

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний авіаційний університет

А.І. Вавіленкова

**АНАЛІЗ І СИНТЕЗ ЛОГІКО-ЛІНГВІСТИЧНИХ
МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ**

КИЇВ 2017

В.А. Широков – академік НАН України (Український мовно-інформаційний фонд НАН України);

Д.В. Ланде – д-р техн. наук, професор (Інститут проблем реєстрації інформації НАН України).

Рекомендовано до видання вченою радою Національного авіаційного університету (протокол №5 від 31.05.2017 р.).

Вавіленкова А.І.

В 12 Аналіз і синтез логіко-лінгвістичних моделей речень природної мови: монографія. – К.:ТОВ «СІК ГРУП УКРАЇНА», 2017. – 152 с.

ISBN 978-617-7457-33-5

Надано класифікацію окремих форм логіко-лінгвістичних моделей, що адаптовані до певного типу речень природної мови. Запропоновано алгоритм побудови логіко-лінгвістичної моделі електронного текстового документу. Сформульовано основні принципи та правила синтезу, на яких базується побудова формальної логіко-лінгвістичної моделі електронного текстового документу, описані абстрактні моделі логічних зв'язків між частинами текстових документів. Сформульовано критерії аналізу та умови тотожності логіко-лінгвістичних моделей простих речень природної мови. Описано структуру інформаційної технології порівняльного аналізу електронних текстових документів за змістом.

Усі наведені результати належать автору особисто.

Для студентів та науковців, які спеціалізуються в галузі інформаційного пошуку, комп'ютерної лінгвістики та інших сферах інформаційних технологій, де необхідною є автоматична обробка текстової інформації.

ЗМІСТ

ВСТУП	5
I. ОКРЕМІФОРМИ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ	9
1.1 Логіко-лінгвістичні моделі речень природної мови та особливості їх побудови	9
1.2 Класифікація окремих форм логіко-лінгвістичних моделей ...	15
1.3 Логіко-лінгвістична модель текстового документу	43
II. СИНТЕЗ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ	55
2.1 Основні принципи синтезу логіко-лінгвістичних моделей речень природної мови.....	56
2.2 Правила синтезу логіко-лінгвістичних моделей текстових фрагментів	80
2.3 Абстрактні моделі логічних зв'язків між частинами текстових документів	86
III. АНАЛІЗ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ	97
3.1 Критерії аналізу логіко-лінгвістичних моделей речень природної мови	97
3.2 Умови тотожності логіко-лінгвістичних моделей простих речень природної мови.....	103
3.3 Алгоритм аналізу логіко-лінгвістичних моделей текстових документів	111
IV. ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕЛЕКТРОННИХ ТЕКСТОВИХ ДОКУМЕНТІВ НА ОСНОВІ ВИКОРИСТАННЯ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ	121

4.1 Основні способи виявлення рерайту за допомогою логіко-лінгвістичного моделювання.....	121
4.2 Метод порівняльного аналізу електронних текстових документів за змістом	126
4.3 Інформаційна технологія порівняльного аналізу електронних текстових документів за змістом	133
СПИСОК ЛІТЕРАТУРНИХ ДЖЕРЕЛ	143

ВСТУП

Необхідність створення унікального контенту, боротьба за авторські права, поява величезної кількості компаній та сервісів із здійснення ререйтингу, а також безупинний плагіат наукових робіт спричиняють необхідність створення нових якісних механізмів для виявлення копій електронних текстових документів на парадигматичному рівні [1]. Задачі змістовної автоматичної обробки текстової інформації пов'язані з проблемами реєстрації інформації, захистом інтелектуальної власності, інноваційною діяльністю, відповідністю документації юридичних та інформаційно-аналітичних організацій певним встановленим нормам [2].

Основною проблемою на шляху змістовної автоматичної обробки текстової інформації є вирішення протиріччя між існуванням різноманітних підходів і алгоритмів та відсутністю якісно нових програмних продуктів, що забезпечили б автоматичну екстракцію знань із електронних текстових документів. Вирішення цієї проблеми можливе завдяки створенню такого формального апарату, який дозволив би аналізувати текстову інформацію за єдиним принципом, систематизувати процес змістовного пошуку та уникнути неоднозначності при аналізі текстових документів. Описана проблема підпадає під пріоритетний тематичний напрям фундаментальних досліджень і науково-технічних розробок на період до 2020 року «Технології та засоби розробки програмних продуктів і систем» постанови Кабінету Міністрів України №556 від 23.08.2016 р. [3]. У монографії пропонується вирішувати цю проблему шляхом логіко-лінгвістичного моделювання.

Для можливості автоматичної обробки текстової інформації необхідна розробка:

- моделей представлення знань, які враховують змістовні зв'язки між структурними елементами тексту;
- алгоритмів аналізу поверхневих структур та конструювання відносно простої та жорсткої семантичної конфігурації;
- моделей та алгоритмів змістовного аналізу тотожних природно мовних конструкцій;
- алгоритмів відновлення окремих об'єктів тексту та їх взаємозв'язків;

– методів автоматичного синтаксичного, семантичного та лінгвістичного аналізу;

– алгоритмів та методів змістовного автоматичного порівняння електронних текстових документів.

Проблематикою змістовного аналізу текстів займаються у наш час в першу чергу комп'ютерні лінгвісти. Так, у роботах Широкова В. А. "Лінгвістичні та технологічні основи тлумачної лексикографії" [4], "Корпусна лінгвістика" [5] та "System semantics of explanatory dictionaries" [6] описується концепція станів мовних одиниць, що лежить в основі створення та функціонування лексикографічних систем. Визначенню лексичних компонент та спробам моделювання змісту речень природної мови присвячено роботу Вівьєна Еванса та Мелані Грін «Когнітивна лінгвістика» [7]. Формальні моделі представлення знань описані у роботах Поспелова Д. А. [8], Осуга С. [9], Вагіна В.М. [10].

Дослідженню проблем та методів виявлення контексту текстових документів присвячено статті Чанга К. та Кларка С. "Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method" [11], а також Занга І. "Discriminative syntax-based word ordering for text generation" [12]. Хорошу теоретичну базу для досліджень структури тексту та логічних зв'язків у ньому дає Р. І. Гальперін у своїй монографії «Текст як об'єкт лінгвістичного дослідження» [13]. Проблемам аналізу тексту та виявленню у ньому змістовних зв'язків присвячено роботи Лайонза Дж. "Лингвистическая семантика" [14], Нікітіна М. В. "Курс лингвистической семантики" [15], Апресьяна Ю. Д. "Лексическая семантика" [16], Чернявської З. Е. "Лингвистика текста: поликодовость, интертекстуальность, интердискурсивность" [17], Рубашкіна В.Л. [18], Гладкого А.В. [19]. Проблемам обробки природної мови присвячено курс лекцій американських комп'ютерних лінгвістів Дурафски Д. та Манінга С. "Natural Language Processing" [20]. Провідними вченими, які мають справу з обробкою текстової інформації у сфері інформаційних технологій, є: Ланде Д. В., роботи якого "Интернетика: навигация в сложных сетях: модели и алгоритмы" [21], "Аналіз інформаційних потоків у глобальних комп'ютерних мережах" [22] присвячені моделям представлення знань, методам обробки природно мовної інформації та інформаційним

технологіям потокової обробки текстової інформації; Глибовець М. М. та Глибовець А. М., які у своїх роботах "Системи штучного інтелекту" [23], "Один підхід до побудови інтелектуальної пошукової системи" [24], "Інтелектуальні мережі" [25] піднімають питання розуміння тексту, інтерпретуючи процес сприйняття природної мови через рівні, кожному з яких ставиться у відповідність конкретний вид абстрактної моделі.

Проте задача вилучення знань з текстової інформації при її автоматичній обробці досі залишається не розв'язаною.

Комп'ютерна програма АBBYU Compreno [26] здійснює аналіз текстової інформації, використовує дерева розбору, отримані семантико-синтаксичним парсером, онтології, правила вилучення інформації та правила ідентифікації. Також проаналізовано роботу існуючих сервісів перевірки текстових документів на плагіат [27-30]. Як виявилось, основною проблемою функціонування існуючих сервісів визначення відсотку унікальності електронних текстових документів є відсутність лінгвістичної компоненти порівняння, тобто механізмів, що здійснювали б зіставлення текстів не лише за кількісними показниками, але й за змістом.

У монографії ж запропоновано методику змістовного аналізу електронних текстових документів, що на базі аналізу та синтезу логіко-лінгвістичних моделей речень природної мови дає можливість структурувати текстову інформацію, починаючи з найнижчого рівня побудови логічних зв'язків (словосполучень) і закінчуючи текстом в цілому.

Метою даної монографії є викладення результатів, отриманих автором у процесі обробки та дослідження логічних зв'язків в електронних текстових документах для їх подальшого порівняння.

Основні результати досліджень висвітлено у чотирьох розділах.

У першому розділі запропоновано уніфіковану форму логіко-лінгвістичних моделей речень природної мови, описано принципи їх побудови. Також у першому розділі надано класифікацію окремих форм логіко-лінгвістичних моделей, що адаптовані до певного типу речень природної мови. Проаналізовано процес побудови окремих форм логіко-лінгвістичних моделей одних і тих самих речень, написаних на різних флективних мовах

(українській, російській та англійській). Описано алгоритм побудови логіко-лінгвістичної моделі електронного текстового документу.

У другому розділі дано визначення синтезу логіко-лінгвістичних моделей речень природної мови, описано основні принципи та правила синтезу, на яких базується побудова формальної логіко-лінгвістичної моделі електронного текстового документу. Також у другому розділі описані абстрактні моделі логічних зв'язків між частинами текстових документів, що відповідають певному типу тематичного зв'язку, та здійснено їх геометричну інтерпретацію.

Третій розділ присвячено аналізу логіко-лінгвістичних моделей речень природної мови. Тут сформульовано критерії їх аналізу, приведено умови тотожності логіко-лінгвістичних моделей простих речень природної мови та запропоновано алгоритм аналізу логіко-лінгвістичної моделі електронного текстового документу, на основі якого з формальної моделі можна відтворити текст природною мовою.

У четвертому розділі описано метод порівняльного аналізу електронних текстових документів, що базується на методі порівняльного аналізу речень природної мови довільної складності, що, у свою чергу, базується на способах виявлення рерайту за допомогою логіко-лінгвістичного моделювання. Описано структуру інформаційної технології порівняльного аналізу електронних текстових документів.

Монографія буде корисною для студентів та науковців, які спеціалізуються в галузі інформаційного пошуку, комп'ютерної лінгвістики та інших сферах інформаційних технологій, де необхідною є автоматична обробка текстової інформації.

Виходом цієї монографії завдячую моїй сім'ї, що безупинно підтримувала мене у непередбачених обставинах розвитку моєї наукової діяльності.

I. ОКРЕМІ ФОРМИ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ

*"Якщо ти знаєш, що ось це рука,
то це потягне за собою і все інше."*

Людвіг Вітгенштейн

Речення є основною синтаксичною одиницею, за допомогою якої формується та виражається думка. Речення складається із граматично оформлених слів і словосполучень, які становлять одне ціле висловлювання, що характеризується смисловою завершеністю і цілісністю сприйняття [31].

Просте речення має одну граматичну основу, у складному – виділяються дві і більше частин, кожна з яких має свою граматичну основу. Частини складного речення утворюють єдине ціле: вони пов'язуються за змістом. Тут зміст – це цілісне розуміння певного висловлювання, що не зводиться до значення його складових частин та елементів, а саме визначає ці значення [32].

Члени простого речення поділяються на головні (підмет і присудок) та другорядні (додаток, означення та обставина) [31]. Відношення між підметом і присудком, а також їх спільне відношення до того, що вони виражають в дійсності, формують суть (зміст) речення, – його предикативність [33].

1.1 Логіко-лінгвістичні моделі речень природної мови та особливості їх побудови

На відміну від стандартної методики поділу речень природної мови на групи слів, що формуються згідно з їх місцем розташування у реченні, логіко-лінгвістичне моделювання базується на ідеї знаходження груп слів, що пов'язані між собою за змістом, тобто словосполучень.

Керуючись гіпотезою "правило на правило", згідно з якою між граматичною структурою речення та логічною формою існує повна відповідність [34], можливістю застосування теореми про повноту [35], а також враховуючи те, що кожне речення природної мови характеризується граматичною організацією:

будовою та граматичним вираженням (оформленням) членів речення [36], маємо таку градацію:

- підмет – суб'єкт x ;
- присудок – відношення p ;
- додаток – об'єкт y або предмет z відношення;
- означення – характеристика суб'єкта g , об'єкта q або предмета відношення r ;
- обставина – характеристика відношення h .

Група слів, пов'язаних між собою логічними зв'язками, буде позначатися $sp_j, j = \overline{1, m}$, де m – кількість словосполучень у реченні. За правилами української мови словосполучення можуть утворювати між собою наступні члени речення (за їх наявності) [37, 38]:

- "означення – підмет" – $sp_j = g \cup x$;
- "присудок – додаток" – $sp_j = p \cup y$;
- "означення – додаток" – $sp_j = q \cup y$;
- "додаток – додаток" – $sp_j = y \cup z$;
- "означення – додаток" – $sp_j = r \cup z$;
- "обставина – присудок" – $sp_j = h \cup p$.

Наприклад, нехай потрібно виявити словосполучення у простому реченні: *"Нові дисципліни регулярно включають до навчальних планів"*. Спочатку відбувається ідентифікація ролей для кожного із слів речення:

- підмет – суб'єкт x – \emptyset ;
- присудок – відношення p – *включають*;
- додаток – об'єкт y – *дисципліни*;
- додаток – предмет z відношення – *планів*;
- означення – характеристика суб'єкта g – \emptyset ;
- означення – характеристика об'єкта q – *нові*;
- означення – характеристика предмета відношення r – *навчальних*;
- обставина – характеристика відношення h – *регулярно*.

Таким чином, у заданому реченні наявні такі словосполучення:

- "присудок – додаток" – $sp_j = p \cup y$ – включають дисципліни;
- "означення – додаток" – $sp_j = q \cup y$ – нові дисципліни;
- "додаток – додаток" – $sp_j = y \cup z$ – дисципліни планів;
- "означення – додаток" – $sp_j = r \cup z$ – навчальних планів;
- "обставина – присудок" – $sp_j = h \cup p$ – включають регулярно.

На основі структуризації словосполучень, які входять до речення як його складові частини і є будівельним матеріалом, створюється формальна логіко-лінгвістична модель [39], що об'єднує в собі всі члени речення природної мови. Вона дозволяє представити довільне речення природної мови у вигляді кон'юнкції атомарних предикатів, кожен із яких описує змістовно неподільну частину речення:

$$L^S = \bigwedge_{p \in P^S} \bigwedge_{h \in H_p^S} L_p^S(h), \quad (1.1)$$

$$L_p^S(h) = \bigwedge_{x \in X_p^S(h)} \bigwedge_{g \in G_p^S(x,h)} L_p^S(x,g,h), \quad (1.2)$$

$$L_p^S(x,g,h) = \bigwedge_{y \in Y_p^S(x,g,h)} \bigwedge_{q \in Q_p^S(x,g,y,h)} L_p^S(x,g,y,q,h), \quad (1.3)$$

$$L_p^S(x,g,y,q,h) = \bigwedge_{z \in Z_p^S(x,g,y,q,h)} \bigwedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x,g,y,q,z,r,h), \quad (1.4)$$

де S – речення природної мови;

p – відношення, що пов'язує суб'єкти, об'єкти та предмети відношень у реченні S , $p \in P^S$ належить множині відношень, що входять до речення S ;

h – характеристика p -го відношення речення S , $h \in H_p^S$ – множина характеристик p -го відношення у реченні S ;

$L_p^S(h)$ – предикат (предикативний вираз) [40], який описує p -е відношення з h -ю характеристикою і пов'язує суб'єкти, об'єкти та предмети відношення p у реченні S ;

x – суб'єкт речення S ;

$x \in X_p^S(h)$ – множина суб'єктів, що пов'язані з об'єктами речення S p -м відношенням, яке володіє h -ю характеристикою;

g – характеристика суб'єкта x речення S , $g \in G_p^S(x, h)$ – множина характеристик суб'єкта $x \in X_p^S(h)$;

$L_p^S(x, g, h)$ – предикат (предикативний вираз), який описує p -е відношення з h -ю характеристикою між суб'єктом $x \in X_p^S(h)$ з характеристикою $g \in G_p^S(x, h)$, об'єктами та предметами p -го відношення у реченні S ;

y – об'єкт речення S , $y \in Y_p^S(x, g, h)$ – множина об'єктів, що пов'язані з суб'єктами речення S p -м відношенням, яке володіє h -ю характеристикою;

q – характеристика об'єкта y речення S , $q \in Q_p^S(x, g, y, h)$ – множина характеристик об'єкта $y \in Y_p^S(x, g, h)$;

$L_p^S(x, g, y, q, h)$ – предикат (предикативний вираз), який описує p -е відношення з h -ю характеристикою між суб'єктом $x \in X_p^S(h)$ з характеристикою $g \in G_p^S(x, h)$ і об'єктом $y \in Y_p^S(x, g, h)$ з характеристикою $q \in Q_p^S(x, g, y, h)$ та предмети p -го відношення у реченні S ;

z – предмет p -го відношення речення S , $z \in Z_p^S(x, g, y, q, h)$ – множина предметів p -го відношення, яке володіє h -ю характеристикою, між суб'єктом $x \in X_p^S(h)$ з характеристикою $g \in G_p^S(x, h)$ та об'єктом $y \in Y_p^S(x, g, h)$ з характеристикою $q \in Q_p^S(x, g, y, h)$;

r – характеристика предмета p -го відношення речення S , $r \in R_p^S(x, g, y, q, z, h)$ – множина характеристик предмета $z \in Z_p^S(x, g, y, q, h)$;

$L_p^S(x, g, y, q, z, r, h)$ – простий предикат, який описує частину речення, що має закінчений зміст та відображає у реченні S p -е відношення з h -ю характеристикою між суб'єктом $x \in X_p^S(h)$ з характеристикою $g \in G_p^S(x, h)$ і об'єктом $y \in Y_p^S(x, g, h)$ з характеристикою $q \in Q_p^S(x, g, y, h)$, предмет якого $z \in Z_p^S(x, g, y, q, h)$ володіє характеристикою $r \in R_p^S(x, g, y, q, z, h)$.

Формули (1.1) – (1.4) інтерпретують синтаксичну структуру речень природної мови та являються уніфікованою формою завдяки тому, що враховують можливість існування:

- декількох відношень між суб'єктами та об'єктами речення;
- декількох характеристик кожного з відношень;
- множини суб'єктів, об'єктів та предметів відношень між ними;
- декількох характеристик суб'єктів, об'єктів та предметів відношень між ними.

Кількість v^S частин речення S , що мають закінчений зміст і описуються простим предикатом $L_p^S(x, g, y, q, z, r, h)$, розраховується за формулами:

$$v^S = \sum_{p \in P^S} \sum_{h \in H_p^S} v_p^S(h), \quad (1.5)$$

$$v_p^S(h) = \sum_{x \in X_p^S(h)} \sum_{g \in G_p^S(x, h)} v_p^S(x, g, h), \quad (1.6)$$

$$v_p^S(x, g, h) = \sum_{y \in Y_p^S(x, g, h)} \sum_{q \in Q_p^S(x, g, y, h)} v_p^S(x, g, y, q, h), \quad (1.7)$$

$$v_p^S(x, g, y, q, h) = \sum_{z \in Z_p^S(x, g, y, q, h)} |R_p^S(x, g, y, q, z, h)|. \quad (1.8)$$

Тобто кількість простих предикатів v^S у формулі (1.1) дорівнює сумі можливих комбінацій узгодження суб'єктів, об'єктів, предметів відношень між ними та їх характеристик.

Логіко-лінгвістична модель L^S речення S відображається сукупністю формул (1.1) – (1.4) та формально представляє собою

послідовність восьми кон'юнкцій, що входять до цих формул. Перехід від загальної формули (1.1) до предикату $L_p^S(x, g, y, q, z, r, h)$ є декомпозицією проблеми формального опису довільного речення природної мови та відображає системний підхід до її вирішення.

