

КАТЕГОРІЇ ЧАСУ ТА ПРОСТОРУ В ІНФОРМАЦІЙНИХ ПРИРОДНО-МОВНИХ ТЕХНОЛОГІЯХ

Анотація: В роботі аналізуються мовні засоби відтворення просторових та часових відношень для потреб автоматичного опрацювання мовного матеріалу. Акцентується увага на структурній організації цих засобів, які багато в чому подібні між собою.

Ключові слова: мовного матеріалу, природно-мовні технології, індивідуальною мовною системою, базу знань.

Вступ

Мовленнева діяльність актуалізується індивідуальною мовною системою (ІМС) конкретної людини – отже, всі сучасні інформаційні природно-мовні технології (ПМТ), що намагаються моделювати окремі прояви багатогранної мовленнєвої діяльності, повинні спиратися на модель ІМС, яка (як відомо за Л.В.Щербою [1]) визначається двома складовими: знаннями про мовну організацію (на свідомому чи підсвідомому рівнях) – лінгвістичним процесором (ЛП) та накопиченим досвідом про середовище свого існування – нашою базою знань (БЗ), де формується модель світу, в якому ми живемо. Без моделювання ІМС з урахуванням обох складових – ЛП та БЗ – всі наші спроби розбудови моделей, адекватних у певному сенсі особливостям мовленнєвої діяльності людини, приречені на невдачу.

В роботі автора [2] послідовно формувалось нове бачення організації мови, яке виводить нас на системний рівень мовної організації, що дозволяє вже тепер досить кваліфіковано ставити питання щодо створення програмного продукту ЛП. Наріжним каменем системної організації постає формально визначене поняття базової семантико-синтаксичної структури (БССС), витoki якої, багато в чому, визначаються структурно-функціональним рівнем нейроорганізації зорового тракту людини [3] та зумовлюють структурні особливості мови взагалі. Базова структура за визначенням – це вербалізована форма відтворення довільної ситуації реального або віртуального світу (дивись [2]), що існує у триєдності часу, простору та дії. Отже, координати часу та простору – це необхідні компоненти БССС, тому їх структурна організація постає однією з найважливіших проблем на шляху створення ЛП для автоматичного опрацювання природно-мовної інформації.

Платформа дослідження

Категорії часу та простору настільки важливі атрибути нашого буття, що для розуміння їх використання бажано звернутися до кваліфікованих джерел. Звернімося, наприклад, до фізичної енциклопедії, де

подаються найбільш узагальненні їх визначення та аналізуються їх особливості. “Простір та час у фізиці визначаються у загальному вигляді як фундаментальні структури координації матеріальних об’єктів та їх станів: система відношень, що визначає координацію співіснуючих об’єктів (відстань, орієнтація тощо) формують простір, а система відношень, що відтворює координацію змінюючих один одного станів чи явищ (послідовність, протяжність тощо) формують час. Простір та час постають організуючими структурами різних рівнів фізичного пізнання і відіграють важливу роль у міжрівневій взаємодії” [4].

Таке визначення сучасної платформи класичної фізики стосовно головних атрибутів нашого існування постає конструктивним підґрунтям інтуїтивних поглядів гуманітарної спільноти (психологів, філософів, лінгвістів), що наше довкілля існує в триєдності часу, простору та дії. Коли ж головна функція мови визначається як адекватне відтворення довкілля, то на перший план, природно, виступають питання аналізу мовних засобів для відтворення просторових та часових відношень і фіксації послідовності їх станів або змін. Отже, нас цікавитимуть, в першу чергу, мовні засоби та їх організація для відтворення різноманітних просторово-часових (П/Ч) відношень з метою подальшого їх використання для автоматичного опрацювання інформації в сучасних інформаційних природно-мовних технологіях.

