. кровно-родинні відносини для людей) та кінцем існування x : *народити, мати, дід по батьковій лінії* (деякі мови мають окремі слова на позначення батьків батька і матері). Цими відношеннями пов’язані також лексеми *дім – архітектор* (приклад поєднання функцій), *фірма - банкрутство*.

3) Холоніми x (ціле, складовою частиною котрого є x).

4) В окремих випадках також мероніми x - унікальні частини x (відносно їх функції, розміщення або поєднання цих та інших ознак).

5) Гіпероніми x (родо-видові відношення – поняття роду відносно x).

Ситуаційні

Події теж можуть бути опорним пунктом функцій з „предметними учасниками”, котрі визначаються однозначно на відповідний проміжок часу. При цьому критерій часової співвіднесеності зберігається і є такий самий, як і у випадку відношень між предметами⁹. Якщо подія представлена як одинична, то її “оточення” як правило є означеним. Назва ситуації може бути виражена дієслівною формою в аористивній або перфективній функції, віддієслівним іменником або іменником із дієслівною семантикою.

До типових ситуаційних можна віднести наступні відношення: причина дії::наслідок дії, дія::виконавець, дія::інструмент, дія::результат (*писання – ручка, автор, лист; відкривання дверей – ключ, протяг, і т. д.*)

⁹ Визначальним є вибір опорного поняття. Одна і та ж сама людина може брати шлюб декілька разів, але в одній і тій же шлюбній церемонії може брати участь тільки один наречений (одна наречена).

Специфіка кожної окремої ситуації однак не дає можливості типізувати усіх її учасників (усі аргументи). Тому іноді доводиться задовільнятися самим переліком цих аргументів, напр.:

Шлюб – наречена, наречений, дружба, свекруха, розлучення.

Війна – переможець, оголошення війни, мирний пакт.

Одні аргументи ситуації можуть бути також однозначні відносно інших або аргументи можуть бути взаємнооднозначні, що дозволяє при потребі змінювати опорне поняття ситуації шляхом перегрупування її елементів, напр.:

Книжка – назва, автор, зміст, тематика.

Анкета – заповнення, бланк, набір питань.

Дисертація – захист, керівник, опонент, автореферат, диплом.

До ситуаційних віднесемо також відношення володіння (у належить до x) і репрезентації x (маємо на увазі тут найважливіший елемент якоїсь множини або його замітник). Ця функція входить у семантичну структуру наступних слів: *керує, править, керівник, президент, король, директор, ректор, голова, відповідальний (за), назва, символ, столиця* і т. д. Хоча практично відношення репрезентації реалізується найчастіше як стале у часі, теоретично однозначність відношення напр. між установою та її керівником є зумовлена і обмежена у часі саме ситуацією керування, а не предметами-репрезентантами.

Визначення понять, пов'язаних семантичними функціями, на мовному матеріалі великих обсягів

В якості корпусу для визначання різного роду семантичних зв'язків ми використовували тлумачний словник. Словникові статті розглядаються нами як оптимальний матеріал для екстракції значень абсолютної означеності з огляду на їх безконтекстний характер. Було проведено аналіз

на матеріалі частини (розділ із іменниками, біля 75 тис. ЛСВ) Словника Української Мови у формі Microsoft Access Database, котрий опрацьовується Українським Мовно-Інформаційним Фондом Національної Академії Наук України¹⁰. Формат бази даних СУМ дозволив виконувати запитання у текстах тлумачень мовою SQL (*Structured Query Language*). Наведемо нижче декілька спостережень, пов'язаних із результатами пошуку.

Репрезентація

Пошук слів, до тлумачень котрих входить слово *найважливіший* (у формі „найважливіш*“, що повинно було дати можливість знайти приклади в усіх можливих граматичних формах цього прикметника), дав 14 відповідей, серед них: *серце, основа 2, ключ, основа 3, домінанта, центр, серцевина*.

Набагато численішою була відповідь на більш загально сформульоване запитання „най*“ у тлумаченні слова – 678 іменників. Однак цей список мав занадто високий рівень інформаційного шуму, наприклад „*візник – найманий кінний екіпаж*“, де підкреслена частина слова є не префіксом найвищого ступеня прикметника, а частиною кореня дієприкметника.