При побудові логіко-лінгвістичних моделей речень природної мови необхідно дотримуватися наступних правил.

1. Просте речення природної мови інтерпретується простим предикатом, у якому відсутні верхні індекси.

2. Якщо у реченні є слова, що тільки в сукупності виконують певну синтаксичну роль, то в логіко-лінгвістичній моделі вони відображаються як елементи однієї множини з різними індексами, об'єднані знаком кон'юнкції "&".

3. Якщо потужність множини певної компоненти логіко-лінгвістичної моделі більше одного, то елементи множини упорядковуються за індексами (i, j) , де i – номер речення, а j – номер елемента у відповідній множині.

4. Номери елементів j у множинах призначаються відповідно до порядку вживання цих елементів у реченні та не залежать від кількості пов'язаних з ними компонент логіко-лінгвістичної моделі.

5. Якщо суб'єкт відношення пов'язаний у реченні з іншими членами речення, з якими не пов'язане відношення, то створюється атомарний предикат, відношенням у якому буде суб'єкт логіко-лінгвістичної моделі головного речення.

6. Якщо у реченні йдеться про протиставлення, то кон'юнкція "&" в логіко-лінгвістичних моделях замінюється на диз'юнкцію " \vee ".

7. Верхній індекс в формулах (у більшості випадків, штрих або декілька штрихів) означає номер простого предикату.

8. Атомарні предикати простого речення у складі складного об'єднуються квадратними скобками.

9. При побудові логіко-лінгвістичних моделей складних або простих, ускладнених дієприкметниковим або дієприслівниковим зворотом, речень природної мови елементи, що замінюються, вживаються з граматичними параметрами у початковій формі.

1.2. Класифікація окремих форм логіко-лінгвістичних моделей

До основних ознак будь-якого речення відноситься [33]:

- змістовна завершеність – речення є одиницею повідомлення та виражає закінчену думку;
- граматична організація – будова речення (кількість граматичних основ) та граматичне вираження його членів (способи вираження членів речення і синтаксичних відношень між частинами складного речення та порядок слів).

У реченні, крім підрядних зв'язків узгодження, керування і прилягання, може мати місце і сурядний зв'язок, яким поєднуються однорідні члени речення. У реченні може бути кілька рядів однорідних членів, можуть вживатися відокремлені другорядні члени речення. Крім того, існують різноманітні види складних речень, для поєднання складових частин яких вживаються різні види сполучників та знаків пунктуації [41]. Формалізація опису таких різноманітних за структурою речень природної мови спонукає до здійснення класифікації окремих форм уніфікованої логіко-лінгвістичної моделі (1.1) – (1.4). Усі окремі форми адаптовані до певного типу речення природної мови, у якому одна або декілька із множин відношень, суб'єктів, об'єктів, предметів відношень та їх характеристик містять більше, ніж один елемент. При цьому $W(S) = \{1, \dots, w_l, \dots, 15\}$ – множина типів окремих форм, елементи якої набувають значення $w_l = \overline{1,15}$ [42].

Тип 1. Описує просте речення природної мови, що містить лише одну частину, яка має закінчений зміст:

$$L^S = L_p^S(x, g, y, q, z, r, h). \quad (1.9)$$

При цьому $|P^S|=1$; $|H_p^S|=1$; $|X_p^S(h)|=1$; $|G_p^S(x, h)|=1$;
 $|Y_p^S(x, g, h)|=1$; $|Q_p^S(x, g, y, h)|=1$; $|Z_p^S(x, g, y, q, h)|=1$;
 $|R_p^S(x, g, y, q, z, h)|=1$, кількість частин речення S , що мають закінчений зміст $v^S = 1$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.9) для конкретних значень елементів множин, що формують просте, не ускладнене речення природної мови S_i , має вигляд:

$$L^S = p_i(x_i, g_i, y_i, q_i, z_i, r_i, h_i).$$

Наприклад, для простого речення української мови «Нові дисципліни регулярно включають до навчальних планів» буде побудована логіко-лінгвістична модель типу 1 і буде мати вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(0, 0, y_1, q_1, z_1, r_1, h_1),$$

$L^S =$ включають (0, 0, дисципліни, нові, планів, навчальних, регулярно).

Для простого речення російської мови «Новые дисциплины регулярно включают в учебные планы» логіко-лінгвістична модель також буде типу 1 і матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(0, 0, y_1, q_1, z_1, r_1, h_1),$$

$L^S =$ включают (0, 0, дисциплины, новые, планы, учебные, регулярно).

Для простого речення англійської мови «New disciplines are regularly included in curricula» логіко-лінгвістична модель також типу 1 і має вигляд:

$$L^S = p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, 0, h_1),$$

$L^S =$ are&included (0, 0, disciplines, new, curricula, 0, regularly).

Тип 2. Описує просте речення з однорідними присудками, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює потужності множини відношень:

$$L^S = \bigwedge_{p \in P^S} L_p^S(x, g, y, q, z, r, h). \quad (1.10)$$

Таким чином, множина відношень P^S , що пов'язують суб'єкт, об'єкт та предмет відношень у реченні S містить декілька елементів $|P^S| \geq 2$, а $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$;

$|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |P^S|$, так як:

$$\begin{aligned} L^S &= \bigwedge_{p \in P^S} L_p^S(h), \\ L_p^S(h) &= L_p^S(x, g, h), \\ L_p^S(x, g, h) &= L_p^S(x, g, y, q, h), \\ L_p^S(x, g, y, q, h) &= L_p^S(x, g, y, q, z, r, h). \end{aligned}$$

Логіко-лінгвістична модель (1.10) конкретного речення S_i з двома однорідними присудками має вигляд:

$$L^{S_i} = p_{i1}(x_i, g_i, y_i, q_i, z_i, r_i, h_i) \& p_{i2}(x_i, g_i, y_i, q_i, z_i, r_i, h_i).$$

Наприклад, для речення української мови «Нові дисципліни регулярно враховують та включають до навчальних планів» буде побудована логіко-лінгвістична модель типу 2, що матиме вигляд:

$$\begin{aligned} L^{S_i} &= p_{11}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = \\ & p_{11}(0, 0, y_1, q_1, z_1, r_1, h_1) \& p_{12}(0, 0, y_1, q_1, z_1, r_1, h_1), \end{aligned}$$

$L^{S_i} =$ *враховують (0, 0, дисципліни, нові, планів, навчальних, регулярно) & включають (0, 0, дисципліни, нові, планів, навчальних, регулярно).*

Для речення російської мови «Новые дисциплины регулярно учитывают и включают в учебные планы» буде побудована логіко-лінгвістична модель типу 2, що матиме вигляд:

$$\begin{aligned} L^{S_i} &= p_{11}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = \\ & p_{11}(0, 0, y_1, q_1, z_1, r_1, h_1) \& p_{12}(0, 0, y_1, q_1, z_1, r_1, h_1), \end{aligned}$$

$L^{S_i} =$ *учитывают (0, 0, дисциплины, новые, планы, учебные, регулярно) & включают (0, 0, дисциплины, новые, планы, учебные, регулярно).*

Для речення англійської мови «New disciplines are regularly considered and included in curricula» буде побудована логіко-лінгвістична модель типу 2, що матиме вигляд:

$$\begin{aligned} L^{S_i} &= p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{11} \& p_{13}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = \\ & p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, 0, h_1) \& p_{11} \& p_{13}(0, 0, y_1, q_1, z_1, 0, h_1), \end{aligned}$$

$L^S = \text{are\&considered } (0, 0, \text{ disciplines, new, curricula, } 0, \text{ regularly})\&$
are\&included (0, 0, disciplines, new, curricula, 0, regularly).

Тип 3. Описує просте речення з однорідними обставинами, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює потужності множини характеристик відношення:

$$L^S = \bigwedge_{h \in H_p^S} L_p^S(x, g, y, q, z, r, h). \quad (1.11)$$

Множина характеристик відношення H_p^S у реченні S містить декілька елементів $|H_p^S| \geq 2$, а $|P^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |H_p^S|$, так як:

$$\begin{aligned} L^S &= \bigwedge_{h \in H_p^S} L_p^S(h), \\ L_p^S(h) &= L_p^S(x, g, h), \\ L_p^S(x, g, h) &= L_p^S(x, g, y, q, h), \\ L_p^S(x, g, y, q, h) &= L_p^S(x, g, y, q, z, r, h). \end{aligned}$$

Логіко-лінгвістична модель (1.11) конкретного речення S_i з двома однорідними обставинами має вигляд:

$$L^{S_i} = p_i(x_i, g_i, y_i, q_i, z_i, r_i, h_{i1}) \& p_i(x_i, g_i, y_i, q_i, z_i, r_i, h_{i2}).$$

Наприклад, для речення природної мови «Нові дисципліни регулярно та скрупульозно включають до навчальних планів» буде побудована логіко-лінгвістична модель типу 3, що матиме вигляд:

$$\begin{aligned} L^{S_i} &= p_i(x_i, g_i, y_i, q_i, z_i, r_i, h_{i1}) \& p_i(x_i, g_i, y_i, q_i, z_i, r_i, h_{i2}) = \\ & p_i(0, 0, y_i, q_i, z_i, r_i, h_{i1}) \& p_i(0, 0, y_i, q_i, z_i, r_i, h_{i2}), \end{aligned}$$

$L^{S_i} = \text{включають } (0, 0, \text{ дисципліни, нові, планів, навчальних, регулярно})\&$
включають (0, 0, дисципліни, нові, планів, навчальних, скрупульозно).

Для речення російської мови «*Новые дисциплины регулярно и скрупулезно включают в учебные планы*» буде побудована логіко-лінгвістична модель типу 3, що матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_{11}) \& p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_{12}) = \\ p_1(0, 0, y_1, q_1, z_1, r_1, h_{11}) \& p_1(0, 0, y_1, q_1, z_1, r_1, h_{12}), \\ L^S = \text{включают} (0, 0, \text{дисциплины, новые, планы, учебные, регулярно}) \& \\ \text{включают} (0, 0, \text{дисциплины, новые, планы, учебные, скрупулезно}).$$

Для речення англійської мови «*New disciplines are regularly and scrupulously included in curricula*» буде побудована логіко-лінгвістична модель типу 3, що матиме вигляд:

$$L^S = p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_{11}) \& p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_{12}) = \\ p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, 0, h_{11}) \& p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, 0, h_{12}), \\ L^S = \text{are\&included} (0, 0, \text{disciplines, new, curricula, 0, regularly}) \& \\ \text{are\&included} (0, 0, \text{disciplines, new, curricula, 0, scrupulously}).$$

Тип 4. Описує просте речення з однорідними підметами, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює потужності множини суб'єктів:

$$L^S = \bigwedge_{x \in X_p^S(h)} L_p^S(x, g, y, q, z, r, h). \quad (1.12)$$

Множина суб'єктів $X_p^S(h)$, що пов'язані з об'єктом речення S p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів $|X_p^S(h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |X_p^S(h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = \bigwedge_{x \in X_p^S(h)} L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.12) конкретного речення S_i з двома однорідними підметами має вигляд:

$$L^{S_i} = p_i(x_{i1}, g_i, y_i, q_i, z_i, r_i, h_i) \& p_i(x_{i2}, g_i, y_i, q_i, z_i, r_i, h_i).$$

Наприклад, для речення української мови «*Провідні викладачі та методисти регулярно включають нові дисципліни до навчальних планів*» буде побудована логіко-лінгвістична модель типу 4 вигляду:

$$L^{S_i} = p_1(x_{11}, g_1, y_1, q_1, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_1, z_1, r_1, h_1) =$$

$$p_1(x_{11}, g_1, y_1, q_1, z_1, r_1, h_1) \& p_1(x_{12}, 0, y_1, q_1, z_1, r_1, h_1),$$

L^{S_i} = включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно) & включають (методисти, 0, дисципліни, нові, планів, навчальних, регулярно).

Для речення російської мови «*Ведущие преподаватели и методисты регулярно включают новые дисциплины в учебные планы*» буде побудована логіко-лінгвістична модель типу 4, що матиме вигляд:

$$L^{S_i} = p_1(x_{11}, g_1, y_1, q_1, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_1, z_1, r_1, h_1) =$$

$$p_1(x_{11}, g_1, y_1, q_1, z_1, r_1, h_1) \& p_1(x_{12}, 0, y_1, q_1, z_1, r_1, h_1),$$

L^{S_i} = включают (преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно) & включают (методисты, 0, дисциплины, новые, планы, учебные, регулярно).

Для речення англійської мови «*Leading teachers and methodologists regularly include new disciplines in curricula*» буде побудована логіко-лінгвістична модель типу 4, що матиме вигляд:

$$L^{S_i} = p_1(x_{11}, g_1, y_1, q_1, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_1, z_1, r_1, h_1) =$$

$$p_1(x_{11}, g_1, y_1, q_1, z_1, 0, h_1) \& p_1(x_{12}, 0, y_1, q_1, z_1, 0, h_1),$$

L^{S_i} = include (teachers, leading, disciplines, new, curricula, 0, regularly) & include (methodologists, 0, disciplines, new, curricula, 0, regularly).

Тип 5. Описує просте речення з однорідними означеннями, що характеризують підмет, внаслідок чого кількість частин, які мають закінчений зміст, дорівнює потужності множини характеристик суб'єкта відношення:

$$L^S = \bigwedge_{g \in G_p^S(x,h)} L_p^S(x, g, y, q, z, r, h). \quad (1.13)$$

Множина характеристик суб'єкта відношення $G_p^S(x, h)$ у реченні S містить декілька елементів $|G_p^S(x, h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |G_p^S(x, h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = \bigwedge_{g \in G_p^S(x,h)} L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.13) конкретного речення S_i з двома однорідними означеннями має вигляд:

$$L^{S_i} = p_i(x_i, g_{i1}, y_i, q_i, z_i, r_i, h_i) \& p_i(x_i, g_{i2}, y_i, q_i, z_i, r_i, h_i).$$

Наприклад, для речення природної мови «Провідні висококваліфіковані викладачі регулярно включають нові дисципліни до навчальних планів» буде побудована логіко-лінгвістична модель типу 5, що матиме вигляд:

$$L^{S_i} = p_1(x_1, g_{11}, y_1, q_1, z_1, r_1, h_1) \& p_1(x_1, g_{12}, y_1, q_1, z_1, r_1, h_1),$$

$L^{S_i} =$ включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно) & включають (викладачі, висококваліфіковані, дисципліни, нові, планів, навчальних, регулярно).

Для речення російської мови «Ведущие высококвалифицированные преподаватели регулярно включают

нові дисципліни в учебні плани» буде побудована логіко-лінгвістична модель типу 5, що матиме вигляд:

$$L^S = p_1(x_1, g_{11}, y_1, q_1, z_1, r_1, h_1) \& p_1(x_1, g_{12}, y_1, q_1, z_1, r_1, h_1),$$

L^S = включають (преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно) & включають (преподаватели, высококвалифицированные, дисциплины, новые, планы, учебные, регулярно).

Для речення англійської мови «*Leading teachers highly qualified regularly include new disciplines in curricula*» буде побудована логіко-лінгвістична модель типу 5, що матиме вигляд:

$$L^S = p_1(x_1, g_{11}, y_1, q_1, z_1, r_1, h_1) \& p_1(x_1, g_{12} \& g_{13}, y_1, q_1, z_1, r_1, h_1) =$$

$$p_1(x_1, g_{11}, y_1, q_1, z_1, 0, h_1) \& p_1(x_1, g_{12} \& g_{13}, y_1, q_1, z_1, 0, h_1),$$

L^S = include (teachers, leading, disciplines, new, curricula, 0, regularly) & include (teachers, highly & qualified, disciplines, new, curricula, 0, regularly).

Тип 6. Описує просте речення з однорідними додатками, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює потужності множини об'єктів:

$$L^S = \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.14)$$

Множина об'єктів $Y_p^S(x, g, h)$, що пов'язані з суб'єктом речення S p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів $|Y_p^S(x, g, h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |Y_p^S(x, g, h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.14) конкретного речення S_i з двома однорідними додатками має вигляд:

$$L^S_i = p_i(x_i, g_i, y_{i1}, q_i, z_i, r_i, h_i) \& p_i(x_i, g_i, y_{i2}, q_i, z_i, r_i, h_i).$$

Наприклад, для речення української мови «Нові дисципліни та оновлені кредити регулярно враховують у навчальних планах» буде побудована логіко-лінгвістична модель типу 6, що матиме вигляд:

$$L^S_i = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1) =$$

$$p_1(0, 0, y_{11}, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_{12}, q_{12}, z_1, r_1, h_1),$$

$L^S_i =$ *враховують (0, 0, дисципліни, нові, планах, навчальних, регулярно)& враховують (0, 0, кредити, оновлені, планах, навчальних, регулярно).*

У розглянутому реченні потужність множини характеристик об'єкту також дорівнює двом, проте характеристики за своїм значенням не однакові, тому тип окремої форми логіко-лінгвістичної моделі – 6.

Для речення російської мови «Новые дисциплины и обновленные кредиты регулярно учитывают в учебных планах» буде побудована логіко-лінгвістична модель типу 6, що матиме вигляд:

$$L^S_i = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1) =$$

$$p_1(0, 0, y_{11}, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_{12}, q_{12}, z_1, r_1, h_1),$$

$L^S_i =$ *учитывают (0, 0, дисциплины, новые, планах, учебных, регулярно)& учитывают (0, 0, кредиты, обновленные, планах, учебных, регулярно).*

Для речення англійської мови «New disciplines and updated loans are regularly considered in curricula» буде побудована логіко-лінгвістична модель типу 6, що матиме вигляд:

$$L^S_i = p_{11} \& p_{12}(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1) =$$

$$p_{11} \& p_{12}(0, 0, y_{11}, q_{11}, z_1, 0, h_1) \& p_{11} \& p_{12}(0, 0, y_{12}, q_{12}, z_1, 0, h_1),$$

$L^S_i =$ *are&considered (0, 0, disciplines, new, curricula, 0, regularly)&*

are&considered (0, 0, loans, updated, curricula, 0, regularly).

Тип 7. Описує просте речення з однорідними означеннями, що характеризують додаток, внаслідок чого кількість частин, які мають закінчений зміст, дорівнює потужності множини характеристик об'єкта відношення:

$$L^S = \bigwedge_{q \in Q_p^S(x, g, y, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.15)$$

Множина характеристик об'єкта відношення $Q_p^S(x, g, y, h)$ у реченні S містить декілька елементів $|Q_p^S(x, g, y, h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |Q_p^S(x, g, y, h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = \bigwedge_{q \in Q_p^S(x, g, y, h)} L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.15) конкретного речення S_i з двома однорідними означеннями має вигляд:

$$L^S = p_i(x_i, g_i, y_i, q_{i1}, z_i, r_i, h_i) \& p_i(x_i, g_i, y_i, q_{i2}, z_i, r_i, h_i).$$

Наприклад, для речення природної мови «Нові сучасні дисципліни регулярно включають до навчальних планів» буде побудована логіко-лінгвістична модель типу 7, що матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_1, g_1, y_1, q_{12}, z_1, r_1, h_1) =$$

$$p_1(0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_1, q_{12}, z_1, r_1, h_1),$$

$L^S =$ включають (0, 0, дисципліни, нові, планів, навчальних, регулярно)&

включають (0, 0, дисципліни, сучасні, планів, навчальних, регулярно).

Для речення російської мови «*Новые современные дисциплины регулярно включают в учебные планы*» буде побудована логіко-лінгвістична модель типу 7, що матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_1, g_1, y_1, q_{12}, z_1, r_1, h_1) =$$

$$p_1(0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_1, q_{12}, z_1, r_1, h_1),$$

$L^S =$ *включают (0, 0, дисциплины, новые, планы, учебные, регулярно)* &
включают (0, 0, дисциплины, современные, планы, учебные, регулярно).

Для речення англійської мови «*New modern disciplines are regularly included in curricula*» буде побудована логіко-лінгвістична модель типу 7, що матиме вигляд:

$$L^S = p_{11} \& p_{12}(x_1, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_1, y_1, q_{12}, z_1, r_1, h_1) =$$

$$p_{11} \& p_{12}(0, 0, y_1, q_{11}, z_1, 0, h_1) \& p_{11} \& p_{12}(0, 0, y_1, q_{12}, z_1, 0, h_1),$$

$L^S =$ *are&included (0, 0, disciplines, new, curricula, 0, regularly)* &
are&included (0, 0, disciplines, modern, curricula, 0, regularly).