Існує досить широкий пласт літератури подібного кшталту, де з різних позицій (філософських, лінгвістичних, кібернетичних) розглядаються питання відтворення просторових та часових відношень у мовному матеріалі. Увага до цієї проблеми набула особливої гостроти з початком формування інформаційних технологій, орієнтованих на опрацювання природно-мовної (ПМ) інформації. Звісно, не можна не згадати в цьому контексті роботи, що пов’язані зі спробами моделювання мовленнєвої діяльності. Одна з перших спроб була здійснена під керівництвом Д.О. Поспелова, що презентована роботою [5]. Слід підкреслити намагання авторів проаналізувати та подати узагальнену модель часу та простору як визначальних атрибутів нашого існування з метою подальшого їх використання в програмних продуктах. В роботі підкреслені такі властивості цих категорій як: ненаправленість простору, однонаправленість часу, безперервність часу та простору, нескінченність часу та простору, гомогенність простору тощо. Вказано також, що залежно від потреб розрізняють системи відтворення відношень: абсолютні, відносні, розмиті. Проте невідомо, наскільки ефективно авторам вдалося перекласти на плечі комп’ютера цю віртуальну модель.

Інша спроба використання просторових відношень досить вузької сфери презентована була роботою автора [2], де на програмному рівні були реалізовані процедури синтезу та аналізу просторових відношень. На особливу увагу заслуговує запропонована графічна інтерпретація П-відношень (образотворчі засоби семантики – за Ю.Д. Апресяном [6]), що дозволило побудувати систему, яка працює зі “сміслом” окремих просторових термінів і за текстом синтезує П-відношення (режим аналізу пові-

домлення) або, навпаки, за графічним поданням просторових відношень визначає його мовну інтерпретацію (режим синтезу повідомлення).

Звісно, є ще багато інших досліджень часу та простору як важливих категорій нашого існування, що зачіпають інші аспекти їх актуалізації на мовному рівні, проте нас цікавитимуть більше структурні особливості відтворення часу та простору у мовному матеріалі. Фахово проблемами часу та простору (абсолютного часу та абсолютного простору) цікавляться фізики та математики, нас же більше цікавитимуть різні схеми відтворення П/Ч-відношень на мовному рівні. У загальному випадку, коли наш життєвий простір заповнений множиною об'єктів/суб'єктів, різних станів та ситуацій, існує і нескінченна кількість просторово-часових відношень; проте, важливо підкреслити, що, практично, кожна мова має засоби відтворення найтонших і найрізноманітніших П/Ч-відношень, і наша мета – простежити хоча б основні схеми та структури формування таких відношень.

Одиниці виміру та їх особливості

Якщо головна функція мови – адекватне відтворення навколишнього середовища, що існує в триєдності часу, простору та дії, то, зазвичай, продуктивним постає аналіз співвідношення “Дійсність-Текст”; наше завдання – проаналізувати особливості організації мовних засобів для адекватного відтворення просторово-часових відношень. Звісно, треба враховувати, що категорії існування світу – час та простір – це категорії безперервні, тоді як мовні засоби мають суто дискретну природу. Отже, можемо вважати, що наш життєвий простір є тривимірним (просторові координати X, Y, Z), а час вважатимемо четвертою координатою T . Мабуть, промовчимо стосовно початку відліку часу та простору, відсилаючи їх до моменту створення всесвіту.

Тут доречно ще раз звернутися за допомогою до фахівців – фізиків [4]: “Абсолютний, істинний, математичний час сам по собі і за своєю суттю, без будь-якого відношення до будь-чого зовнішнього, протікає рівномірно і інакше називається *тривалістю*. Абсолютний простір за самою своєю суттю безвідносно до будь-чого зовнішнього залишається завжди однако-вим та нерухомим”. Тож, практично, при аналізі часових та просторових відношень йдеться про фіксацію або поточних П/Ч координат існування окремої складової довкілля, або відстані між об'єктами/суб'єктами, подіями, ситуаціями тощо. Всі одиниці виміру часу та простору умовно можемо віднести до одиниць кількісних та якісних; сфера використання перших – це технічні напрями застосування, математичні та теоретичні розробки тощо, тоді як інші – частіше використовуються для відтворення просторово-часових відношень на побутовому рівні. Кількісні одиниці визначаються чіткою логікою організації та певними сферами використання, тоді як якісні – частіше відносяться до нечітких множин, де не зовсім однозначно окреслені межі їх застосування, а семантичне навантаження визначається конкретною ситуацією. Приклади використання одиниць виміру різних рівнів наведені у Табл.1. Слід тільки зазначити,