Тлумачення, котрі починаються на „*головн**“, належать до 50 ЛСВ. Серед них було 9 випадків із омонімічним відповідником ключового слова – „*головний убір, головна хустка*“, котрі однак відрізняються своєю синтаксичною структурою, і при належному поділі тлумачень на синтаксичні складові (наприклад шляхом провізоричного усунення прикметників) такий вид тлумачень можна було би принаймні частково відокремити автоматично. До властивих результатів пошуку належить наступне типове тлумачення: „*Кореневище – головний корінь дерева, що є*

¹⁰ Авторка хотіла би подякувати керівництву УМІФ за надання можливості користування даними електронними ресурсами.

продовженням стовбура”. Після ключового слова йде гіперонім у називному відмінку, узгоджений з ним також по роду, потім холонім у родовому відмінку, а за ним т. зв. *differentia specifica*, диференційна ознака даного гіпоніма. Серед інших прикладів:

серце *головна, найважливіша частина, центр чого-небудь.*

собор *головна чи велика церква міста; головна церква в монастирі.*

столиця *головне місто, адміністративно-політичний центр держави, республіки і т. ін.*

Холоніми

Пошук за ознакою наявності слова *частина* у тлумаченні дав 1180 результатів, які встановлюють холоніми пояснюваних слів (іменник у родовому відмінку, який іде відразу після ключового слова).

Інхоативи і термінативи

Пошук за ознакою *початок* дав 54 результати, напр.:

світанок *початок, ранній період, зародження чого-небудь; зоря.*

старт *початок чого-небудь.*

Наявність неозначеного займенника *чого-небудь* у наведених прикладах свідчить про те, що дані поняття можуть мати досить універсальне застосування і вживатися по відношенню до багатьох різних понять.

До інформаційного шуму належать тлумачення, у котрих *початок* не характеризує базове слово дефініції, а є складовою частиною диференційної ознаки.

Пошук за ознакою *кінець* дав 70 результатів. До інформаційного шуму окрім випадків, аналогічних до згаданих вище, належать тлумачення із сталими словосполученнями, напр. *кінець кінцем*.

Дія::зряддя

Пошук понять, пов'язаних вказаним видом семантичного зв'язку, ми проводили у словнику за допомогою запитань за декількома ключовими

словами (фразами): *для* – 4840; *за допомогою якого* – 78; *за допомогою якої* – 28. Інформаційний шум у випадку ключових фраз є практично відсутній. Великий обсяг результатів у випадку *для* вимагає додаткового аналізу.

Верифікація даних

Припущення про те, що онтологічно однозначні системні (а тому і загальновідомі) семантичні співвідношення переважно представлені у мовленні у формі означених іменникових фраз, має під собою прагматичний ґрунт і базується на відомих максимах успішної комунікації Грайса. Засада кількості інформації говорить про те, що повідомлення повинно бути достатньо і не більш, ніж достатньо, інформативним¹¹. Інакше кажучи, надмір інформації шкодить комунікації так само, як і її нестача, і цим можна пояснити неанафоричне вживання означених ІФ у тексті – як відомо, будинки/квартири мають кухні (пор. приклад вище) згідно із самою дефініцією цих понять. Ця інформація є надлишковою і у дискурсі не вербалізується, або ця експлікація відбувається у специфічних комунікативних умовах, наприклад дидактичного характеру.

До таких типів текстів належать тлумачні словники. Порівняння вживання тих самих слів у словникових статтях і звичайних середньостатистичних текстах показало значну розбіжність у використанні артиклів. Текст перекладу згаданої повісті Кобилянської (69 520 слів) містить 2844 означених артиклів і 1189 неозначених (1081 *a* і 108 *an*). Подібний за об'ємом текст з Pocket Oxford Dictionary (від „a” до „bridle”, 65 992 слова) містить відповідно 1150 означених і 2111 неозначених (1797 *a* і 314 *an*) артиклів. Цю значну розбіжність пов'язуємо із типом комунікативного реєстру, а точніше – із ступенем узагальнення іменникової фрази

¹¹ Horn Laurence R.: *Presupposition and Implicature*, “The handbook of contemporary semantic theory”, Oxford, Blackwell reference, 1996, стор. 311.