Тип 8. Описує просте речення з однорідними додатками, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює потужності множини предметів відношення:

$$L^S = \bigwedge_{z \in Z_p^S(x, g, y, q, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.16)$$

Множина предметів p -го відношення $Z_p^S(x, g, y, q, h)$, яке володіє h -ю характеристикою, містить декілька елементів $|Z_p^S(x, g, y, q, h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |Z_p^S(x, g, y, q, h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = \bigwedge_{z \in Z_p^S(x, g, y, q, h)} L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.16) конкретного речення S_i з двома однорідними додатками має вигляд:

$$L^S = p_i(x_i, g_i, y_i, q_i, z_{i1}, r_i, h_i) \& p_i(x_i, g_i, y_i, q_i, z_{i2}, r_i, h_i).$$

Наприклад, для речення природної мови «Нові дисципліни регулярно включають до навчальних планів та програм» буде побудована логіко-лінгвістична модель типу 8, що матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_1, z_{11}, r_1, h_1) \& p_1(x_1, g_1, y_1, q_1, z_{12}, r_1, h_1) =$$

$$p_1(0, 0, y_1, q_1, z_{11}, r_1, h_1) \& p_1(0, 0, y_1, q_1, z_{12}, r_1, h_1),$$

$L^S =$ включають (0, 0, дисципліни, нові, планів, навчальних, регулярно)&

включають (0, 0, дисципліни, нові, програм, навчальних, регулярно).

Для речення російської мови «Новые дисциплины регулярно включают в учебные планы и программы» буде побудована логіко-лінгвістична модель типу 8, що матиме вигляд:

$$L^S = p_1(x_1, g_1, y_1, q_1, z_{11}, r_1, h_1) \& p_1(x_1, g_1, y_1, q_1, z_{12}, r_1, h_1) =$$

$$p_1(0, 0, y_1, q_1, z_{11}, r_1, h_1) \& p_1(0, 0, y_1, q_1, z_{12}, r_1, h_1),$$

$L^S =$ включают (0, 0, дисциплины, новые, планы, учебные, регулярно)&

включают (0, 0, дисциплины, новые, программы, учебные, регулярно).

Для речення англійської мови «New disciplines are regularly included in curricula and educational programs» буде побудована логіко-лінгвістична модель типу 8, що матиме вигляд:

$$L^S = p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_{12}, r_1, h_1) =$$

$$p_{11} \& p_{12}(0, 0, y_1, q_1, z_{11}, 0, h_1) \& p_{11} \& p_{12}(0, 0, y_1, q_1, z_{12}, r_1, h_1),$$

$L^S =$ are&included (0, 0, disciplines, new, curricula, 0, regularly)&

are&included (0, 0, disciplines, new, programs, educational, regularly).

Тип 9. Описує просте речення з однорідними означеннями, що характеризують додаток, внаслідок чого кількість частин, які

мають закінчений зміст, дорівнює потужності множини характеристик предмета відношення:

$$L^S = \bigwedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x,g,y,q,z,r,h). \quad (1.17)$$

Множина характеристик об'єкта відношення $Q_p^S(x,g,y,h)$ у реченні S містить декілька елементів $|R_p^S(x,g,y,q,z,h)| \geq 2$, а $|P^S| = 1$; $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x,h)| = 1$; $|Y_p^S(x,g,h)| = 1$; $|Q_p^S(x,g,y,h)| = 1$; $|Z_p^S(x,g,y,q,h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |R_p^S(x,g,y,q,z,h)|$, так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = L_p^S(x,g,h),$$

$$L_p^S(x,g,h) = L_p^S(x,g,y,q,h),$$

$$L_p^S(x,g,y,q,h) = \bigwedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x,g,y,q,z,r,h).$$

Логіко-лінгвістична модель (1.17) конкретного речення S_i з двома однорідними означеннями має вигляд:

$$L^{S_i} = p_i(x_i, g_i, y_i, q_i, z_i, r_{i1}, h_i) \& p_i(x_i, g_i, y_i, q_i, z_i, r_{i2}, h_i).$$

Наприклад, для речення природної мови «Нові дисципліни регулярно включають до навчальних та робочих планів» буде побудована логіко-лінгвістична модель типу 9, що матиме вигляд:

$$L^{S_i} = p_1(x_1, g_1, y_1, q_1, z_1, r_{11}, h_1) \& p_1(x_1, g_1, y_1, q_1, z_1, r_{12}, h_1) =$$

$$p_1(0, 0, y_1, q_1, z_1, r_{11}, h_1) \& p_1(0, 0, y_1, q_1, z_1, r_{12}, h_1),$$

$L^{S_i} =$ включають (0, 0, дисципліни, нові, планів, навчальних, регулярно) &

включають (0, 0, дисципліни, нові, планів, робочих, регулярно).

Для речення російської мови «Новые дисциплины регулярно включают в учебные и рабочие планы» буде побудована логіко-лінгвістична модель типу 9, що матиме вигляд:

$$L^{S_i} = p_1(x_1, g_1, y_1, q_1, z_1, r_{11}, h_1) \& p_1(x_1, g_1, y_1, q_1, z_1, r_{12}, h_1) =$$

$$p_1(0, 0, y_1, q_1, z_1, r_{11}, h_1) \& p_1(0, 0, y_1, q_1, z_1, r_{12}, h_1),$$

L^S = включають (0, 0, дисципліни, нові, плани, навчальні, регулярно)&

включають (0, 0, дисципліни, нові, плани, робочі, регулярно).

Для речення англійської мови «*New disciplines are regularly included in educational and work plans*» буде побудована логіко-лінгвістична модель типу 9, що матиме вигляд:

$$L^S = p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_{11}, h_1) \& p_{11} \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_{12}, h_1) = p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, r_{11}, h_1) \& p_{11} \& p_{12}(0, 0, y_1, q_1, z_1, r_{12}, h_1),$$

L^S = are&included (0, 0, disciplines, new, plans, educational, regularly)&
are&included (0, 0, disciplines, new, plans, work, regularly).

Тип 10. Характеризує типове сполучення членів речення природної мови та описує просте речення з однорідними підметами, додатками та обставинами, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює сумі потужностей множин суб'єктів, об'єктів та характеристик відношення:

$$L^S = \bigwedge_{h \in H_p^S} \bigwedge_{x \in X_p^S(h)} \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.18)$$

У цьому випадку множини елементів логіко-лінгвістичної моделі мають такі потужності: $|H_p^S| \geq 2$; $|X_p^S(h)| \geq 2$; $|Y_p^S(x, g, h)| \geq 2$, а $|P^S| = 1$; $|G_p^S(x, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |H_p^S| + |X_p^S(h)| + |Y_p^S(x, g, h)|$, так як:

$$L^S = \bigwedge_{h \in H_p^S} L_p^S(h),$$

$$L_p^S(h) = \bigwedge_{x \in X_p^S(h)} L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.18) конкретного речення S_i з двома однорідними підметами, додатками та обставинами має вигляд:

$$L^{S_i} = p_i(x_{i1}, g_i, y_{i1}, q_i, z_i, r_i, h_{i1}) \& p_i(x_{i1}, g_i, y_{i2}, q_i, z_i, r_i, h_{i1}) \& \\ p_i(x_{i1}, g_i, y_{i1}, q_i, z_i, r_i, h_{i2}) \& p_i(x_{i2}, g_i, y_{i1}, q_i, z_i, r_i, h_{i1}) \& \\ p_i(x_{i2}, g_i, y_{i2}, q_i, z_i, r_i, h_{i1}) \& p_i(x_{i2}, g_i, y_{i2}, q_i, z_i, r_i, h_{i2}).$$

Наприклад, для речення української мови «*Провідні викладачі та методисти регулярно і скрупульозно включають нові дисципліни та оновлені кредити до навчальних планів*» буде побудована логіко-лінгвістична модель типу 10 вигляду:

$$L^{S_i} = p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, r_1, h_{11}) \& p_1(x_{11}, g_{11}, y_{12}, q_{12}, z_1, r_1, h_{11}) \& \\ p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, r_1, h_{12}) \& p_1(x_{12}, 0, y_{11}, q_{11}, z_1, r_1, h_{11}) \& \\ p_1(x_{12}, 0, y_{12}, q_{12}, z_1, r_1, h_{11}) \& p_1(x_{12}, 0, y_{12}, q_{12}, z_1, r_1, h_{12}), \\ L^{S_i} = \text{включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, провідні, кредити, оновлені, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, провідні, дисципліни, нові, планів, навчальних, скрупульозно)} \& \\ \text{включають (методисти, 0, дисципліни, нові, планів, навчальних, регулярно)} \& \\ \text{включають (методисти, 0, кредити, оновлені, планів, навчальних, регулярно)} \& \\ \text{включають (методисти, 0, дисципліни, нові, планів, навчальних, скрупульозно)}.$$

Для речення російської мови «*Ведущие преподаватели и методисты регулярно и скрупулезно включают новые дисциплины и обновленные кредиты в учебные планы*» буде побудована логіко-лінгвістична модель типу 10, що матиме вигляд:

$$L^{S_i} = p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, r_1, h_{11}) \& p_1(x_{11}, g_{11}, y_{12}, q_{12}, z_1, r_1, h_{11}) \& \\ p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, r_1, h_{12}) \& p_1(x_{12}, 0, y_{11}, q_{11}, z_1, r_1, h_{11}) \& \\ p_1(x_{12}, 0, y_{12}, q_{12}, z_1, r_1, h_{11}) \& p_1(x_{12}, 0, y_{12}, q_{12}, z_1, r_1, h_{12}), \\ L^{S_i} = \text{включают (преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно)} \&$$

включают (преподаватели, ведущие, кредиты, обновленные, планы, учебные, регулярно) &
 включают (преподаватели, ведущие, дисциплины, новые, планы, учебные, скрупулезно) &
 включают (методисты, 0, дисциплины, новые, планы, учебные, регулярно) &
 включают (методисты, 0, кредиты, обновленные, планы, учебные, регулярно) &
 включают (методисты, 0, дисциплины, новые, планы, учебные, скрупулезно).

.Для речення англійської мови «*Leading teachers and methodologists regularly and scrupulously include new disciplines and updated loans in curricula*» буде побудована логіко-лінгвістична модель типу 10, що матиме вигляд:

$$L^S = p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, 0, h_{11}) \& p_1(x_{11}, g_{11}, y_{12}, q_{12}, z_1, 0, h_{11}) \&$$

$$p_1(x_{11}, g_{11}, y_{11}, q_{11}, z_1, 0, h_{12}) \& p_1(x_{12}, 0, y_{11}, q_{11}, z_1, 0, h_{11}) \&$$

$$p_1(x_{12}, 0, y_{12}, q_{12}, z_1, 0, h_{11}) \& p_1(x_{12}, 0, y_{12}, q_{12}, z_1, 0, h_{12}),$$

L^S = include (teachers, leading, disciplines, new, curricula, 0, regularly) &
include (teachers, leading, loans, updated, curricula, 0, regularly) &
include (teachers, leading, disciplines, new, curricula, 0, scrupulously) &
include (methodologists, 0, disciplines, new, curricula, 0, regularly) &
include (methodologists, 0, loans, updated, curricula, 0, regularly) &
include (methodologists, 0, disciplines, new, curricula, 0, scrupulously).

Тип 11. Відображає типове сполучення членів речення природної мови та описує просте речення з однорідними присудками і додатками, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює сумі потужностей множин відношень та об'єктів відношень:

$$L^S = \bigwedge_{p \in P^S} \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.19)$$

У цьому випадку множини елементів логіко-лінгвістичної моделі мають такі потужності: $|P^S| \geq 2$; $|Y_p^S(x, g, h)| \geq 2$, а $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |P^S| + |Y_p^S(x, g, h)|$, так як:

$$L^S = \bigwedge_{p \in P^S} L_p^S(h),$$

$$L_p^S(h) = L_p^S(x, g, h),$$

$$L_p^S(x, g, h) = \bigwedge_{y \in Y_p^S(x, g, h)} L_p^S(x, g, y, q, h),$$

$$L_p^S(x, g, y, q, h) = L_p^S(x, g, y, q, z, r, h).$$

Логіко-лінгвістична модель (1.19) конкретного речення S_i з двома однорідними присудками та додатками має вигляд:

$$L^{S_i} = p_{i1}(x_i, g_i, y_{i1}, q_i, z_i, r_i, h_i) \& p_{i1}(x_i, g_i, y_{i2}, q_i, z_i, r_i, h_i) \& p_{i2}(x_i, g_i, y_{i1}, q_i, z_i, r_i, h_i) \& p_{i2}(x_i, g_i, y_{i2}, q_i, z_i, r_i, h_i).$$

Наприклад, для речення української мови «Провідні викладачі регулярно враховують і включають нові дисципліни та оновлені кредити до навчальних планів» буде побудована логіко-лінгвістична модель типу 11 вигляду:

$$L^{S_i} = p_{11}(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_{11}(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1),$$

L^{S_i} = враховують (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно) & враховують (викладачі, провідні, кредити, оновлені, планів, навчальних, регулярно) & включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно) & включають (викладачі, провідні, кредити, оновлені, планів, навчальних, регулярно).

Для речення російської мови «Ведущие преподаватели регулярно учитывают и включают новые дисциплины и обновленные кредиты в учебные планы» буде побудована логіко-лінгвістична модель типу 11, що матиме вигляд:

$$L^S = p_{11}(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_{11}(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1) \& \\ p_{12}(x_1, g_1, y_{11}, q_{11}, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_{12}, q_{12}, z_1, r_1, h_1),$$

L^S = учитывают (преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно) & учитывают (преподаватели, ведущие, кредиты, обновленные, планы, учебные, регулярно) & включают (преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно) & включают (преподаватели, ведущие, кредиты, обновленные, планы, учебные, регулярно).

Для речення англійської мови «*Leading teachers regularly consider and include new disciplines and updated loans in curricula*» буде побудована логіко-лінгвістична модель типу 11, що матиме вигляд:

$$L^S = p_{11}(x_1, g_1, y_{11}, q_{11}, z_1, 0, h_1) \& p_{11}(x_1, g_1, y_{12}, q_{12}, z_1, 0, h_1) \& \\ p_{12}(x_1, g_1, y_{11}, q_{11}, z_1, 0, h_1) \& p_{12}(x_1, g_1, y_{12}, q_{12}, z_1, 0, h_1),$$

L^S = consider (teachers, leading, disciplines, new, curricula, 0, regularly) & consider (teachers, leading, loans, updated, curricula, 0, regularly) & include (teachers, leading, disciplines, new, curricula, 0, regularly) & include (teachers, leading, loans, updated, curricula, 0, regularly).

Тип 12. Відображає типове сполучення членів речення природної мови та описує просте речення з однорідними означеннями і додатками, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює сумі потужностей множин характеристик суб'єкта, об'єктів та предметів відношень:

$$L^S = \bigwedge_{g \in G_p^S(x, h)} \bigwedge_{y \in Y_p^S(x, g, h)} \bigwedge_{z \in Z_p^S(x, g, y, q, h)} L_p^S(x, g, y, q, z, r, h). \quad (1.20)$$

У цьому випадку множини елементів логіко-лінгвістичної моделі мають такі потужності: $|G_p^S(x, h)| \geq 2$; $|Y_p^S(x, g, h)| \geq 2$; $|Z_p^S(x, g, y, q, h)| \geq 2$; а $|P^S| = 1$; $|H^S| = 1$; $|X_p^S(h)| = 1$; $|Q_p^S(x, g, y, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. У такому випадку кількість простих

предикатів у реченні дорівнює $v^S = |G_p^S(x,h)| + |Y_p^S(x,g,h)| + |Z_p^S(x,g,y,q,h)|$,

так як:

$$L^S = L_p^S(h),$$

$$L_p^S(h) = \bigwedge_{g \in G_p^S(x,h)} L_p^S(x,g,h),$$

$$L_p^S(x,g,h) = \bigwedge_{y \in Y_p^S(x,g,h)} L_p^S(x,g,y,q,h),$$

$$L_p^S(x,g,y,q,h) = \bigwedge_{z \in Z_p^S(x,g,y,q,h)} L_p^S(x,g,y,q,z,r,h).$$

Логіко-лінгвістична модель (1.20) конкретного речення S_i з двома однорідними означеннями та додатками має вигляд:

$$L^{S_i} = p_i(x_i, g_{i1}, y_{i1}, q_i, z_{i1}, r_i, h_i) \& p_i(x_i, g_{i1}, y_{i2}, q_i, z_{i1}, r_i, h_i) \& \\ p_i(x_i, g_{i1}, y_{i1}, q_i, z_{i2}, r_i, h_i) \& p_i(x_i, g_{i2}, y_{i1}, q_i, z_{i1}, r_i, h_i) \& \\ p_i(x_i, g_{i2}, y_{i2}, q_i, z_{i1}, r_i, h_i) \& p_i(x_i, g_{i2}, y_{i2}, q_i, z_{i2}, r_i, h_i).$$

Наприклад, для речення української мови «Провідні висококваліфіковані викладачі регулярно включають нові дисципліни та оновлені кредити до навчальних планів і програм» буде побудована логіко-лінгвістична модель типу 12 вигляду:

$$L^{S_i} = p_1(x_1, g_{11}, y_{11}, q_{11}, z_{11}, r_1, h_1) \& p_1(x_1, g_{11}, y_{12}, q_{12}, z_{11}, r_1, h_1) \& \\ p_1(x_1, g_{11}, y_{11}, q_{11}, z_{12}, r_1, h_1) \& p_1(x_1, g_{12}, y_{11}, q_{11}, z_{11}, r_1, h_1) \& \\ p_1(x_1, g_{12}, y_{12}, q_{12}, z_{11}, r_1, h_1) \& p_1(x_1, g_{12}, y_{12}, q_{12}, z_{12}, r_1, h_1), \\ L^{S_i} = \text{включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, провідні, кредити, оновлені, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, провідні, дисципліни, нові, програм, навчальних, регулярно)} \& \\ \text{включають (викладачі, висококваліфіковані, дисципліни, нові, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, висококваліфіковані, кредити, оновлені, планів, навчальних, регулярно)} \& \\ \text{включають (викладачі, висококваліфіковані, дисципліни, нові, програм, навчальних, регулярно)}.$$

Для речення російської мови «*Ведущие высококвалифицированные преподаватели регулярно включают новые дисциплины и обновленные кредиты в учебные планы и программы*» буде побудована логіко-лінгвістична модель типу 12, що матиме вигляд:

$$L^S = p_1(x_1, g_{11}, y_{11}, q_{11}, z_{11}, r_1, h_1) \& p_1(x_1, g_{11}, y_{12}, q_{12}, z_{11}, r_1, h_1) \&$$

$$p_1(x_1, g_{11}, y_{11}, q_{11}, z_{12}, r_1, h_1) \& p_1(x_1, g_{12}, y_{11}, q_{11}, z_{11}, r_1, h_1) \&$$

$$p_1(x_1, g_{12}, y_{12}, q_{12}, z_{11}, r_1, h_1) \& p_1(x_1, g_{12}, y_{12}, q_{12}, z_{12}, r_1, h_1),$$

$L^S =$ *включают(преподаватели, ведущие, дисциплины, новые, планы, учебные, регулярно) & включают (преподаватели, ведущие, кредиты, обновленные, планы, учебные, регулярно) & включают (преподаватели, ведущие, дисциплины, новые, программы, учебные, регулярно) & включают (преподаватели, высококвалифицированные, дисциплины, новые, планы, учебные, регулярно) & включают (преподаватели, высококвалифицированные, кредиты, обновленные, планы, учебные, регулярно) & включают (преподаватели, высококвалифицированные, дисциплины, новые, программы, учебные, регулярно).*

Для речення англійської мови «*Leading highly qualified teachers regularly include new disciplines and updated loans in curricula and work programs*» буде побудована логіко-лінгвістична модель типу 12, що матиме вигляд:

$$L^S = p_1(x_1, g_{11}, y_{11}, q_{11}, z_{11}, 0, h_1) \& p_1(x_1, g_{11}, y_{12}, q_{12}, z_{11}, 0, h_1) \&$$

$$p_1(x_1, g_{11}, y_{11}, q_{11}, z_{12}, r_1, h_1) \& p_1(x_1, g_{12} \& g_{13}, y_{11}, q_{11}, z_{11}, 0, h_1) \&$$

$$p_1(x_1, g_{12} \& g_{13}, y_{12}, q_{12}, z_{11}, 0, h_1) \& p_1(x_1, g_{12} \& g_{13}, y_{12}, q_{12}, z_{12}, r_1, h_1),$$

$L^S =$ *include (teachers, leading, disciplines, new, curricula, 0, regularly)& include (teachers, leading, loans, updated, curricula, 0, regularly)& include (teachers, leading, disciplines, new, programs, work, regularly)& include (teachers, highly&qualified, disciplines, new, curricula, 0, regularly)&*

include (teachers, highly&qualified, loans, updated, curricula, 0, regularly)&

include (teachers, highly&qualified, disciplines, new, programs, work, regularly).