що коли наш простір визначається як тривимірний (3D – за сучасними уподобаннями), то і одиниці виміру також повинні відповідати його особливостям, тобто, бути одновимірними, двовимірними та тривимірними.

Важливо тут підкреслити суттєву різницю між кількісними та якісними одиницями виміру: кількісні одиниці, вкладаючись одна в одну, дозволяють відтворити весь континуат (безперервність та нескінченність) категорій часу та простору, тоді як якісні одиниці фіксують лише окремі їх проміжки, що пов’язані з певними якісними оцінками, ситуаціями, подіями, важливими для існування людини та її життєдіяльності.

Система СІ подає нам вихідні одиниці часу “секунда” та простору “метр” і вказує алгоритм формування всіх похідних одиниць.

Таблиця 1

Одиниці виміру часу та простору

Кількісні одиниці		Якісні одиниці	
Час	секунда, хвилина,	день, ніч, ранок, вечір, північ	час доби
	година, доба, тиждень, місяць,	сніданок, обід, ланч, вечеря	час прийому їжі
	декада, квартал, рік, десятиріччя	понеділок, вівторок, середа, четвер, п’ятниця, субота, неділя	послідовний порядок днів у тижня
Простір	метр, кілометр, сантиметр, міліметр	лікоть, сажень, крок	міри довжин

Система СІ подає нам вихідні одиниці часу “секунда” та простору “метр” і вказує алгоритм формування всіх похідних одиниць. Якщо нас цікавлять, в першу чергу, мовні засоби відтворення просторово-часових відношень, то не завадило б враховувати можливість ідентифікації всіх одиниць, які визначаються стандартною системою префіксів (див. Табл.2), які додаються до базових, формуючи весь спектр похідних одиниць, що можуть використовуватися в мовній інформації [7].

Слід підкреслити ще одну особливість використання одиниць виміру, що пов’язана саме з мовними особливостями їх використання. По-перше, ідентифікація окремих проміжків часу та простору у загальному випадку відбувається шляхом використання певної множини різномасштабних одиниць (“12 годин”, “15 діб”, “5 км” тощо); по-друге, система відтворення П/Ч відношень формується таким чином, щоб послідовність одиниць була здатною з певною точністю презентувати довільний проміжок (“12 год. 18 хв.”, “5 км. 200м.” тощо); по-третє, важливим елементом використання кількісних та якісних одиниць виміру є можливість урахування порядку використання однотипних одиниць виміру при іден-

Префікси системи СІ для формування похідних одиниць

Назва префікса	Позначення префікса		Множник	Назва множника
	Українське	Міжна родне		
екса	Е	E	10^{18}	Квінтиліон
пета	П	P	10^{15}	Квадриліон
тетра	Т	T	$100000000000=10^{12}$	Трильйон
гіга	Г	G	$100000000=10^9$	Мільярд
мега	М	M	$100000=10^6$	Мільйон
кіло	К	K	$1000=10^3$	Тисяча
гекто	г	h	$100=10^2$	Сто
дека	да	da	$10=10^1$	Десять
деци	Д	d	$0.1=10^{-1}$	одна десята
санти	с	c	$0.01=10^{-2}$	одна сота
мілі	м	m	$0.001=10^{-3}$	одна тисячна
мікро	мк	mc	$0.000001=10^{-6}$	одна мільйонна
нано	н	n	$0.000000001=10^{-9}$	одна мільярдна
піко	п	p	$0.000000000001=10^{-12}$	одна трильйонна
фемто	ф	f	10^{-15}	одна квадроліонна
отто	а	a	10^{-18}	одна квінтиліонна

тифікації проміжків часу та простору (що притаманно, до речі, багатьом системам відліку).