(речення). Словники, тексти наукового і енциклопедичного характеру, прислів'я та афоризми належать до генеритивного реєстру мови¹², і в них ми, як правило, знаходимо відношення до типів або збірних понять, а не до конкретних предметів. Переважна більшість типів текстів має репродуктивно-інформативний характер (з можливими узагальненнями, напр. *All good things come to an end*, котрі однак мають відмінні синтаксичні структури і їх часто можна ідентифікувати формально). Типовим контекстом для неозначеної ІФ в них є декларація існування предмета, що часто є пов'язане із паралельним віднесенням даного предмета до певного класу понять. До таких класифікаційних і екзистенціальних контекстів та їх похідних належать фрази типу *існує x, x належить до класу P, сталося P, x має y*; а також опис способу робити щось, міра чогось (відстань, проміжок часу), напр.: ... *tomorrow, those of you who desire it, can bring an end to her*.

Вживання неозначених артиклів із словами, котрі теоретично є однозначно співвіднесені із іншими означеними словами, є можливе, але обмежене специфічними типами контекстів. Зокрема, в жодному з результатів пошуку в Інтернеті (за допомогою пошукової системи Google) конструкція “*a husband of* + власна назва” не зустрілася нам у ролі підмета. В основному це були частини іменного присудка, прикладки або неповні речення (підписи під портретами, на надгробках і т. д.), які фактично виконують роль іменного присудка (*тут спочиває x, тут намальований x, це x*). Статистичне співвідношення у вживанні означеного і неозначеного артикля з потенційно означеними ІФ виглядає наступним чином:

a husband of Julie (1) :: *the husband of Julie* (253) (*Tony Jones is a husband (of Julie) and a dad (of Tanner and Lily), and they're all now living in New Jersey*);

¹² Золотова, Г. А., Онипенко, Н. К., Сидорова, М. Ю.: *Коммуникативная грамматика русского языка*, Москва, РАН, 1998, 528 с.

для імені *Jane* це співвідношення становило 6::404; *Mary* 29::59000 (переважна кількість контекстів – біблійні).

Як видно, статистична розбіжність у вживанні подібної конструкції із означеним або неозначеним артиклем є досить велика, що свідчить про винятковий і досить специфічний характер останніх. Пояснення цього явища може бути наступне: однозначно залежна лексема включає в себе 1) значення відношення і 2) віднесення до предмету, який визначається через це відношення, а у згаданих “виняткових” контекстах реалізується саме перша з цих сем, тобто стверджується саме існування даного відношення в стосунку до *x* (тобто, що *x* є жонатий).

Так само спостерігається кількісна розбіжність у пропорціях вживання *a(n)::the* із деякими “нефункційними” словами, котрі однак дуже часто реалізуються як однозначні, що є зумовлено онтологічно. Однак у цьому випадку частка конструкцій з неозначеним артиклем є значно більша і вони мають вже не винятковий, а системний характер:

a/the daughter of Julie (13::436); *Jane* (281::780); *Mary* (911::32000)

a/the son of Julie (36::307); *Jane* (197::853); *Mary* (11800::195000)

На відміну від функційних, дані конструкції можна зустріти у позиції підмета, прямого та непрямого додатка.

Практичне застосування отриманих результатів

Інформація про отримані пари однозначно пов’язаних слів після відповідної семантичної редакції може бути придатна для побудови лексико-семантичної бази даних на зразок WordNet¹³ із більш розширеним набором типів семантичних зв’язків.

Її також можна закодувати як додаткові дані в електронних дво- та багатомовних словниках, котрі служать за лексичний матеріал для

¹³ WordNet: <http://wordnet.princeton.edu/>.

автоматичного (або автоматизованого) перекладу. Можна припустити, що слова із асоціативного ряду ключової лексеми вживатимуться у тексті із означеним артиклем, якщо знаходяться поблизу неї, за умови, що не підпадають під згадані вище обмеження.

Abstract

In this article we discuss non-anaphoric cases of definite noun phrases in English texts translated from Ukrainian. We claim that the use of first mention definites is connected with the notion of semantic function, by which we understand an unambiguous relation. We make an attempt to present a provisional classification of them and suggest some ways of computer-aided extraction of the ‘definiteness’ knowledge from explanatory dictionaries.

Анотація

У статті розглядаються неанафоричні означені іменникові фрази з англійських текстів, перекладених з української мови. Ми намагаємось показати, що значна частина неанафоричних випадків означеності є зумовлена семантичними функціями, під якими ми розуміємо однозначні відношення між поняттями. Ми також робимо спробу попередньої класифікації семантичних функцій і пропонуємо автоматизований спосіб екстракції з тлумачних словників лексичних одиниць, в семантичну структуру котрих входить сема означеності.

Natalia Kotsyba, Institute of Slavic Studies, Polish Academy of Sciences, tel. +48 503 488 804