Тип 13. Відображає типове сполучення членів речення природної мови та описує просте речення з однорідними додатками та обставинами, внаслідок чого кількість частин, що мають закінчений зміст, дорівнює сумі потужностей множин характеристик відношень, характеристик об'єктів та предметів відношень:

$$L^S = \bigwedge_{h \in H_p^S} \bigwedge_{q \in Q_p^S(x,g,y,h)} \bigwedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x,g,y,q,z,r,h). \quad (1.21)$$

У цьому випадку множини елементів логіко-лінгвістичної моделі мають такі потужності: $|H_p^S| \geq 2$; $|Q_p^S(x,g,y,h)| \geq 2$; $|R_p^S(x,g,y,q,z,h)| \geq 2$, а $|P^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x,h)| = 1$; $|Y_p^S(x,g,h)| \geq 2$; $|Z_p^S(x,g,y,q,h)| = 1$. У такому випадку кількість простих предикатів у реченні дорівнює $v^S = |H_p^S| + |Q_p^S(x,g,y,h)| + |R_p^S(x,g,y,q,z,h)|$, так як:

$$L^S = \bigwedge_{h \in H_p^S} L_p^S(h),$$

$$L_p^S(h) = L_p^S(x,g,h),$$

$$L_p^S(x,g,h) = \bigwedge_{q \in Q_p^S(x,g,y,h)} L_p^S(x,g,y,q,h),$$

$$L_p^S(x,g,y,q,h) = \bigwedge_{r \in R_p^S(x,g,y,q,z,h)} L_p^S(x,g,y,q,z,r,h).$$

Логіко-лінгвістична модель (1.21) конкретного речення S_i з двома однорідними обставинами та означеннями має вигляд:

$$L^{S_i} = p_i(x_i, g_i, y_i, q_{i1}, z_i, r_{i1}, h_{i1}) \& p_i(x_i, g_i, y_i, q_{i1}, z_i, r_{i2}, h_{i1}) \& \\ p_i(x_i, g_i, y_i, q_{i2}, z_i, r_{i1}, h_{i2}) \& p_i(x_i, g_i, y_i, q_{i2}, z_i, r_{i1}, h_{i1}) \& \\ p_i(x_i, g_i, y_i, q_{i2}, z_i, r_{i2}, h_{i1}) \& p_i(x_i, g_i, y_i, q_{i2}, z_i, r_{i2}, h_{i2}).$$

Наприклад, для речення української мови «Нові сучасні дисципліни регулярно і скрупульозно включають до навчальних та

робочих планів» буде побудована логіко-лінгвістична модель типу 13 вигляду:

$$L^S = p_1(0, 0, y_1, q_{11}, z_1, r_{11}, h_{11}) \& p_1(0, 0, y_1, q_{11}, z_1, r_{12}, h_{11}) \&$$

$$p_1(0, 0, y_1, q_{11}, z_1, r_{11}, h_{12}) \& p_1(0, 0, y_1, q_{12}, z_1, r_{11}, h_{11}) \&$$

$$p_1(0, 0, y_1, q_{12}, z_1, r_{12}, h_{11}) \& p_1(0, 0, y_1, q_{12}, z_1, r_{12}, h_{12}),$$

$L^S =$ включають (0, 0, дисципліни, нові, планів, навчальних, регулярно)&

включають (0, 0, дисципліни, нові, планів, робочих, регулярно)&

включають (0, 0, дисципліни, нові, планів, навчальних, скрупульозно)&

включають (0, 0, дисципліни, сучасні, планів, навчальних, регулярно)&

включають (0, 0, дисципліни, сучасні, планів, робочих, регулярно)&

включають (0, 0, дисципліни, сучасні, планів, навчальних, скрупульозно).

Для речення російської мови «*Новые современные дисциплины регулярно и скрупулезно включают в учебные и рабочие планы*» буде побудована логіко-лінгвістична модель типу 13, що матиме вигляд:

$$L^S = p_1(0, 0, y_1, q_{11}, z_1, r_{11}, h_{11}) \& p_1(0, 0, y_1, q_{11}, z_1, r_{12}, h_{11}) \&$$

$$p_1(0, 0, y_1, q_{11}, z_1, r_{11}, h_{12}) \& p_1(0, 0, y_1, q_{12}, z_1, r_{11}, h_{11}) \&$$

$$p_1(0, 0, y_1, q_{12}, z_1, r_{12}, h_{11}) \& p_1(0, 0, y_1, q_{12}, z_1, r_{12}, h_{12}),$$

$L^S =$ включают(0, 0, дисциплины, новые, планы, учебные, регулярно) &

включают (0, 0, дисциплины, новые, планы, рабочие, регулярно) &

включают (0, 0, дисциплины, новые, планы, учебные, скрупулезно) &

включают (0, 0, дисциплины, современные, планы, учебные, регулярно) &

включают 0, 0, дисциплины, современные, планы, рабочие, регулярно) &

включают (0, 0, дисциплины, современные, планы, учебные, скрупулезно).

Для речення англійської мови «*New modern disciplines are regularly and scrupulously included in educational and work*

plans» буде побудована логіко-лінгвістична модель типу 13, що матиме вигляд:

$L^S = p_{11} \& p_{12} (0, 0, y_1, q_{11}, z_1, r_{11}, h_{11}) \& p_{11} \& p_{12} (0, 0, y_1, q_{11}, z_1, r_{12}, h_{11}) \&$
 $p_{11} \& p_{12} (0, 0, y_1, q_{11}, z_1, r_{11}, h_{12}) \& p_{11} \& p_{12} (0, 0, y_1, q_{12}, z_1, r_{11}, h_{11}) \&$
 $p_{11} \& p_{12} (0, 0, y_1, q_{12}, z_1, r_{12}, h_{11}) \& p_{11} \& p_{12} (0, 0, y_1, q_{12}, z_1, r_{12}, h_{12}),$
L^S = are&included (0, 0, disciplines, new, plans, educational, regularly)&
are&included (0, 0, disciplines, new, plans, work, regularly)&
are&included (0, 0, disciplines, new, plans, educational, scrupulously)&
are&included (0, 0, disciplines, modern, plans, educational, regularly)&
are&included (0, 0, disciplines, modern, plans, work, regularly)&
are&included (0, 0, disciplines, modern, plans, educational, scrupulously).

Тип 14. Відображає типові зв'язки в складносурядному, безсполучниковому та складнопідрядному реченні природної мови, частини якого рівноправні за змістом, при цьому потужність множини відношень та множини суб'єктів відношень повинні бути не менше двох, тобто $|P^S| \geq 2$; $|X_p^S(h)| \geq 2$, а логіко-лінгвістична модель (1.1) – (1.4) набуває вигляду:

$$\begin{aligned}
 L^S_i &= L^S_{p^{(1)}}(x_i^{(1)}, g_i^{(1)}, y_i^{(1)}, q_i^{(1)}, z_i^{(1)}, r_i^{(1)}, h_i^{(1)}) \& \dots \& \\
 L^S_{p^{(\lambda)}}(x_i^{(\lambda)}, g_i^{(\lambda)}, y_i^{(\lambda)}, q_i^{(\lambda)}, z_i^{(\lambda)}, r_i^{(\lambda)}, h_i^{(\lambda)}) \& \dots \& & (1.22) \\
 L^S_{p^{(\phi)}}(x_i^{(\phi)}, g_i^{(\phi)}, y_i^{(\phi)}, q_i^{(\phi)}, z_i^{(\phi)}, r_i^{(\phi)}, h_i^{(\phi)}),
 \end{aligned}$$

де $L^S_{p^{(\lambda)}}(x_i^{(\lambda)}, g_i^{(\lambda)}, y_i^{(\lambda)}, q_i^{(\lambda)}, z_i^{(\lambda)}, r_i^{(\lambda)}, h_i^{(\lambda)})$ – предикатний вираз, що описує частину складного речення S_i , яка відображає закінчений зміст, виражається за допомогою простого речення рівня вкладеності $\lambda = \overline{1, \phi}$ і може набувати вигляд типу 1-13;

i – номер речення природної мови у тексті;

ϕ – кількість простих речень у складному.

Наприклад, для речення української мови «*Нові сучасні дисципліни регулярно включають до навчальних планів, але рівень викладання не змінюється*» логіко-лінгвістична модель (1.22) набуде вигляду:

$$L^S = [p_1(0,0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0,0, y_1, q_{12}, z_1, r_1, h_1)] \vee \neg p'_1(x'_1, g'_1, 0, 0, 0, 0, 0),$$

$$L^S = [включають(0, 0, дисципліни, нові, планів, навчальних, регулярно) \& включають(0, 0, дисципліни, сучасні, планів, навчальних, регулярно)] \vee \neg змінюється(рівень, викладання, 0, 0, 0, 0, 0).$$

Для речення російської мови «*Новые современные дисциплины регулярно включают в учебные планы, но уровень преподавания не изменяется*» буде побудована логіко-лінгвістична модель типу 14, що матиме вигляд:

$$L^S = [p_1(0,0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0,0, y_1, q_{12}, z_1, r_1, h_1)] \vee \neg p'_1(x'_1, g'_1, 0, 0, 0, 0, 0),$$

$$L^S = [включают(0, 0, дисциплины, новые, планы, учебные, регулярно) \& включают(0, 0, дисциплины, современные, планы, учебные, регулярно)] \vee \neg изменяется(уровень, преподавания, 0, 0, 0, 0, 0).$$

Для речення англійської мови «*New modern disciplines are regularly included in curricula, but the level of teaching does not change*» буде побудована логіко-лінгвістична модель типу 14, що матиме вигляд:

$$L^S = [p_{11} \& p_{12}(0,0, y_1, q_{11}, z_1, 0, h_1) \& p_{11} \& p_{12}(0,0, y_1, q_{12}, z_1, 0, h_1)] \vee \neg p'_1(x'_1, g'_1, 0, 0, 0, 0, 0),$$

$$L^S = [are \& included(0, 0, disciplines, new, curricula, 0, regularly) \& are \& included(0, 0, disciplines, modern, curricula, 0, regularly)] \vee \neg change(level, teaching, 0, 0, 0, 0, 0).$$

Ще один приклад побудови логіко-лінгвістичної моделі для складнопідрядного означального речення української мови «*Нові сучасні дисципліни регулярно включають до навчальних планів, які*

затверджуються Міністерством освіти і науки». Логіко-лінгвістична модель (1.22) набуде вигляду:

$$L^S = [p_1(0,0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0,0, y_1, q_{12}, z_1, r_1, h_1)] \&$$
$$[p'_1(x'_1, g'_{11}, z_1, r_1, 0, 0, 0) \& p'_1(x'_1, g'_{12}, z_1, r_1, 0, 0, 0),$$
$$L^S = [включають (0, 0, дисципліни, нові, планів, навчальних, регулярно) \&$$
$$включають (0, 0, дисципліни, сучасні, планів, навчальних, регулярно) \&$$
$$[затверджує (Міністерство, освіти, плани, навчальні, 0, 0, 0) \&$$
$$затверджує (Міністерство, науки, плани, навчальні, 0, 0, 0)].$$

Для речення російської мови «Новые современные дисциплины регулярно включают в учебные планы, которые утверждаются Министерством образования и науки» буде побудована логіко-лінгвістична модель типу 14, що матиме вигляд:

$$L^S = [p_1(0,0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0,0, y_1, q_{12}, z_1, r_1, h_1)] \&$$
$$[p'_1(x'_1, g'_{11}, z_1, r_1, 0, 0, 0) \& p'_1(x'_1, g'_{12}, z_1, r_1, 0, 0, 0),$$
$$L^S = [включают (0, 0, дисциплины, новые, планы, учебные, регулярно) \&$$
$$включают (0, 0, дисциплины, современные, планы, учебные, регулярно) \&$$
$$[утверждает (Министерство, образования, планы, учебные, 0, 0, 0) \&$$
$$утверждает (Министерство, науки, планы, учебные, 0, 0, 0)].$$

Для речення англійської мови «New modern disciplines are regularly included in curricula that are approved by the Ministry of education and science» буде побудована логіко-лінгвістична модель типу 14, що матиме вигляд:

$$L^S = [p_{11} \& p_{12}(0,0, y_1, q_{11}, z_1, 0, h_1) \& p_{11} \& p_{12}(0,0, y_1, q_{12}, z_1, 0, h_1)] \&$$
$$[p'_1(x'_1, g'_{11}, z_1, 0, 0, 0, 0) \& p'_1(x'_1, g'_{12}, z_1, 0, 0, 0, 0),$$
$$L^S = [are \& included (0, 0, disciplines, new, curricula, 0, regularly) \&$$
$$are \& included (0, 0, disciplines, modern, curricula, 0, regularly) \&$$
$$[approve (Ministry, education, curricula, 0, 0, 0, 0) \&$$
$$approve (Ministry, science, curricula, 0, 0, 0, 0)].$$

Тип 15. Відображає типові зв'язки в підрядному реченні природної мови, у якому йдеться про умови, мету, причину,

допуск, наслідок та ін. і для якого головна та залежна частини з'єднуються в логіко-лінгвістичній моделі логічною операцією імплікації " \rightarrow ", при цьому потужність множини відношень та множини суб'єктів відношень повинні бути не менше двох, тобто $|P^S| \geq 2$; $|X_p^S(h)| \geq 2$, а логіко-лінгвістична модель (1.1) – (1.4) набуває вигляду:

$$L_i^{S_i} = L_{p_i^{(\lambda)}}^{S_i}(x_i^{(\lambda)}, g_i^{(\lambda)}, y_i^{(\lambda)}, q_i^{(\lambda)}, z_i^{(\lambda)}, r_i^{(\lambda)}, h_i^{(\lambda)}) \rightarrow L_{p_i^{(\lambda+1)}}^{S_i}(x_i^{(\lambda+1)}, g_i^{(\lambda+1)}, y_i^{(\lambda+1)}, q_i^{(\lambda+1)}, z_i^{(\lambda+1)}, r_i^{(\lambda+1)}, h_i^{(\lambda+1)}), \quad (1.23)$$

де $L_{p_i^{(\lambda)}}^{S_i}(x_i^{(\lambda)}, g_i^{(\lambda)}, y_i^{(\lambda)}, q_i^{(\lambda)}, z_i^{(\lambda)}, r_i^{(\lambda)}, h_i^{(\lambda)})$ – предикатний вираз, що описує частину складного речення S_i , яка відображає закінчений зміст, виражається за допомогою простого речення рівня вкладеності $\lambda = \bar{1}, \varphi$ і може набувати вигляд типу 1-13;

i – номер речення природної мови у тексті;

φ – кількість простих речень у складному.

Наприклад, для речення української мови «Нові сучасні дисципліни регулярно включають до навчальних планів, внаслідок чого рівень навчання повинен піднятися» логіко-лінгвістична модель (1.23) набуває вигляду:

$$L_i^{S_i} = [p_1(0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_1, q_{12}, z_1, r_1, h_1)] \rightarrow p'_{11} \& p'_{12}(x'_1, g'_1, 0, 0, 0, 0, 0),$$

$L_i^{S_i} =$ [включають (0, 0, дисципліни, нові, планів, навчальних, регулярно)& включають (0, 0, дисципліни, сучасні, планів, навчальних, регулярно)] \rightarrow повинен&піднятися (рівень, навчання, 0, 0, 0, 0, 0).

Для речення російської мови «Новые современные дисциплины регулярно включают в учебные планы, вследствие чего уровень обучения должен подняться» буде побудована логіко-лінгвістична модель типу 15, що матиме вигляд:

$$L_i^{S_i} = [p_1(0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1(0, 0, y_1, q_{12}, z_1, r_1, h_1)] \rightarrow p'_{11} \& p'_{12}(x'_1, g'_1, 0, 0, 0, 0, 0),$$

$L_i^{S_i} =$ [включают(0, 0, дисциплины, новые, планы, учебные, регулярно) &

включают (0, 0, дисциплины, современные, планы, рабочие, регулярно)] &

должен&подняться (уровень, обучения, 0,0,0,0,0).

Для речення англійської мови «*New modern disciplines are regularly included in curricula, and as a result the level of education must rise*» буде побудована логіко-лінгвістична модель типу 15, що матиме вигляд:

$$L^S = [p_{11} \& p_{12} (0, 0, y_1, q_{11}, z_1, 0, h_1) \&$$
$$p_{11} \& p_{12} (0, 0, y_1, q_{12}, z_1, 0, h_1)] \rightarrow p'_{11} \& p'_{12} (x'_1, g'_1, 0, 0, 0, 0, 0),$$
$$L^S = [are\&included (0, 0, disciplines, new, curricula, 0, regularly)\&$$
$$are\&included (0, 0, disciplines, modern, curricula, 0, regularly)] \rightarrow$$
$$must\&rise (level, education, 0, 0, 0, 0, 0).$$

У випадку, коли обидві частини складного речення містять однорідні члени, для речення української мови «*Нові сучасні дисципліни регулярно включають до навчальних планів, що впливає на статус вищого навчального закладу*» логіко-лінгвістична модель (1.23) набуде вигляду:

$$L^S = [p_1 (0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1 (0, 0, y_1, q_{12}, z_1, r_1, h_1)] \rightarrow$$
$$[p'_1 (0, 0, y'_1, 0, z_1, r'_{11}, 0) \& p'_1 (0, 0, y'_1, 0, z_1, r'_{12}, 0)],$$
$$L^S = [включают (0, 0, дисципліни, нові, планів, навчальних, регулярно)\&$$
$$включают (0, 0, дисципліни, сучасні, планів, навчальних, регулярно)]\&$$
$$[впливає (0, 0, статус, 0, закладу, вищого, 0)\&$$
$$впливає (0, 0, статус, 0, закладу, навчального, 0)].$$

Для речення російської мови «*Новые современные дисциплины регулярно включают в учебные планы, что влияет на статус высшего учебного заведения*» буде побудована логіко-лінгвістична модель типу 15, що матиме вигляд:

$$L^S = [p_1 (0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_1 (0, 0, y_1, q_{12}, z_1, r_1, h_1)] \rightarrow$$
$$[p'_1 (0, 0, y'_1, 0, z_1, r'_{11}, 0) \& p'_1 (0, 0, y'_1, 0, z_1, r'_{12}, 0)],$$
$$L^S = [включают (0, 0, дисциплины, новые, планы, учебные, регулярно) \&$$
$$включают (0, 0, дисциплины, современные, планы, учебные, регулярно)] \&$$

[влияет (0, 0, статус, 0, заведения, высшего, 0)&
влияет (0, 0, статус, 0, заведения, учебного, 0)].