Ця властивість дозволяє додатково визначати не лише відстань, а й враховувати порядок використання окремих одиниць на певному відрізку, коли позиція одиниці виміру пов'язана з певними особливостями відтворюваної ситуації (**“на 20-му кілометрі траси...”**, **“о 18-й годині...”**, **“на 5-й хвилині...”** тощо). Важливо тут підкреслити семантичну розбіжність у використанні кількісних та порядкових оцінок відліку. Кількісний характер використання одиниць виміру називає лише множину одиниць певного кшталту, враховуючи їх кінцеві оцінки. Так термін **“5 кілометрів”** означає всю відстань від умовного початку відліку до кінцевої позначки кількісної оцінки. Термін же **“5-й кілометр”** визначає відстань лише однієї одиниці виміру, починаючи від кінця четвертого до початку шостого. Слід зауважити, що ця дихотомія “кількість/порядок” використовується досить ефективно в мовній практиці для ідентифікації не лише просторово-часових залежностей, а й окремих складових множини об'єктів/суб'єктів із загальної їх кількості, акцентуючи увагу читача саме на конкретному об'єкті серед йому подібних. Ввівши систему одиниць часу та простору, можемо вже перейти до аналізу різних схем формування П/Ч-відношень. Але спочатку розгля-

немо варіанти реалізації П/Ч-відношень, що не використовують одиниць виміру.

Фіксація поточних координат часу та простору

Перший варіант формування просторово-часових відношень на шляху філогенезу мови був пов’язаний, мабуть, зі спробою актуалізації цих відношень без урахування будь-яких одиниць виміру (не знала ще первісна людина як вимірювати час та простір!). Аналогічна ситуація спостерігається в онтогенезі мови на етапах опанування дитиною категорій часу та простору (дивись Гвоздева А.Н. [8]). Час та простір визначаються прислівниками *“тут”, “там”, “ось”, “тепер”*, що визначають, власне, поточні координати мовця (див. Табл.3). Термін *“ось”* навіть важко з впевненістю віднести до часової чи просторової оцінки. Трохи пізніше вже фіксуються П/Ч-відношення в більш конкретній формі: *“на цьому місці”, “в певний момент часу”* тощо. Наступний етап пов’язаний вже з формуванням та використанням одиниць виміру, мабуть, спочатку якісних і лише значно пізніше – кількісних, які вже дозволяють виокремлювати певні відстані у часі та просторі. Для аналізу структури категорій з використанням вже одиниць виміру введемо певну нотацію.

Таблиця 3

Поточні координати часу та простору

	Минулий час	Теперішній час	Майбутній час
Час	попередньої миті	тепер	наступної миті
	вчора	цієї миті	Завтра
	минулого тижня	сьогодні	наступного тижня
	минулого року	цього тижня	наступного місяця
	минулої декади	цього року	наступного кварталу
	минулого століття	цього ранку	наступного півріччя
Простір	<i>Простір абсолютний (одновимірний)</i>		
		на цьому місці	
		на тому місці	
		безвідносно subj/obj	

Нотація формування просторово-часових відношень

Для аналізу структури П/Ч-відношень введемо відповідні позначення: часову відстань (відрізок часової осі) позначатимемо через *T*, просторову відстань подамо через *S*, тоді, відповідно, кількісні одиниці виміру часу приймаємо за *t_i*, а кількісні одиниці виміру простору – як *s_i*, кількість одиниць виміру певного кшталту позначимо – *n*.