Для речення англійської мови «*New modern disciplines are regularly included in curricula, which affects the status of higher education institution*» буде побудована логіко-лінгвістична модель типу 15, що матиме вигляд:

$$L^S = [p_{11} \& p_{12} (0, 0, y_1, q_{11}, z_1, r_1, h_1) \& p_{11} \& p_{12} (0, 0, y_1, q_{12}, z_1, r_1, h_1)] \rightarrow [p'_1(0, 0, y'_1, 0, z_1, r'_{11}, 0) \& p'_1(0, 0, y'_1, 0, z_1, r'_{12}, 0)],$$

$$L^S = [are\&included (0, 0, disciplines, new, curricula, 0, regularly)\& are\&included (0, 0, disciplines, modern, curricula, 0, regularly)] \rightarrow [affects (0, 0, status, 0, institution, higher, 0)\& affects (0, 0, status, 0, institution, education, 0)].$$

Формули (1.1) – (1.4) інтерпретуються як матриця L^S , розмірністю $(v^S \times 8)$, у якій кількість рядків v^S варіюється за рахунок потужності множин елементів логіко-лінгвістичних моделей ($\mu=1, v^S$) та обчислюється за допомогою виразів (1.5) – (1.8):

$$L^S = \begin{pmatrix} p^{(1)} & x^{(1)} & g^{(1)} & y^{(1)} & q^{(1)} & z^{(1)} & r^{(1)} & h^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p^{(\lambda)} & x^{(\lambda)} & g^{(\lambda)} & y^{(\lambda)} & q^{(\lambda)} & z^{(\lambda)} & r^{(\lambda)} & h^{(\lambda)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p^{(v^S)} & x^{(v^S)} & g^{(v^S)} & y^{(v^S)} & q^{(v^S)} & z^{(v^S)} & r^{(v^S)} & h^{(v^S)} \end{pmatrix}.$$

Також проведені експерименти з побудови окремих форм одного і того самого речення, записаного українською, російською та англійською флективними мовами. Дослідження показали, що тип окремих форм логіко-лінгвістичних моделей для різномовних речень природної мови з однаковим змістом співпадає. Це дає підставу для подальших досліджень у сфері машинного перекладу.

1.3 Логіко-лінгвістична модель текстового документу

Враховуючи досвід створення існуючих формальних моделей представлення знань [43 - 46] та моделей пошуку [47 – 49], для формального опису текстового документу запропоновано формальну модель, що містить лінгвістичну та семантико-синтаксичну складову [50].

Лінгвістична складова розуміє під собою абстрактну модель, що формується на основі правил побудови зв'язного тексту та його змістовних категорій, які виражаються цими правилами.

Синтаксис в даній роботі передбачає формування логіко-лінгвістичних моделей речень природної мови на основі формальних правил побудови словосполучень природної мови та принципів виявлення засобів вираження змістовних відношень у цих словосполученнях [31, 33].

Семантика в даній роботі – це правила та принципи об'єднання логіко-лінгвістичних моделей речень природної мови для визначення цілісної структури пов'язаних понять у вигляді формальної моделі текстового документу.

Текст – це завершена з точки зору його автора, проте відкрита для множини інтерпретацій в змістовному та інтенціональному плані, лінійна послідовність мовних знаків, виражених графічно, семантико-змістовна взаємодія яких створює композиційну єдність, підтримуючи лексико-граматичні відношення між окремими елементами створеної таким чином структури [51].

У граматичному аспекті зв'язність тексту визначається законами узгодження, правилами побудови висловлювань з використанням морфологічних та синтаксичних засобів мови. Тому у даній монографії електронний текстовий документ розглядається як множина взаємопов'язаних логіко-лінгвістичних моделей речень природної мови, що входять до складу тексту, тобто його семантико-синтаксична складова [52].

Логіко-лінгвістична модель текстового документу – це абстрактна модель, яка об'єднує в собі основні властивості тексту та його складових частин, відображає основні взаємозв'язки між структурними компонентами, представляє собою впорядковану четвірку та масив логіко-лінгвістичних моделей речень природної мови, що входять до тексту [53].

Лінгвістична складова формального опису тексту:

$$t = \langle CQ, F, B, A \rangle, \quad (1.24)$$

Де $t \in T$ – конкретний електронний текст із всієї множини текстів T ;

$CQ \in C$ – тип тексту, що може приймати одне із значень елементів множини типів текстів $C = \{1, 2, 3, 4\}$, що розглядаються (1 – науковий, 2 – публіцистичний, 3 – художній, 4 – офіційно-діловий);

$F = \{f_1, \dots, f_j, \dots, f_m\}$ – множина складних синтаксичних частин тексту, $j = \overline{1, m}$, m – кількість складних синтаксичних частин;

B – текстова база, що складається з набору ключових слів тексту та взаємопов'язаних пропозицій, і яку можна представити у вигляді трійки: $B = \langle K, SJ, D \rangle$, K – множина ключових слів тексту; SJ – множина ключових словосполучень тексту; D – множина пропозицій;

$A = \{a_1, \dots, a_k, \dots, a_q\}$ – множина абзаців тексту, $k = \overline{1, q}$, q – кількість абзаців. Кожен абзац у свою чергу описується трійкою: $a_k = \langle H, Y, R \rangle$, $H = \{1, 2\}$ – множина типів зв'язків між реченнями (ланцюговий чи паралельний); $Y = \{1, 2, 3, 4, 5\}$ – множина типів тематичних прогресій, що вжиті у абзаці $a_k \in A$; $R = \{1, 2, 3, 4, 5, 6\}$ – множина рематичних доміант у абзаці $a_k \in A$ [54, 55].

Довільний текстовий фрагмент складається з пов'язаних між собою логічними, змістовними зв'язками речень природної мови. Створення семантико-синтаксичної складової логіко-лінгвістичної моделі текстового документу базується на створенні масивів логіко-лінгвістичних моделей речень природної мови. Тому формально текстовий фрагмент можна представити за допомогою одновимірного масиву логіко-лінгвістичних моделей $t = \{L^{\delta_1}, L^{\delta_2}, \dots, L^{\delta_\delta}, \dots, L^{\delta_{N(t)}}\}$, де $\delta = \overline{1, N(t)}$, $N(t)$ – загальна кількість речень у текстовому фрагменті t .

Тоді **семантико-синтаксична складова** формального опису тексту:

$$t' = \bigwedge_{\delta=1}^{N(t)} L^{\delta_\delta}, \quad (1.25)$$

де L^{δ} – логіко-лінгвістична модель речення S_{δ} , $\delta = \overline{1, N(t)}$, сформована за формулами (1.1) – (1.4).

Модель (1.24) – (1.25) містить вичерпну інформацію про текст та зв'язки у ньому. Побудова такої логіко-лінгвістичної моделі для довільного типу тексту дає змогу перейти до аналізу текстової інформації, порівняння текстів за змістом, пошуку протиріч та збігів.

Алгоритм побудови логіко-лінгвістичної моделі тексту складається з таких кроків (рис. 1.1).

1. Членування тексту. Цей крок відповідає за розбиття усієї текстової інформації (електронного документу), на декілька рівнів. Членування тексту – функція загального композиційного плану документу, основну роль в якому відіграють розмір частин та змістовно-фактуальна інформація [56].

За Гальперіном можна виділити два типи членування текстів: об'ємно-прагматичне та контекстно-варіативне [57]. В основі першого членування лежить кількісний параметр (поділ на розділи, частини, абзаци і т.д.).

Другий тип членування дає можливість встановити типи зв'язків між складними синтаксичними частинами, тобто побудувати схематичну структуру тексту, що аналізується (у тексті виділяють три основні частини: зачин, розгортання думки і висновок). Залежно від типу зв'язного тексту таких складових може бути більше, тому для конкретної текстової бази буде свій набір складних синтаксичних частин F .

На даному етапі відбувається об'ємно-прагматичне членування, тобто побудова схематичної структури документу: текст розбивається на розділи, частини, абзаци, а абзаци в свою чергу – на речення. Членування тексту – це суто технічний крок, на якому не враховуються синтаксичні та семантичні зв'язки.

Після виконання етапу членування тексту t буде отримано множину складних синтаксичних частин $F = \{f_1, \dots, f_i, \dots, f_m\}$, $j = \overline{1, m}$ (розділів, підрозділів та ін. за наявності), множину абзаців тексту $A = \{a_1, \dots, a_k, \dots, a_q\}$, $k = \overline{1, q}$, q – кількість абзаців та

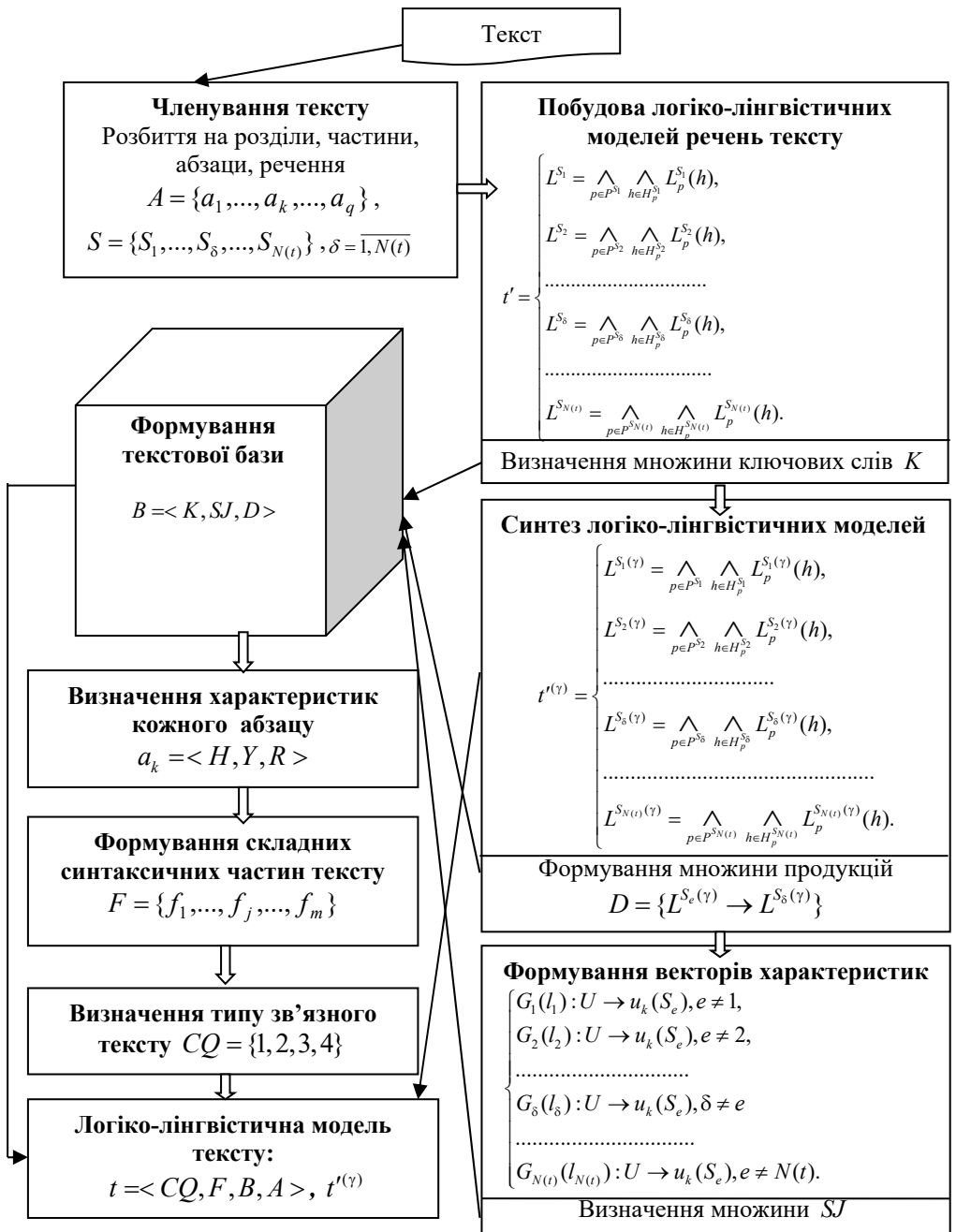


Рис. 1.1 – Алгоритм побудови логіко-лінгвістичної моделі тексту

множину речень природної мови $S = \{S_1, \dots, S_\delta, \dots, S_{N(t)}\}$, де $\delta = \overline{1, N(t)}$.

2. Побудова логіко-лінгвістичних моделей речень тексту.

На даному етапі до тексту t застосовується метод автоматизованої побудови логіко-лінгвістичних моделей речень природної мови, в результаті чого кожне речення тексту перетворюється в логічну формулу (1.1) – (1.4).

Виконання даного етапу алгоритму забезпечує формування семантико-синтаксичної складової формального опису тексту (1.25), тобто:

$$t' = \begin{cases} L^{S_1} = \bigwedge_{p \in P^{S_1}} \bigwedge_{h \in H_p^{S_1}} L_p^{S_1}(h), \\ L^{S_2} = \bigwedge_{p \in P^{S_2}} \bigwedge_{h \in H_p^{S_2}} L_p^{S_2}(h), \\ \dots \\ L^{S_\delta} = \bigwedge_{p \in P^{S_\delta}} \bigwedge_{h \in H_p^{S_\delta}} L_p^{S_\delta}(h), \\ \dots \\ L^{S_{N(t)}} = \bigwedge_{p \in P^{S_{N(t)}}} \bigwedge_{h \in H_p^{S_{N(t)}}} L_p^{S_{N(t)}}(h), \end{cases} \quad (1.26)$$

а простий предикат (1.4) для кожного речення природної мови має вигляд: $L_{p^{(\lambda)}}^{S_\delta}(x_\delta^{(\lambda)}, g_\delta^{(\lambda)}, y_\delta^{(\lambda)}, q_\delta^{(\lambda)}, z_\delta^{(\lambda)}, r_\delta^{(\lambda)}, h_\delta^{(\lambda)})$, де δ – номер речення природної мови у тексті; λ – рівень вкладеності простого речення у складному реченні S_δ , тобто:

$$t' = \begin{cases} L_{p^{(\lambda)}}^{S_1}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}) \\ L_{p^{(\lambda)}}^{S_2}(x_2^{(\lambda)}, g_2^{(\lambda)}, y_2^{(\lambda)}, q_2^{(\lambda)}, z_2^{(\lambda)}, r_2^{(\lambda)}, h_2^{(\lambda)}) \\ \dots \\ L_{p^{(\lambda)}}^{S_\delta}(x_\delta^{(\lambda)}, g_\delta^{(\lambda)}, y_\delta^{(\lambda)}, q_\delta^{(\lambda)}, z_\delta^{(\lambda)}, r_\delta^{(\lambda)}, h_\delta^{(\lambda)}), \\ \dots \\ L_{p^{(\lambda)}}^{S_{N(t)}}(x_{N(t)}^{(\lambda)}, g_{N(t)}^{(\lambda)}, y_{N(t)}^{(\lambda)}, q_{N(t)}^{(\lambda)}, z_{N(t)}^{(\lambda)}, r_{N(t)}^{(\lambda)}, h_{N(t)}^{(\lambda)}). \end{cases} \quad (1.27)$$

3. Синтез логіко-лінгвістичних моделей. На третьому кроці алгоритму побудови логіко-лінгвістичної моделі текстового документу здійснюється об'єднання та заміна структурних компонентів логіко-лінгвістичних моделей (1.26), отриманих на попередньому етапі алгоритму. Це відбувається на основі виявлення способів логічного зв'язку між реченнями природної мови.

На основі використання принципів та правил синтезу, описаних у розділі 2 цієї монографії, логіко-лінгвістичні моделі (1.27) набувають такого вигляду:

$$t^{(\gamma)} = \left\{ \begin{array}{l} L^{S_1(\gamma)} = \bigwedge_{p \in P^{S_1}} \bigwedge_{h \in H_p^{S_1}} L_p^{S_1(\gamma)}(h), \\ L^{S_2(\gamma)} = \bigwedge_{p \in P^{S_2}} \bigwedge_{h \in H_p^{S_2}} L_p^{S_2(\gamma)}(h), \\ \dots \dots \dots \\ L^{S_\delta(\gamma)} = \bigwedge_{p \in P^{S_\delta}} \bigwedge_{h \in H_p^{S_\delta}} L_p^{S_\delta(\gamma)}(h), \\ \dots \dots \dots \\ L^{S_{N(t)}(\gamma)} = \bigwedge_{p \in P^{S_{N(t)}}} \bigwedge_{h \in H_p^{S_{N(t)}}} L_p^{S_{N(t)}(\gamma)}(h), \end{array} \right. \quad (1.28)$$

а, окрім заміни в формулах (1.27), кожній логіко-лінгвістичній моделі $L^{S_\delta(\gamma)}$ приписується масив характеристик l_δ , кожна з яких відповідає певній компоненті логіко-лінгвістичної моделі речення S_δ , пов'язаного з ним за змістом:

$$\left\{ \begin{array}{l} G_1(l_1): U \rightarrow u_k(S_e), e \neq 1, e = \overline{1, N(t)}, \\ G_2(l_2): U \rightarrow u_k(S_e), e \neq 2, \\ \dots \dots \dots \\ G_\delta(l_\delta): U \rightarrow u_k(S_e), \delta \neq e \\ \dots \dots \dots \\ G_{N(t)}(l_{N(t)}): U \rightarrow u_k(S_e), e \neq N(t). \end{array} \right. \quad (1.29)$$

4. Формування текстової бази. Завдяки синтезу логіко-лінгвістичних моделей формується текстова база

4.1. Множина ключових слів тексту K формується з елементів масивів l_{δ} , $\delta = \overline{1, N(t)}$, тобто кожен елемент логіко-лінгвістичної моделі, що був замінений внаслідок використання принципів та правил синтезу (1.29), буде входити до множини ключових слів: $u_k(S_e) \subseteq K$, де $e = \overline{1, N(t)}$, при $e \neq \delta$.

4.2. Множина ключових словосполучень тексту SJ формується формальними правилами утворення словосполучень тексту [55], до яких входять ключові слова із утвореної множини K . Тобто, якщо слово $H_l = u_k(S_{\delta})$, $l = \overline{1, n}$, де n – кількість слів у реченні S_{δ} , утворює словосполучення зі словом H_k , $k = \overline{1, n}$, $k \neq l$ цього ж речення: $SJ_{lk} = H_l \cup H_k$, то $SJ_{lk} \in SJ$.

4.3. Для формування множини пропозицій D застосовується інтерпретатор продукцій, який працює циклічно. Основними даними для його роботи є отримані внаслідок синтезу логіко-лінгвістичних моделей, масиви характеристик (1.29). Кожен елемент масиву характеристик l_{δ} речення природної мови S_{δ} інтерпретує зв'язок простого предиката $L^{S_{\delta}(\gamma)}$ із формули (1.28) речення S_{δ} та простого предиката $L^{S_e(\gamma)}$ речення природної мови S_e , при цьому $e \neq \delta$.

Тобто, в результаті роботи інтерпретатора продукцій шукаються такі моделі $L^{S_e(\gamma)}$, які за змістом передують або з яких випливають моделі $L^{S_{\delta}(\gamma)}$:

$$L^{S_e(\gamma)} \rightarrow L^{S_{\delta}(\gamma)}.$$

Етап побудови текстової бази дає змогу сформувати змістовні зв'язки між реченнями тексту, незалежно від того, до якого абзацу чи складної синтаксичної частини вони належать.

5. Визначення характеристик кожного з абзаців. Визначення типів зв'язків між реченнями в кожному абзаці є одним із найважливіших етапів проведення лінгвістичного аналізу тексту [55] і відбувається на основі принципів та правил синтезу логіко-лінгвістичних моделей.

5.1. Визначення типу зв'язку в абзаці, тобто першого параметра із трійки характеристик $H = \{1, 2\}$ із множини типів зв'язків між реченнями (ланцюговий чи паралельний) [58].

5.2. Визначення типу тематичної прогресії із множини $Y = \{1, 2, 3, 4, 5\}$, $a_k \in A$ (1 – проста лінійна прогресія, 2 – прогресія з наскрізною темою, 3 – прогресія з похідними темами, 4 – прогресія з розщепленою темою, 5 – прогресія з тематичним стрибком);

5.3. Визначення рематичної домінанти $R = \{1, 2, 3, 4, 5, 6\}$ – множина рематичних домінант у абзаци $a_k \in A$ (1 – предметна, 2 – статальна, 3 – динамічна, 4 – якісна, 5 – імпресивна, 6 – комбінована) [54, 55].

Визначення рематичних домінант відбувається шляхом узагальнення результатів синтезу (отриманої множини продукцій та масивів характеристик кожного речення), а також на основі отриманих з бази даних характеристик ключових слів тексту із множини K .

6. Визначення типу тексту CQ . Стилістичні, семантичні та синтаксичні особливості кожного тексту формують глобальний зв'язок. Тому визначення типу тексту, як і рематичних домінант, відбувається на основі отриманих з бази даних характеристик ключових слів тексту K , а також статистичної оцінки типу отриманих характеристик ключових слів до типу характеристик, що властиві конкретному типу тексту. В базі даних для кожного типу на основі мовних та когнітивних особливостей [59] побудовано текстову базу та ситуаційну модель [55], на основі чого і відбувається оцінка.

Приведемо приклад формування семантико-синтаксичної складової логіко-лінгвістичної моделі текстового фрагменту українською мовою: *"Цікава лекція заохоче лінєвих і працьовитих студентів та практикантів до подальшого навчання. Від її змісту, динаміки викладання та демонстраційних матеріалів залежить те, чи отримає знання студент. Тому викладачі повинні ретельно готуватися до лекцій, постійно удосконалювати свої знання та стежити за сучасними тенденціями у науці. Провідні висококваліфіковані викладачі регулярно включають нові сучасні дисципліни до навчальних планів"*.

Логіко-лінгвістична модель текстового фрагменту (1.25) набуває вигляду:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$$\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \&$$

$$p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0),$$

$$L^{S_2} = [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \&$$

$$p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0),$$

$$L^{S_3} = p_{31} \& p_{32}(x_3, 0, x_1, 0, 0, 0, h_{31}) \&$$

$$p_{31} \& p_{33}(x_3, 0, y'_2, q_{31}, 0, 0, h_{32}) \& p_{31} \& p_{34}(x_3, 0, y_{33}, q_{32}, z_3, 0, 0),$$

$$L^{S_4} = p_4(x_3, g_{41}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{41}, y_4, q_{32}, z_4, r_4, h_4) \&$$

$$p_4(x_3, g_{42}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{42}, y_4, q_{32}, z_4, r_4, h_4)\}.$$

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$\{L^{S_1} = \text{заохочує (лекція, цікава, студентів, лінивих, навчання, подальшого, 0) \& заохочує (лекція, цікава, практикантів, лінивих, навчання, подальшого, 0) \&$

$\text{заохочує (лекція, цікава, студентів, працюючих, навчання, подальшого, 0) \&$

$\text{заохочує (лекція, цікава, практикантів, працюючих, навчання, подальшого, 0),}$

$L^{S_2} = [\text{залежить (те, 0, змісту, лекції, 0,0,0) \&}$

$\text{залежить (те, 0, динаміки, 0, викладання, 0,0) \&}$

$\text{залежить (те, 0, матеріалів, демонстраційних, 0,0,0)] \rightarrow$

$\text{отримає (студент, 0, знання, 0,0,0,0),}$

$L^{S_3} = \text{повинні \& готуватися (викладачі, 0, лекцій, 0,0,0, ретельно) \&}$

$\text{повинні \& удосконалювати (викладачі, 0, знання, свої, 0,0, постійно) \&}$

$\text{повинні \& стежити (викладачі, 0, тенденціями, сучасними, науці, 0, 0),}$

$L^{S_4} = \text{включають (викладачі, провідні, дисципліни, нові, планів, навчальних, регулярно) \&}$

$\text{включають (викладачі, провідні, дисципліни, сучасні, планів, навчальних, регулярно) \&}$

$\text{включають (викладачі, висококваліфіковані, дисципліни, нові, планів, навчальних, регулярно) \&}$

включают (викладачі, висококваліфіковані, дисципліни, сучасні, планів, навчальних, регулярно).

Для цього ж текстового фрагменту російською мовою "Интересная лекция стимулирует ленивых и трудолюбивых студентов и практикантов к последующему обучению. От ее содержания, динамики преподавания и демонстрационных материалов зависит то, получит ли знания студент. Поэтому преподаватели должны тщательно готовиться к лекциям, постоянно совершенствовать свои знания и следить за современными тенденциями в науке. Ведущие высококвалифицированные преподаватели регулярно включают новые современные дисциплины в учебные планы" семантико-синтаксична складова матиме вигляд:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$$\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \&$$

$$p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0),$$

$$L^{S_2} = [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \&$$

$$p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0),$$

$$L^{S_3} = p_{31} \& p_{32}(x_3, 0, x_1, 0, 0, 0, h_{31}) \&$$

$$p_{31} \& p_{33}(x_3, 0, y'_2, q_{31}, 0, 0, h_{32}) \& p_{31} \& p_{34}(x_3, 0, y_{33}, q_{32}, z_3, 0, 0),$$

$$L^{S_4} = p_4(x_3, g_{41}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{41}, y_4, q_{32}, z_4, r_4, h_4) \&$$

$$p_4(x_3, g_{42}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{42}, y_4, q_{32}, z_4, r_4, h_4)\}.$$

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$\{L^{S_1}$ = стимулирует (лекция, интересная, студентов, ленивых, обучению, последующему, 0) &

стимулирует (лекция, интересная, практикантов, ленивых, обучению, последующему, 0) &

стимулирует (лекция, интересная, студентов, трудолюбивых, обучению, последующему, 0) &

стимулирует (лекция, интересная, практикантов, трудолюбивых, обучению, последующему, 0),

L^{S_2} = [зависит (то, 0, содержания, лекции, 0,0,0) &

зависит (то, 0, динамики, преподавания, 0,0) &

зависит (то, 0, материалов, демонстрационных, 0,0,0)] →

получит (студент, 0, знания, 0,0,0,0),

$L^3 = \text{должны\&готовиться (преподаватели, 0, лекциям, 0,0,0,}$
тщательно)\&

$\text{должны\&усовершенствовать (преподаватели, 0, знания, свои,0,0,}$
тщательно)\&

$\text{должны\&следить (преподаватели, 0, тенденциями,}$
современными, науке, 0, 0),

$L^4 = \text{включают (преподаватели, ведущие, дисциплины, новые,}$
планы, учебные, регулярно) \&

$\text{включают (преподаватели, ведущие, дисциплины, современные,}$
планы, учебные, регулярно) \&

$\text{включают (преподаватели, высококвалифицированные,}$
дисциплины, новые, планы, учебные, регулярно) \&

$\text{включают (преподаватели, высококвалифицированные,}$
дисциплины, современные, планы, учебные, регулярно)\}.

Для цього ж текстового фрагменту англійською мовою "An interesting lecture stimulates lazy and industrious students and trainees for subsequent training. From its content, the dynamics of teaching and demonstration materials, it depends on whether the student will gain knowledge. Therefore, teachers should carefully prepare for lectures, constantly improve their knowledge and follow modern trends in science. Leading teachers highly qualified regularly include new modern disciplines in curricula" семантико-синтаксична складова матиме вигляд:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$$\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \&$$

$$p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0),$$

$$L^{S_2} = [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \&$$

$$p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_{21} \& p'_{22}(y_{11}, 0, y'_2, 0, 0, 0, h'_2),$$

$$L^{S_3} = p_{31} \& p_{32}(x_3, 0, x_1, 0, 0, 0, h_{31}) \&$$

$$p_{31} \& p_{33}(x_3, 0, y'_2, q_{31}, 0, 0, h_{32}) \& p_{31} \& p_{34}(x_3, 0, y_{33}, q_{32}, z_3, 0, 0),$$

$$L^{S_4} = p_4(x_3, g_{41}, y_4, q_{41}, z_4, 0, h_4) \& p_4(x_3, g_{41}, y_4, q_{32}, z_4, 0, h_4) \&$$

$$p_4(x_3, g_{42} \& q_{43}, y_4, q_{41}, z_4, 0, h_4) \& p_4(x_3, g_{42} \& q_{43}, y_4, q_{32}, z_4, 0, h_4)\}.$$

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} =$$

$\{L^{S_1} = \text{stimulates (lecture, interesting, students, lazy, training,}$
subsequent, 0) \&

stimulates (lecture, interesting, students, industrious, training, subsequent, 0) &

{L^{S1} = stimulates (lecture, interesting, trainees, lazy, training, subsequent, 0) &

stimulates (lecture, interesting, trainees, industrious, training, subsequent, 0),

L^{S2} = [depends (it, 0, content, lecture's, 0,0,0)&

depends (it, 0, dynamics, 0, teaching,0,0)&

depends (it, 0, materials, demonstration, 0,0,0)]→

will&gain (student, 0, knowledge,0,0,0, whether),

L^{S3} = should&prepare (teachers, 0, lectures, 0,0,0, carefully)&

should&improve (teachers, 0, knowledge, their,0,0, constantly)&

should&follow (teachers, 0, trends, modern, science,0, 0),

L^{S4} = include (teachers, leading, disciplines, new, curricula, 0, regularly)&

include (teachers, leading, disciplines, modern, curricula, 0, regularly)&

include (teachers, leading, disciplines, new, curricula, 0, regularly)&

include (teachers, highly&qualified, disciplines, modern, curricula, 0, regularly)}.

Із продемонстрованого прикладу побудови логіко-лінгвістичних моделей текстових фрагментів, що однакові за змістом, та написані на різних флективних мовах (українській, російській, англійській), видно, що їх семантико-синтаксичні складові однакові, а логіко-лінгвістична модель (1.25) – універсальна.

Отже, при побудові масиву логіко-лінгвістичних моделей застосовуються основні принципи їх синтезу, що дозволяються здійснити заміну складових логіко-лінгвістичних моделей речень природної мови, які пов'язані між собою за змістом.

Структуризація текстової інформації шляхом логіко-лінгвістичного моделювання передбачає виявлення і пошук логічних зв'язків, починаючи з найнижчого рівня – побудови словосполучень та закінчуючи пошуком логічних зв'язків у текстах.

II. СИНТЕЗ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ

*"Мета математичної строгості
полягає у тому, щоб санкціонувати
та узаконити завоювання інтуїції."*

Жак Адамар

Логічні зв'язки у тексті можна виявити як шляхом аналізу складного речення, так і завдяки лексико-граматичним механізмам (різноманітним мовним відношенням між реченнями), які забезпечують формально-граматичну зв'язність тексту, тобто когезію [60].

Когезія – це семантико-когнітивна зв'язність в її природно-наслідковому, часовому та преференційному аспектах [51]. Лексичні засоби когезії виконують функцію просторової організації тексту, функцію одночасного зв'язку фрагментів тексту та текстоутворюючу функцію. Формалізація правил використання локальної та глобальної зв'язності тестів дає можливість автоматизувати процес аналізу текстових документів за змістом.

У сучасній лінгвістиці текст представляє собою когерентний дискурс, який внаслідок взаємодії суб'єктів наповнює його контентом. Семантико-синтаксичний аналіз тексту передбачає визначення ознак тексту як цілого та виявлення засобів його локальної і глобальної зв'язності [61]. У граматичному аспекті зв'язність тексту визначається законами узгодження, правилами побудови висловлювань з використанням морфологічних та синтаксичних засобів мови. У прагматичному аспекті зв'язність обумовлена загальною комунікативною функцією тексту, реалізується у суб'єктивній організації тексту, в системі просторових та часових характеристик, що пронизують текст з початку і до кінця [62]. Між структурним та семантичним планами тексту немає тотожності, але є паралелізм, який проявляється у зв'язках. Так як довільна синтаксична структура існує для того, щоб передати деякий зміст, то призначення структурного плану – зробити можливим вираження думки – об'єктів семантичного плану [63]. Виходячи із цього твердження, можна дати визначення синтезу логіко-лінгвістичних моделей.

Синтез логіко-лінгвістичних моделей – це об'єднання та заміна їх структурних компонентів на основі виявлення способів логічного зв'язку між реченнями природної мови з метою простеження єдиної тематичної прогресії текстового документу [64].

Синтез логіко-лінгвістичних моделей передбачає отримання деякої канонічної форми, яка містить у собі елементарні відношення, що носять тотожний зміст і дозволяють однозначно інтерпретувати речення природної мови довільної структури.

Побудова формальної логіко-лінгвістичної моделі (1.24) – (1.25) електронного текстового документу можлива завдяки синтезу логіко-лінгвістичних моделей речень природної мови. Для його здійснення розроблено основні принципи, формальні правила, а також умови визначення типів тематичного зв'язку у текстових фрагментах.

2.1 Основні принципи синтезу логіко-лінгвістичних моделей речень природної мови

З метою отримання формальної моделі, за допомогою якої можна описати фрагмент тексту та врахувати зміст речень природної мови, що його складають, використовуються основні принципи синтезу логіко-лінгвістичних моделей.

Нехай є два речення природної мови S_1 та S_2 , кожному з яких ставиться у відповідність логіко-лінгвістична модель, побудована згідно з формулами (1.1) – (1.4):

$$L^{S_1} = \bigwedge_{\lambda=1}^{v^{S_1}} L_{p^{(\lambda)}}^{S_1}(h), \quad (2.1)$$

$$L^{S_2} = \bigwedge_{\mu=1}^{v^{S_2}} L_{p^{(\mu)}}^{S_1}(h), \quad (2.2)$$

де $L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)})$ – простий предикат, що описує частину речення S_1 , яка відображає закінчений зміст;

$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)})$ – простий предикат, що описує частину речення S_2 , яка відображає закінчений зміст;

$\lambda = 1, v^{S_1}, v^{S_1}$ – кількість частин речення S_1 , які мають закінчений зміст, і обчислюється за формулами (1.5) – (1.8);

$\mu = 1, v^{S_2}, v^{S_2}$ – кількість частин речення S_2 , які мають закінчений зміст, і обчислюється за формулами (1.5) – (1.8).

Тоді існує такий оператор $G(l_\delta)$, який ставить у відповідність елементам масиву характеристик l_δ кожного речення природної мови $S_\delta, \delta = \overline{1, N(t)}$ певну компоненту логіко-лінгвістичної моделі $u_k(S_e) \in U, k = \overline{1, 8}$ (1 – відношення, 2 – суб'єкт, 3 – характеристика суб'єкта, 4 – об'єкт, 5 – характеристика об'єкта, 6 – предмет відношення, 7 – характеристика предмета відношення, 8 – характеристика відношення) речення природної мови $S_e, e = \overline{1, N(t)}, e \neq \delta$, пов'язаного за змістом з реченням S_δ :

$$G(l_\delta): U \rightarrow u_k(S_e),$$

де U – множина можливих значень компонент логіко-лінгвістичної моделі (1.1) – (1.4);

l_δ – одновимірний масив з кількістю елементів $N(t) - 1$.

Принцип 1. Принцип відповідає ситуації, коли для зв'язку речень у тексті використовується семантичне повторення (рекуренція) – засіб, який дає змогу простежити ряд речень з однаковим змістом (синтаксична структура таких речень буде однаковою, проте відношення, суб'єкт або об'єкт можуть бути замінені синонімами) [65, 66]. Семантичне повторення може бути здійснене декількома шляхами.

Повне тотожне повторення – це найбільш простий механізм зв'язку, що полягає у повторенні однакових слів або слівформ, які мають один корінь.

Тематичне повторення – слова у складі тексту виявляють спільність лексичних значень, позначають різні сторони, компоненти, складові частини елементів однієї ситуації. Слова однієї тематичної групи утворюють всередині тексту парадигму, яка пов'язує різні частини тексту.

Експресивне повторення – вживання слів, які формують загальний план оцінки.

Синонімічне повторення – вживання синонімів для забезпечення міжфразового зв'язку, об'єднання змісту різних речень.

Повторення різних стилістичних варіантів одного і того ж слова.

Зв'язність через відношення частина-ціле шляхом родово-видової заміни. Тобто, слова, не будучи синонімами, стають позначеннями одного і того ж предмета, – контекстуальними синонімами.

Зв'язність через колокацію або антонімічне повторення – дає змогу підкреслити протиріччя, конфліктність подій, станів, створити антитезу образів. Такі відношення можуть бути виявлені шляхом статистики частого вживання слів.

Порядок вживання речень у текстовому фрагменті при виявленні першого принципу не береться до уваги.

Твердження 1.1. Якщо відношення двох речень природної мови $p_1^{(\lambda)} \equiv p_2^{(\mu)}$ тотожні або $\widehat{p}_1^{(\lambda)} \equiv \widehat{p}_2^{(\mu)}$, де $\widehat{p}_1^{(\lambda)} \in \widehat{R}_p$ і $\widehat{p}_2^{(\mu)} \in \widehat{R}_p$, відносяться до одного і того ж ряду спільнокореневих слів $\widehat{R}_p \in R$ із множини R , або $p_1^{(\lambda)} \in W_p$ і $p_2^{(\mu)} \in W_p$ відносяться до одного і того ж синонімічного ряду $W_q \subseteq W$ із множини W , то необхідно здійснити синтез, зробивши заміну $p_1^{(\lambda)} \equiv p_2^{(\mu)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_1^{(\lambda)}(x_2^{(\mu)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}).$$

При цьому елемент $\alpha_{a_1}(l_1)$ ($a_1 = \overline{1, b_1}$, b_1 – кількість елементів) масиву характеристик l_1 речення природної мови S_1 дорівнює значенню $u_k(S_e) = u_1(S_2) = p_2^{(\mu)}$, а елемент $\alpha_{a_2}(l_2)$ ($a_2 = \overline{1, b_2}$, b_2 – кількість елементів) масиву характеристик l_2 речення природної мови S_2 – значенню $u_k(S_e) = u_1(S_1) = p_1^{(\lambda)}$.

Розглянемо текстовий фрагмент, що складається із двох речень: "Хто говорить надто тихо, той справляє враження

людини, яка не вірить у свої сили. Той же, хто говорить надто голосно, справляє враження агресивної людини" [67].

Логіко-лінгвістична модель фрагмента тексту українською мовою матиме вигляд:

$$L^{S_1} = p_{11}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$$p_1'(x_1', g_1', y_1', q_1', z_1', r_1', h_1') = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$$

$$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& p_1'(z_1, 0, y_1', q_1', 0, 0, 0),$$

$$L^{S_2} = p_{21}(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \& p_{22}(x_2, g_2, y_2, q_2, z_2, r_2, h_2) =$$

$$p_{21}(x_2, 0, 0, 0, 0, 0, h_{21} \& h_{22}) \& p_{22}(x_2, 0, y_2, 0, z_2, r_2, 0),$$

L^{S_1} = говорить (хто, 0, 0, 0, 0, 0, надто&тихо)&
справляє (той, 0, враження, 0, людини, 0, 0)&
¬ вірить (людина, 0, сили, свої, 0, 0, 0),

L^{S_2} = говорить (хто, 0, 0, 0, 0, 0, надто&голосно)&
справляє (той, 0, враження, 0, людини, агресивної, 0).

У фрагменті використано повне тотожне повторення, за рахунок чого у логіко-лінгвістичних моделях виконано заміну відношень за твердженням 1.1, суб'єктів та об'єктів відношення:

$$L^{S_1} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$$

$$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& p_1'(z_1, 0, y_1', q_1', 0, 0, 0),$$

$$L^{S_2} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{22}) \& p_{12}(x_1, 0, y_1, 0, z_1, r_2, 0).$$

Для цього ж текстового фрагменту російською мовою "Кто говорит слишком тихо, тот создает впечатление человека, который не верит в свои силы. Тот же, кто говорит слишком громко, создает впечатление агрессивного человека" маємо:

$$L^{S_1} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$$

$$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& \neg p_1'(z_1, 0, y_1', q_1', 0, 0, 0),$$

$$L^{S_2} = p_{21}(x_2, 0, 0, 0, 0, 0, h_{21} \& h_{22}) \& p_{22}(x_2, 0, y_2, 0, z_2, r_2, 0),$$

L^{S_1} = говорит (кто, 0, 0, 0, 0, 0, слишком&тихо)&
создает (тот, 0, впечатление, 0, человека, 0, 0)&
¬ верит (человек, 0, силы, свои, 0, 0, 0),

L^{S_2} = говорит (кто, 0, 0, 0, 0, 0, слишком&громко)&
создает (тот, 0, впечатление, 0, человека, агрессивного, 0).