Визначення відрізків часу та простору

Узагальнена процедура визначення відстані у часі та просторі добре відома і виглядатиме наступним чином:

$$S = n_1 * s_1 + n_2 * s_2 + ... + n_k * s_k = \sum n_i * s_i$$

“8 метрів 5 сантиметрів 18 міліметрів”

$$T = n_1 * t_1 + n_2 * t_2 + \dots + n_k * t_k = \sum n_i * t_i$$

“2 години 45 хвилин 30 секунд”.

До цього часу ми працювали з кількісними вимірами простору та часу, що визначають певні відрізки цих категорій. Проте там, де існує “Кількість” завжди присутня й інша категорія “Порядок”. Особливість використання цієї категорії – у виокремленні із певної множини однакових одиниць конкретної міри (просторової чи часової) з урахуванням її протяжності від початку до закінчення або до поточного моменту (моменту мовлення). Приклади можуть бути наведені досить широко:

*“останній тиждень місяця”, “сімнадцятий кілометр”,
“1997 рік”, “двадцять перше сторіччя”.*

Кількісна оцінка завжди фіксує наявність кінцевої межі виміру, звернення ж до порядкової форми передбачає актуалізацію всього визначеного проміжку.

Відносні схеми формування просторово-часових відношень

До цього часу ми аналізували паралельно мовні засоби відтворення часових та просторових відношень як одновимірних характеристик, не акцентуючи увагу на системах відліку цих відстаней. Незважаючи на попереднє зауваження стосовно початку відліку світових координат часу та простору, які сприймаємо як невизначені, відлік часу та простору ведеться відносно визначальних координат нашого існування: просторові координати – це наш дім-Земля, а часові – це простір існування людства, зафіксований в календарних системах. Тут не зачіпаємо такі категорії як “абсолютний час” та “абсолютний простір”, які відіграють суттєву роль у формуванні теорії становлення світу та теоретичних розробок, а на побутовому рівні майже не використовуються. Значно важливішими у житті постають відносні системи відліку часу та простору, коли за початок координат приймається окремий об’єкт/суб’єкт (для формування П-відношень) або конкретна подія/ситуація (для ідентифікації Ч-відношень).

У роботі автора [3] свого часу досліджувалися особливості формування Ч-відношень за відносною системою. Було показано, що коли одна подія (ситуація) зміщується відносно іншої, то мовна практика опрацювала відповідні засоби для досить тонкої ідентифікації подібних відношень. Приблизний підрахунок подібних Ч-відношень складав близько шестисот типів. Аналогічним чином, враховуючи структурний рівень формування обставин простору, можемо проаналізувати можливість формування П-відношень за відносною системою, коли відлік відстані ведеться відносно конкретного об’єкта/суб’єкта, просторові координати якого нам добре відомі або з контексту, або з нашого досвіду. Можливі варіанти формування П/Ч-відношень представлені рисунками 1 та 2; при їх зіставленні чітко простежується структурна одноманітність організації П/Ч-відношень

відносно початку відліку. До того ж слід нагадати тезу, що кожна мова має відповідні засоби для відтворення (фіксації) найрізноманітніших П/Ч-відношень нашого існування.


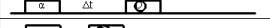
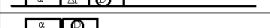

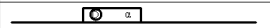

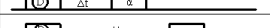
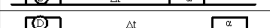

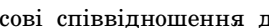
	задовго до ...
	незадовго до ...
	за Δt до ...
	напередодні ...; за день до ...
	з настанням ...; в час ...; тільки що ...; як тільки ...; ледве ...
	протягом ...; в той час ...; коли ...; поки ...
	одразу після ...; незабаром після ...; слідом за ...
	після ...
	через Δt після ...; трохи пізніше після ...
	значно пізніше після ...

Рис. 1 – Часові співвідношення двох подій: α – відтворювана подія середовища, D – довільна подія (момент віднесення), Δt – відстань між α і D , що визначається на континуаті часової осі, t_b – часова вісь.