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою "*Whoever speaks too quietly creates the impression of a person who does not believe in his own strength. The one who speaks too loudly, gives the impression of an aggressive person*" матиме вигляд:

$$L^{S_1} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$$

$$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& \neg p'_1(z_1, 0, y'_1, q'_{11} \& q'_{12}, 0, 0, 0),$$

$$L^{S_2} = p_{21}(x_2, 0, 0, 0, 0, 0, h_{21} \& h_{22}) \& p_{22}(x_2, 0, y_2, 0, z_2, r_2, 0),$$

$L^{S_1} = \textit{speaks (whoever, 0, 0, 0, 0, 0, too \& quietly)} \&$
 $\textit{creates (who, 0, impression, 0, person, 0, 0)} \&$
 $\neg \textit{believe (person, 0, strength, his\&own, 0, 0, 0)},$

$L^{S_2} = \textit{speaks (who, 0, 0, 0, 0, 0, too \& loudly)} \&$
 $\textit{gives (one, 0, impression, 0, person, aggressive, 0)}.$

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за першим принципом (твердження 1.1), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$$

$$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& \neg p'_1(z_1, 0, y'_1, q'_{11} \& q'_{12}, 0, 0, 0),$$

$$L^{S_2} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{22}) \& p_{12}(x_1, 0, y_1, 0, z_1, r_2, 0).$$

Твердження 1.2. Якщо суб'єкти відношення двох речень природної мови тотожні $x_1^{(\lambda)} \equiv x_2^{(\mu)}$ або $\widehat{x}_1^{(\lambda)} \equiv \widehat{x}_2^{(\mu)}$, де $\widehat{x}_1^{(\lambda)} \in \widehat{R}_x$ і $x_2^{(\mu)} \in \widehat{R}_x$, а $\widehat{R}_x \in R$, або $x_1^{(\lambda)} \in W_x$ і $x_2^{(\mu)} \in W_x$ відносяться до одного і того ж синонімічного ряду $W_x \subseteq W$, то необхідно здійснити синтез, зробивши заміну $x_1^{(\lambda)} \equiv x_2^{(\mu)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_1^{(\lambda)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_2(S_2) = x_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_2(S_1) = x_1^{(\lambda)}.$$

Розглянемо текстовий фрагмент, що складається із двох речень: "Він не здатний був на високі посади. Він навіть не був добрим хліборобом" [68].

Логіко-лінгвістична модель текстового фрагменту українською мовою матиме вигляд:

$$L^{S_1} = \neg p_{11} \& p_{12} (x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \neg p_2 (x_2, 0, y_2, q_2, 0, 0, h_2),$$

$$L^{S_1} = \neg \text{здатний\&був (він, 0, посади, високі, 0, 0, 0)},$$

$$L^{S_2} = \neg \text{був (він, 0, хліборобом, добрим, 0, 0, навіть)}.$$

У логіко-лінгвістичних моделях виконано заміну суб'єктів за твердженням 1.2:

$$L^{S_1} = \neg p_{11} \& p_{12} (x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \neg p_{12} (x_1, 0, y_2, q_2, 0, 0, h_2).$$

Для цього ж текстового фрагменту російською мовою "Он не способен был на высокие должности. Он даже не был хорошим земледельцем" маємо наступну логіко-лінгвістичну модель:

$$L^{S_1} = \neg p_{11} \& p_{12} (x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \neg p_2 (x_2, 0, y_2, q_2, 0, 0, h_2),$$

$$L^{S_1} = \neg \text{способен\&был (он, 0, должности, высокие, 0, 0, 0)},$$

$$L^{S_2} = \neg \text{был (он, 0, земледельцем, хорошим, 0, 0, даже)}.$$

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою "He was not capable of high positions. He was not even a good farmer" матиме вигляд:

$$L^{S_1} = \neg p_{11} \& p_{12} (x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \neg p_2 (x_2, 0, y_2, q_2, 0, 0, h_2),$$

$$L^{S_1} = \neg \text{was\&capable (he, 0, positions, high, 0, 0, 0)},$$

$$L^{S_2} = \neg \text{was (he, 0, farmer, good, 0, 0, even)}.$$

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за першим принципом (твердження 1.2), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = \neg p_{11} \& p_{12} (x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \neg p_{12} (x_1, 0, y_2, q_2, 0, 0, h_2).$$

Твердження 1.3. Якщо суб'єкт одного та об'єкт іншого речення природної мови тотожні $x_1^{(\lambda)} \equiv y_2^{(\mu)}$ або $\widehat{x}_1^{(\lambda)} \equiv \widehat{y}_2^{(\mu)}$, де $\widehat{x}_1^{(\lambda)} \in \widehat{R}_x$ і $\widehat{y}_2^{(\mu)} \in \widehat{R}_x$, а $\widehat{R}_x \in R$, або $x_1^{(\lambda)} \in W_x$ і $y_2^{(\mu)} \in W_x$ відносяться до одного і того ж синонімічного ряду $W_x \subseteq W$, то необхідно здійснити синтез, зробивши заміну $x_1^{(\lambda)} \equiv y_2^{(\mu)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$\begin{aligned} L_{p^{(\lambda)}}^{S_1}(h) &= p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}), \\ L_{p^{(\mu)}}^{S_2}(h) &= p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, x_1^{(\lambda)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}), \\ \alpha_{a_1}(l_1) &= u_4(S_2) = y_2^{(\mu)}, \\ \alpha_{a_2}(l_2) &= u_2(S_1) = x_1^{(\lambda)}. \end{aligned}$$

Для текстового фрагменту українською мовою "Орудував він косою, як добрий маляр пензлем, – легко і вправно. Коли б його пустили з косою просто, він обкосив би всю земну кулю" [68] логіко-лінгвістична модель буде мати вигляд:

$$\begin{aligned} L^{S_1} &= [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim \\ & p_1'(x_1', g_1', y_1', 0, 0, 0, 0), \\ L^{S_2} &= p_2(0, 0, y_2, 0, z_2, 0, h_2) \rightarrow [p_2'(x_2', 0, y_2', q_{21}', 0, 0, 0) \& \\ & p_2'(x_2', 0, y_2', q_{22}', 0, 0, 0)]. \\ L^{S_1} &= [\text{орудував (він, 0, косою, 0, 0, 0, легко)} \& \\ & \text{орудував (він, 0, косою, 0, 0, 0, вправно)}] \sim \\ & \text{орудував (маляр, добрий, пензлем, 0, 0, 0, 0)}, \\ L^{S_2} &= \text{пустили (0, 0, його, 0, косою, 0, просто)} \rightarrow \\ & [\text{обкосив (він, 0, кулю, всю, 0, 0, 0)} \& \\ & \text{обкосив (він, 0, кулю, земну, 0, 0, 0)}]. \end{aligned}$$

У фрагменті використано синонімічне повторення, за рахунок чого у логіко-лінгвістичних моделях виконано заміну об'єкта на суб'єкт за твердженням 1.3:

$$\begin{aligned} L^{S_1} &= [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim \\ & p_1(x_1', g_1', y_1', 0, 0, 0, 0), \\ L^{S_2} &= p_2(0, 0, x_1, 0, y_1, 0, h_2) \rightarrow [p_2'(x_1, 0, y_2', q_{21}', 0, 0, 0) \& \\ & p_2'(x_1, 0, y_2', q_{22}', 0, 0, 0)]. \end{aligned}$$

Логіко-лінгвістична модель цього ж текстового фрагменту "Орудовал он косой, как хороший маляр кистью, – легко и умело. Если бы его пустили с косой просто, он обкосил бы весь земной шар" російською мовою:

$$L^{S_1} = [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim p'_1(x'_1, g'_1, y'_1, 0, 0, 0, 0),$$

$$L^{S_2} = p_2(0, 0, y_2, 0, z_2, 0, h_2) \rightarrow [p'_2(x'_2, 0, y'_2, q'_{21}, 0, 0, 0) \& p'_2(x'_2, 0, y'_2, q'_{22}, 0, 0, 0)].$$

$$L^{S_1} = [\text{орудовал (он, 0, косой, 0, 0, 0, легко) \& орудовал (он, 0, косой, 0, 0, 0, умело)}] \sim \text{орудовал (маляр, хороший, кистью, 0, 0, 0, 0)},$$

$$L^{S_2} = \text{пустили (0, 0, его, 0, косой, 0, просто)} \rightarrow [\text{обкосил (он, 0, шар, весь, 0, 0, 0) \& обкосил (он, 0, шар, земной, 0, 0, 0)}].$$

Для фрагменту російською мовою синтез здійснюється за першим принципом (твердження 1.3), а логіко-лінгвістичні моделі набувають вигляду:

$$L^{S_1} = [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim p_1(x'_1, g'_1, y'_1, 0, 0, 0, 0),$$

$$L^{S_2} = p_2(0, 0, x_1, 0, y_1, 0, h_2) \rightarrow [p'_2(x_1, 0, y'_2, q'_{21}, 0, 0, 0) \& p'_2(x_1, 0, y'_2, q'_{22}, 0, 0, 0)].$$

Логіко-лінгвістична модель цього ж текстового фрагменту "He used the scythe as a good painter with a brush - easily and skillfully. If he were allowed to go with a scythe simply, he would have torn the whole globe" англійською мовою має вигляд:

$$L^{S_1} = [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim p'_1(x'_1, g'_1, y'_1, 0, 0, 0, 0),$$

$$L^{S_2} = p_{21} \& p_{22}(0, 0, y_2, 0, z_2, 0, h_2) \rightarrow p'_{21} \& p'_{22} \& p'_{23}(x'_2, 0, y'_2, q'_2, 0, 0, 0),$$

$$L^{S_1} = [\text{used (he, 0, scythe, 0, 0, 0, easily) \& used (he, 0, scythe, 0, 0, 0, skillfull)}] \sim \text{used (painter, good, brush, 0, 0, 0, 0)},$$

$$L^{S_2} = \text{allowed \& go (0, 0, he, 0, scythe, 0, simply)} \rightarrow \text{would \& have \& torn (he, 0, globe, whole, 0, 0, 0)}.$$

Для фрагменту англійською мовою синтез також здійснюється за першим принципом (твердження 1.3), а логіко-лінгвістичні моделі набувають вигляду:

$$L^{S_1} = [p_1(x_1, 0, y_1, 0, 0, 0, h_{11}) \& p_1(x_1, 0, y_1, 0, 0, 0, h_{12})] \sim$$

$$p'_1(x'_1, g'_1, y'_1, 0, 0, 0, 0),$$

$$L^{S_2} = p_{21} \& p_{22}(0, 0, x_1, 0, y_1, 0, h_2) \rightarrow p'_{21} \& p'_{22} \& p'_{23}(x_1, 0, y'_2, q'_2, 0, 0, 0).$$

Твердження 1.4. Якщо об'єкти відношень двох речень природної мови тотожні $y_1^{(\lambda)} \equiv y_2^{(\mu)}$ або $\widehat{y}_1^{(\lambda)} \equiv \widehat{y}_2^{(\mu)}$, де $\widehat{y}_1^{(\lambda)} \in \widehat{R}_y$ і $y_2^{(\mu)} \in \widehat{R}_y$, а $\widehat{R}_y \in R$, або $y_1^{(\lambda)} \in W_y$ і $y_2^{(\mu)} \in W_y$ відносяться до одного і того ж синонімічного ряду $W_y \subseteq W$, то необхідно здійснити синтез, зробивши заміну $y_1^{(\lambda)} \equiv y_2^{(\mu)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L^{S_1}_{p^{(\lambda)}}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L^{S_2}_{p^{(\mu)}}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_1^{(\lambda)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_4(S_2) = y_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_4(S_1) = y_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "Думки треба донести до слухачів. Якщо хочете наголосити на якійсь думці, це можна зробити кількома способами" [33] логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_{11} \& p_{12}(0, 0, y_1, 0, z_1, 0, 0),$$

$$L^{S_2} = p_{21} \& p_{22}(0, 0, y_2, q_2, 0, 0, 0) \rightarrow p'_{21} \& p'_{22}(0, 0, y'_2, 0, z'_2, r'_2, 0),$$

$$L^{S_1} = \text{треба\&донести}(0, 0, \text{думки}, 0, \text{слухачів}, 0, 0),$$

$$L^{S_2} = \text{хочете\&наголосити}(0, 0, \text{думці}, \text{якійсь}, 0, 0, 0) \rightarrow \text{можна\&зробити}(0, 0, \text{це}, 0, \text{способами}, \text{кількома}, 0).$$

За твердженням 1.4 над логіко-лінгвістичними моделями фрагменту було здійснено синтез, внаслідок чого отримано:

$$L^{S_1} = p_{11} \& p_{12}(0, 0, y_1, 0, z_1, 0, 0),$$

$$L^{S_2} = p_{21} \& p_{22}(0, 0, y_1, q_2, 0, 0, 0) \rightarrow p'_{21} \& p'_{22}(0, 0, y'_2, 0, z'_2, r'_2, 0).$$

Логіко-лінгвістична модель цього ж текстового фрагменту "Мысли нужно донести до слушателей. Если хотите выделить

какую-то мысль, это можно сделать несколькими способами" російською мовою:

$$L^{S_1} = p_{11} \& p_{12} (0,0, y_1,0,z_1,0,0),$$

$$L^{S_2} = p_{21} \& p_{22} (0,0, y_2,q_2,0,0,0) \rightarrow p'_{21} \& p'_{22} (0,0, y'_2,0, z'_2, r'_2, 0),$$

$$L^{S_1} = \text{нужно\&донести} (0, 0, \text{мысли}, 0, \text{слушателей}, 0, 0),$$

$$L^{S_2} = \text{хочете\&выделить} (0, 0, \text{мысль, какую-то}, 0, 0, 0) \rightarrow \text{можно\&сделать} (0, 0, \text{это}, 0, \text{способами, несколькими}, 0).$$

Логіко-лінгвістична модель цього ж текстового фрагменту "Thoughts need to be conveyed to listeners. If you want to highlight a thought, you can do it in several ways" англійською мовою матиме вигляд:

$$L^{S_1} = p_{11} \& p_{12} (0,0, y_1,0,z_1,0,0),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2,0, y_2,0,0,0,0) \rightarrow p'_{21} \& p'_{22} (x'_2,0, y'_2,0, z'_2, r'_2, 0),$$

$$L^{S_1} = \text{need\&conveyed} (0, 0, \text{thoughts}, 0, \text{listeners}, 0, 0),$$

$$L^{S_2} = \text{want\&highlight} (\text{you}, 0, \text{thought}, 0, 0, 0, 0) \rightarrow \text{can\&do} (\text{you}, 0, \text{it}, 0, \text{ways, several}, 0).$$

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за першим принципом (твердження 1.4). Після цього логіко-лінгвістичні моделі двох розглянутих речень текстового фрагменту російською набувають вигляду:

$$L^{S_1} = p_{11} \& p_{12} (0,0, y_1,0,z_1,0,0),$$

$$L^{S_2} = p_{21} \& p_{22} (0,0, y_1,q_2,0,0,0) \rightarrow p'_{21} \& p'_{22} (0,0, y'_2,0, z'_2, r'_2, 0),$$

а англійською, відповідно:

$$L^{S_1} = p_{11} \& p_{12} (0,0, y_1,0,z_1,0,0),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2,0, y_1,0,0,0,0) \rightarrow p'_{21} \& p'_{22} (x_2,0, y'_2,0, z'_2, r'_2, 0).$$

Твердження 1.5. Якщо суб'єкт, об'єкт або предмет відношення одного речення тотожний предмету відношення іншого речення, тобто $x_1^{(\lambda)} \equiv z_2^{(\mu)}$ чи $y_{\mu_1} \equiv z_2^{(\mu)}$, чи $z_{\mu_1} \equiv z_2^{(\mu)}$, або $\hat{x}_1^{(\lambda)} \equiv \hat{z}_2^{(\mu)}$ чи $\hat{y}_1^{(\lambda)} \equiv \hat{z}_2^{(\mu)}$, чи $\hat{z}_1^{(\lambda)} \equiv \hat{z}_2^{(\mu)}$, де $\hat{x}_1^{(\lambda)} \in \hat{R}_x$, $\hat{y}_1^{(\lambda)} \in \hat{R}_x$, $\hat{z}_1^{(\lambda)} \in \hat{R}_x$ і $\hat{z}_2^{(\mu)} \in \hat{R}_x$, а $\hat{R}_x \in R$, або $x_1^{(\lambda)} \in W_x$, $y_1^{(\lambda)} \in W_x$, $z_1^{(\lambda)} \in W_x$ і $z_2^{(\mu)} \in W_x$ відносяться до одного і того ж синонімічного ряду $W_x \subseteq W$, то необхідно здійснити синтез, зробивши заміну

$x_1^{(\lambda)} \equiv z_2^{(\mu)}$ або $y_{\mu_1} \equiv z_2^{(\mu)}$, або $z_{\mu_1} \equiv z_2^{(\mu)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_1^{(\lambda)}, q_2^{(\mu)}, x_1^{(\lambda)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_6(S_2) = z_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_2(S_1) = x_1^{(\lambda)} \text{ або}$$

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_1^{(\lambda)}, q_2^{(\mu)}, y_1^{(\lambda)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_6(S_2) = z_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_4(S_1) = y_1^{(\lambda)} \text{ або}$$

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_1^{(\lambda)}, q_2^{(\mu)}, z_1^{(\lambda)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_6(S_2) = z_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_6(S_1) = z_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "Усі вулиці у Вербівці ніби зумисне обсажені високими вербами. Як залле Вербівку літнє палке сонце, то вся кучерява долина здається залитою буйними зеленими морськими хвилями" [69] логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, 0, h_{11} \& h_{12}),$$

$$L^{S_2} = [p_2(x_2, g_2, y_2, 0, 0, 0, 0) \& p_2(x_2, g_2, y_2, 0, 0, 0, 0)] \rightarrow$$

$$[p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{23}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{23}, 0, 0, 0)],$$

L^{S_1} = обсажені (вулиці, усі, вербами, високими, Вербівці, 0, ніби&зумисне),

$$L^{S_2} = [залле (сонце, літнє, Вербівку, 0, 0, 0, 0) \&$$

залле (сонце, палке, Вербівку, 0, 0, 0, 0)] \rightarrow

[здається&залитою (долина, вся, хвилями, буйними, 0, 0, 0)&

здається&залитою (долина, вся, хвилями, зеленими, 0, 0, 0)&
здається&залитою (долина, вся, хвилями, морськими, 0, 0, 0)&
здається&залитою (долина, кучерява, хвилями, буйними, 0, 0, 0)&
здається&залитою (долина, кучерява, хвилями, зеленими, 0, 0, 0)&
здається&залитою (долина, кучерява, хвилями, морськими, 0, 0, 0)].

За твердженням 1.5 над логіко-лінгвістичними моделями фрагменту було здійснено синтез, внаслідок чого отримано:

$$L^S_1 = p_1(x_1, g_1, y_1, q_1, z_1, 0, h_{11} \& h_{12}),$$

$$L^S_2 = [p_2(x_2, g_2, z_1, 0, 0, 0, 0) \& p_2(x_2, g_2, z_1, 0, 0, 0, 0)] \rightarrow$$

$$[p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{23}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{23}, 0, 0, 0)].$$

Логіко-лінгвістична модель цього ж текстового фрагменту "Все улицы в Вербовке как-будто специально осажены высокими вербами. Когда зальет Вербовку летнее жаркое солнце, то вся кучерявая долина кажется залита буйными зелеными морскими волнами" російською мовою:

$$L^S_1 = p_1(x_1, g_1, y_1, q_1, z_1, 0, h_{11} \& h_{12}),$$

$$L^S_2 = [p_2(x_2, g_2, y_2, 0, 0, 0, 0) \& p_2(x_2, g_2, y_2, 0, 0, 0, 0)] \rightarrow$$

$$[p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{23}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{23}, 0, 0, 0)],$$

$L^S_1 =$ осажены (улицы, все, вербами, высокими, Вербовке, 0, как-будто&специально),

$L^S_2 =$ [зальет (солнце, летнее, Вербовку, 0, 0, 0, 0)& зальет (солнце, жаркое, Вербовку, 0, 0, 0, 0)] \rightarrow

[кажется&залита (долина, вся, волнами, буйными, 0, 0, 0)& кажется&залита (долина, вся, волнами, зелеными, 0, 0, 0)& кажется&залита (долина, вся, волнами, морскими, 0, 0, 0)& кажется&залита (долина, кучерявая, волнами, буйными, 0, 0, 0)& кажется&залита (долина, кучерявая, волнами, зелеными, 0, 0, 0)&

кажется&залита (долина, кучерявая, волнами, морскими, 0, 0, 0)].