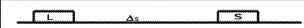
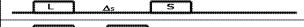

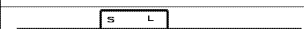
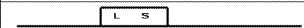

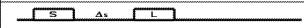
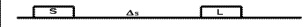
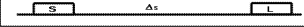
	далеко (попереду) від
	недалеко (попереду) від
	за декілька метрів (попереду) від
	пору (прямо попереду) з порід з
	в тому ж місці де ...; з однаковою позиці від розташований в тому ж місці де ...
	займають однакову площу з ...; рівне співвідношення між еквівалентне місцезнаходження
	впритул до (за) ...; відразу за ...; позаду за ...
	проходять повз ...; недалеко від/позаду від ...; недалеко обгиваючи ...
	за кілька метрів позаду від ...; трошки подаль від ...; заходять недалеко/подаль позаду ...; значно далі позаду ...; проходять далеко повз обминаючи ... на значну відстань

Рис. 2 – Просторові співвідношення двох об’єктів: L – об’єкт, що співвідноситься, S – довільний об’єкт, відносно якого здійснюється відлік (точка віднесення), Δs – відстань між L і S , що визначається на континуаті просторової осі, s_b – просторова вісь.

Визначення відстані між двома об’єктами, подіями

Вище розглядалися системи відліку П/Ч-відношень відносно однієї події/ситуації чи конкретного об’єкта/суб’єкта. Проте у одновимірному варіанті можлива також фіксація відстані між двома об’єктами, подіями тощо. У такому випадку, у повідомленні присутня актуалізація обох елементів, стосовно яких формується просторове чи часове відношення. Наприклад, для фіксації П-відношень матимемо: “Між човном та островам відстань вже складала не більше сотні метрів”; а для часових відношень, відповідно – “Між заручинами і весіллям промайнуло вже кілька місяців”.

Висновки

В роботі ретельно проаналізовано співвідношення “Дійсність-Текст” в плані аналізу мовних засобів відтворення найважливіших категорій існування нашого світу – часу та простору. Дослідження присвячене аналізу структурного рівня організації мовних засобів, де головними складовими виступають: одиниці виміру (кількісні та порядкові), їх параметри, а також використовуються службові елементи мови для відтворення П/Ч-відношень окремих складових нашого оточення. В роботі представлені кількісні та якісні одиниці виміру часу та простору, проаналізовані механізмами фіксації поточних координат моменту мовлення, враховані особливості використання абсолютних та відносних систем відліку. Всі ці питання проаналізовані в площинах 1D, що зумовлює відповідні формати подання П/Ч-відношень.

Література

1. Щерба Л.В. О тройком аспекте языковых явлений и эксперименте в языкознании : избр. труды. /Щерба Л.В. Языковая система и речевая деятельность. – Л.: Наука, 1974. – 24-39 с.
2. Кисленко Ю.І. Архітектура мови (Лінгвістичне забезпечення інтелектуальних інтегрованих систем). – учбовий посібник. – К.: Випол, 1998. – 344 с.
3. Кисленко Ю.И. От мысли к знанию (нейрофизиологические основания): Монография. – К.: “Український літопис”, 2008.-101 с.
4. Физическая энциклопедия (под ред. А.М. Прохорова) Т.4. – М.:1994. – 704 с.
5. Кандрашина Е.Ю., Литвинцева Л.В., Поспелов Д.А. Представление знаний о времени и пространстве в интеллектуальных системах. – М.: Наука, 1989. – 328 с.
6. Апресян Ю.Д. Лексическая семантика. – М.: Наука, 1974. – 307 с.
7. Енохович А.С. Справочник по физике и технике. Учебное пособие для учащихся. – М.: Просвещение, 1989. – 224 с.
8. Гвоздев А.Н. Формирование у ребенка грамматического строя русского языка. – М.: АПН, 1949. – 268 с.

Отримано 11.03.2011 р.