Логіко-лінгвістична модель цього ж текстового фрагменту "All the streets in Verbovka seem to be specially besieged with high willows. When the hot summer sun floods the Verbovka, the whole curly valley seems to be flooded with riotous green sea waves" англійською мовою:

$$L^S_1 = p_1(x_1, g_1, y_1, q_1, z_1, 0, h_{11} \& h_{12}),$$

$$L^S_2 = [p_2(x_2, g_2, y_2, 0, 0, 0, 0) \& p_2(x_2, g_2, y_2, 0, 0, 0, 0)] \rightarrow$$

$$[p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{21}, y'_2, q'_{23}, 0, 0, 0)] \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p'_{22}(x'_2, g'_{22}, y'_2, q'_{23}, 0, 0, 0)],$$

$L^S_1 =$ besieged (streets, all, willows, high, Verbovka, 0, seem & specially),

$$L^S_2 = [floods (sun, hot, Verbovka, 0, 0, 0, 0) \&$$

floods (sun, summer, Verbovka, 0, 0, 0, 0)] \rightarrow

[seems & flooded (valley, whole, waves, riotous, 0, 0, 0) & seems & flooded (valley, whole, waves, green, 0, 0, 0) & seems & flooded (valley, whole, waves sea, 0, 0, 0) & seems & flooded (valley, curly, waves, riotous, 0, 0, 0) & seems & flooded (valley, curly, waves, green, 0, 0, 0) & seems & flooded (valley, curly, waves sea, 0, 0, 0)].

Після здійснення синтезу над логіко-лінгвістичними моделями текстових фрагментів на російській та англійській мові за твердженням 1.5 буде отримано:

$$L^S_1 = p_1(x_1, g_1, y_1, q_1, z_1, 0, h_{11} \& h_{12}),$$

$$L^S_2 = [p_2(x_2, g_2, z_1, 0, 0, 0, 0) \& p_2(x_2, g_2, z_1, 0, 0, 0, 0)] \rightarrow$$

$$[p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{21}, y'_2, q'_{23}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{21}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{22}, 0, 0, 0) \& p'_{21} \& p_2(x'_2, g'_{22}, y'_2, q'_{23}, 0, 0, 0)].$$

Як видно із наведених прикладів, при синтезі логіко-лінгвістичних моделей за першим принципом для всіх тверджень справджується гіпотеза про те, що для різних флективних мов (української, російської та англійської) здійснюється заміна одних

і тих самих компонент логіко-лінгвістичних моделей речень природної мови. Це свідчить про універсальність запропонованого принципу синтезу.

Принцип 2. Принцип базується на дейктичному повторенні, що відноситься до способів контактного зв'язку в текстах і означає, що на вже названі предмети, характеристики та обставини посилаються в подальших реченнях за допомогою спеціальних слів: займенників, прислівників тощо [70, 71].

На відміну від першого принципу, порядок вживання речень у текстовому фрагменті відіграє принципове значення.

Твердження 2.1. Якщо суб'єкт $x_1^{(\lambda)}$ володіє характеристиками $Z(x_1^{(\lambda)}) = \{1, 1, n(x_1^{(\lambda)}), k2(x_1^{(\lambda)}), 0, 0, l(x_1^{(\lambda)}), 1\}$ [72], а суб'єкт $x_2^{(\mu)}$ має характеристики $Z(x_2^{(\mu)}) = \{4, 1, n(x_1^{(\lambda)}), k2(x_1^{(\lambda)}), 0, 0, l(x_1^{(\lambda)}), 1\}$, то необхідно здійснити синтез, зробивши заміну $x_2^{(\mu)} \equiv x_1^{(\lambda)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_1^{(\lambda)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_2(S_2) = x_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_2(S_1) = x_1^{(\lambda)}.$$

Розглянемо текстовий фрагмент, що складається із двох речень: *"Дядько Самійло не був ні професором, ні лікарем, ні інженером. Він не здатний був на високі посади"* [68].

Логіко-лінгвістична модель текстового фрагменту українською мовою матиме вигляд:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_{22}(x_2, 0, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \neg \text{був (дядько, Самійло, } \neg \text{професором, } 0, 0, 0, 0) \& \neg \text{був (дядько, Самійло, } \neg \text{лікарем, } 0, 0, 0, 0) \& \neg \text{був (дядько, Самійло, } \neg \text{інженером, } 0, 0, 0, 0),$$

$$L^{S_2} = \neg \text{здатний} \& \text{був (він, } 0, \text{посади, високі, } 0, 0, 0, 0).$$

У цьому текстовому фрагменті використано дейктичне повторення, внаслідок чого до логіко-лінгвістичних моделей

застосовано другий принцип синтезу твердження 2.1 і виконано заміну суб'єктів:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_1(x_1, g_1, y_2, q_2, 0, 0, 0).$$

Для цього ж текстового фрагменту російською мовою "Дядя Самойло не был ни профессором, ни доктором, ни инженером. Он не способен был на высокие должности" маємо:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_{22}(x_2, 0, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \neg \text{был (дядя, Самойло, } \neg \text{профессором, } 0, 0, 0, 0) \& \neg \text{быв (дядько, Самйло, } \neg \text{доктором, } 0, 0, 0, 0) \& \neg \text{быв (дядько, Самйло, } \neg \text{инженером, } 0, 0, 0, 0),$$

$$L^{S_2} = \neg \text{способен \& был (он, } 0, \text{ должности, высокие, } 0, 0, 0).$$

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою "Uncle Samoilo was neither a professor, nor a doctor, nor an engineer. He was not able to hold high posts" матиме вигляд:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_{22}(x_2, 0, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \neg \text{was (uncle, Samoilo, } \neg \text{professor, } 0, 0, 0, 0) \& \neg \text{was (uncle, Samoilo, } \neg \text{doctor, } 0, 0, 0, 0) \& \neg \text{was (uncle, Samoilo, } \neg \text{engineer, } 0, 0, 0, 0),$$

$$L^{S_2} = \neg \text{was \& capable (he, } 0, \text{ positions, high, } 0, 0, 0).$$

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за другим принципом (твердження 2.1), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_1(x_1, g_1, y_2, q_2, 0, 0, 0).$$

Твердження 2.2. Якщо об'єкт $y_1^{(\lambda)}$ володіє характеристиками $Z(y_1^{(\lambda)}) = \{1, g(y_1^{(\lambda)}) \neq 1, n(y_1^{(\lambda)}), k2(y_1^{(\lambda)}), 0, 0, l(y_1^{(\lambda)}), 3\}$, а суб'єкт $x_2^{(\mu)}$ має характеристики $Z(x_2^{(\mu)}) = \{4, 1, n(y_1^{(\lambda)}), k2(y_1^{(\lambda)}), 0, 0, l(y_1^{(\lambda)}), 1\}$, то необхідно здійснити синтез, зробивши заміну $x_2^{(\mu)} \equiv y_1^{(\lambda)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(y_1^{(\lambda)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_2(S_2) = x_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_4(S_1) = y_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "*Рано починається таврійська весна! Починають її жайворонки*" [73] логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_1(x_1, g_1, 0, 0, 0, 0, h_1),$$

$$L^{S_2} = p_2(x_2, 0, y_2, 0, 0, 0, 0),$$

$$L^{S_1} = \text{починається (весна, таврійська, 0, 0, 0, 0, рано)},$$

$$L^{S_2} = \text{починають (жайворонки, 0, її, 0, 0, 0, 0)}.$$

До текстового фрагменту застосовано другий принцип синтезу твердження 2.2, внаслідок чого логіко-лінгвістична модель набуває вигляду:

$$L^{S_1} = p_1(x_1, g_1, 0, 0, 0, 0, h_1),$$

$$L^{S_2} = p_1(x_1, 0, x_1, 0, 0, 0, 0).$$

Логіко-лінгвістична модель цього ж текстового фрагменту "*Рано начинается таврийская весна! Начинают ее жаворонки*" російською мовою:

$$L^{S_1} = p_1(x_1, g_1, 0, 0, 0, 0, h_1),$$

$$L^{S_2} = p_2(x_2, 0, y_2, 0, 0, 0, 0),$$

$$L^{S_1} = \text{начинается (весна, таврийская, 0, 0, 0, 0, рано)},$$

$$L^{S_2} = \text{начинают (жаворонки, 0, ее, 0, 0, 0, 0)}.$$

Логіко-лінгвістична модель цього ж текстового фрагменту "*Early tavriya spring begins! The larks start her*" англійською мовою:

$$L^{S_1} = p_1(x_1, g_1, 0, 0, 0, 0, h_1),$$

$$L^{S_2} = p_2(x_2, 0, y_2, 0, 0, 0, 0),$$

$$L^{S_1} = \text{begins (spring, tavriya, 0, 0, 0, 0, early)},$$

$$L^{S_2} = \text{start (larks, 0, her, 0, 0, 0, 0)}.$$

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за другим принципом (твердження 2.2), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = p_1(x_1, g_1, 0, 0, 0, 0, h_1),$$

$$L^{S_2} = p_1(x_1, 0, x_1, 0, 0, 0, 0).$$

Твердження 2.3. Якщо предмет відношення $z_1^{(\lambda)}$ володіє характеристиками $Z(z_1^{(\lambda)}) = \{1, g(z_1^{(\lambda)}) \neq 1, n(z_1^{(\lambda)}), k2(z_1^{(\lambda)}), 0, 0, l(z_1^{(\lambda)}), 3\}$, а суб'єкт $x_2^{(\mu)}$ має характеристики $Z(x_2^{(\mu)}) = \{4, 1, n(z_1^{(\lambda)}), k2(z_1^{(\lambda)}), 0, 0, l(z_1^{(\lambda)}), 1\}$, то необхідно здійснити синтез, зробивши заміну $x_2^{(\mu)} \equiv z_1^{(\lambda)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(z_1^{(\lambda)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)}),$$

$$\alpha_{a_1}(l_1) = u_2(S_2) = x_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_6(S_1) = z_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "Вересень бив у золоті дзвони соняшників [33]. Вони стояли важкі, повні чорних коштовностей" логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, z_1, 0, 0),$$

$$L^{S_2} = p_2(x_2, g_{21}, 0, 0, 0, 0, 0) \& p_2(x_2, g_{22}, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \text{бив (вересень, 0, дзвони, золоті, соняшників, 0, 0)},$$

$$L^{S_2} = \text{стояли (вони, важкі, 0, 0, 0, 0, 0) \&}$$

стояли (вони, повні, коштовностей, чорних, 0, 0, 0).

До текстового фрагменту застосовано другий принцип синтезу твердження 2.3, внаслідок чого логіко-лінгвістична модель набуває вигляду:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, z_1, 0, 0),$$

$$L^{S_2} = p_2(z_1, g_{21}, 0, 0, 0, 0, 0) \& p_2(z_1, g_{22}, y_2, q_2, 0, 0, 0).$$

Логіко-лінгвістична модель цього ж текстового фрагменту "Сентябрь бил в золотые колокола подсолнухов. Они стояли тяжелые, полные черных драгоценностей" російською мовою:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, z_1, 0, 0),$$

$$L^{S_2} = p_2(x_2, g_{21}, 0, 0, 0, 0, 0) \& p_2(x_2, g_{22}, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \text{бил (сентябрь, 0, колокола, золотые, подсолнухов, 0, 0)},$$

$$L^{S_2} = \text{стояли (они, тяжелые, 0, 0, 0, 0, 0) \&}$$

$\text{стояли (они, полные, драгоценностей, черных, 0, 0, 0)}$.

Логіко-лінгвістична модель цього ж текстового фрагменту "September beat the golden bells of sunflowers. They stood heavy, full of black jewelry" англійською мовою:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, z_1, 0, 0),$$

$$L^{S_2} = p_2(x_2, g_{21}, 0, 0, 0, 0, 0) \& p_2(x_2, g_{22}, y_2, q_2, 0, 0, 0),$$

$$L^{S_1} = \text{beat (september, 0, bells, golden, sunflowers, 0, 0)},$$

$$L^{S_2} = \text{stood (they, heavy, 0, 0, 0, 0, 0) \&}$$

$\text{stood (they, heavy, jewelry, black, 0, 0, 0)}$.

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за другим принципом (твердження 2.3), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, z_1, 0, 0),$$

$$L^{S_2} = p_2(z_1, g_{21}, 0, 0, 0, 0, 0) \& p_2(z_1, g_{22}, y_2, q_2, 0, 0, 0).$$

Твердження 2.4. Якщо характеристика відношення $h_1^{(\lambda)}$ має параметри $Z(h_1^{(\lambda)}) = \{8, 0, 0, 0, 0, 0, 0, 5\}$, а характеристика відношення $h_1^{(\lambda)} \in V_h$ входить до множини вказівних слів V_h , то необхідно здійснити синтез, зробивши заміну $h_2^{(\mu)} \equiv h_1^{(\lambda)}$ в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_1^{(\lambda)}),$$

$$\alpha_{a_1}(l_1) = u_8(S_2) = h_2^{(\mu)},$$

$$\alpha_{a_2}(I_2) = u_8(S_1) = h_1^{(\lambda)}.$$

Нехай текстовий фрагмент складається із двох речень:
*"Керівництво університетів оновлює навчальні плани навесні.
 Тоді ще не сформоване навчальне навантаження"*.

Логіко-лінгвістична модель текстового фрагменту українською мовою матиме вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, 0, 0, h_1),$$

$$L^{S_2} = \neg p_2(0, 0, y_2, q_2, 0, 0, h_2),$$

L^{S_1} = оновлює (керівництво, університетів, плани, навчальні, 0, 0, навесні),

L^{S_2} = \neg сформоване (0, 0, навантаження, навчальне, 0, 0, тоді).

До текстового фрагменту застосовано другий принцип синтезу твердження 2.4, внаслідок чого логіко-лінгвістична модель набуває вигляду:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, 0, 0, h_1),$$

$$L^{S_2} = \neg p_2(0, 0, y_2, q_1, 0, 0, h_1).$$

Для цього ж текстового фрагменту російською мовою
*"Руководство университетов обновляет учебные планы весной.
 Тогда еще не сформирована учебная нагрузка"* логіко-лінгвістична модель матиме вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, 0, 0, h_1),$$

$$L^{S_2} = \neg p_2(0, 0, y_2, q_2, 0, 0, h_2),$$

L^{S_1} = обновляет (руководство, университетов, планы, учебные, 0, 0, весной),

L^{S_2} = \neg сформирована (0, 0, нагрузка, учебная, 0, 0, тогда).

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою
"University management updates educational plans in spring. Then the academic load has not yet been formed" матиме вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, 0, 0, h_1),$$

$$L^{S_2} = \neg p_2(0, 0, y_2, q_2, 0, 0, h_2),$$

L^{S_1} = updates (management, university, plans, educational, 0, 0, spring),

L^{S_2} = \neg formed (0, 0, load, academic, 0, 0, then).

Для фрагментів російською та англійською мовами синтез аналогічно, як і для фрагменту українською мовою, здійснюється за другим принципом (твердження 2.4), після чого логіко-лінгвістичні моделі двох розглянутих речень набувають вигляду:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, 0, 0, h_1),$$

$$L^{S_2} = -p_2(0, 0, y_2, q_1, 0, 0, h_1).$$

Принцип 3. Принцип формулюється, виходячи з повторення граматичних параметрів в реченнях. Так, у зв'язному тексті дієслівні форми об'єднують речення природної мови спільним часовим планом, не дивлячись на те, що як предикати, так і суб'єкти двох речень не являються тотожними [74]. Тобто, якщо відношення одного із речень $p_1^{(\lambda)}$ володіє характеристиками $Z(p_1^{(\lambda)}) = \{5, 0, n(p_1^{(\lambda)}), k2(p_1^{(\lambda)}), t(p_1^{(\lambda)}), h(p_1^{(\lambda)}), l(p_1^{(\lambda)}), 2\}$, які співпадають з граматичними параметрами відношення $p_2^{(\mu)}$ другого речення $Z(p_2^{(\mu)}) = \{5, 0, n(p_1^{(\lambda)}), k2(p_1^{(\lambda)}), t(p_1^{(\lambda)}), h(p_1^{(\lambda)}), l(p_1^{(\lambda)}), 2\}$, то необхідно здійснити синтез, не роблячи заміну в логіко-лінгвістичних моделях L^{S_1} та L^{S_2} , але заносючи елементи до векторів характеристик:

$$\alpha_{a_1}(l_1) = u_1(S_2) = p_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_1(S_1) = p_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "На Поділлі були традиційними геометричні та геометризовані рушники з малюнками жіночих фігурок, пташок, тварин. Волинські рушники оздоблювалися однією горизонтальною смугою геометричного орнаменту, виконаного нанизанням, хрестиком..." [33] логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{12}, 0, h_1) \&$$

$$p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{13}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{11}, r_1, h_1) \&$$

$$p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{13}, 0, h_1),$$

$$L^{S_2} = [p_2(x_2, g_2, y_2, q_{21}, z_1, r_1, 0) \& p_2(x_2, g_2, y_2, q_{22}, z_1, r_1, 0)] \&$$

$$[p'_2(x'_2, 0, y'_{21}, 0, 0, 0, 0) \& p'_2(x'_2, 0, y'_{22}, 0, 0, 0, 0)],$$

L^{S_1} = були&традиційними (рушники, геометричні, малюнками, 0, фігурок, жіночих, Поділлі)&

були&традиційними (рушники, геометричні, малюнками, 0, пташок, 0, Поділлі)&

були&традиційними (рушники, геометричні, малюнками, 0, тварин, 0, Поділлі)&

були&традиційними (рушники, геометризovanі, малюнками, 0, фігурок, жіночих, Поділлі)&

були&традиційними (рушники, геометризovanі, малюнками, 0, пташок, 0, Поділлі)&

були&традиційними (рушники, геометризovanі, малюнками, 0, тварин, 0, Поділлі),

$L^{S^2} =$ [оздоблювалися (рушники, волинські, смугою, однією, орнаменту, геометричного, 0) &

оздоблювалися (рушники, волинські, смугою, горизонтальною, орнаменту, геометричного, 0)]&

[виконаний (орнамент, 0, нанизуванням, 0, 0, 0, 0)&

виконаний (орнамент, 0, хрестиком, 0, 0, 0, 0)].

Логіко-лінгвістична модель цього ж текстового фрагменту російською мовою "На Подолье традиционными были геометрические и геометризované полотенца с рисунками женских фигурок, птиц, зверей. Волынские полотенца украшались одной горизонтальной полосой геометрического орнамента, исполненного нанизыванием, крестиком..." матиме вигляд:

$L^{S^1} = p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{12}, 0, h_1) \&$

$p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{13}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{11}, r_1, h_1) \&$

$p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{13}, 0, h_1),$

$L^{S^2} = [p_2(x_2, g_2, y_2, q_{21}, z_2, r_1, 0) \& p_2(x_2, g_2, y_2, q_{22}, z_2, r_1, 0)] \&$

$[p'_2(x'_2, 0, y'_{21}, 0, 0, 0, 0) \& p'_2(x'_2, 0, y'_{22}, 0, 0, 0, 0)],$

$L^{S^1} =$ были&традиционными (полотенца, геометрические, рисунками, 0, фигурок, женских, Подолье)&

были&традиционными (полотенца, геометрические, рисунками, 0, птиц, 0, Подолье)&

были&традиционными (полотенца, геометрические, рисунками, 0, зверей, 0, Подолье)&

были&традиционными (полотенца, геометризované, рисунками, 0, фигурок, женских, Подолье)&

были&традиционными (полотенца, геометризованные, рисунками, 0, птиц, 0, Подолье)&

были&традиционными (полотенца, геометризованные, рисунками, 0, зверей, 0, Подолье),

$L^{S_2} =$ [украшались (полотенца, волынские, полосой, одной, орнамента, геометрического, 0) &

украшались (полотенца, волынские, полосой, горизонтальной, орнамента, геометрического, 0)]&

[исполненный (орнамент, 0, нанизыванием, 0, 0, 0, 0)&

исполненный (орнамент, 0, крестиком, 0, 0, 0, 0)].

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою "On Podillia, there were traditional geometric and geometrized towels with drawings of female figures, birds, animals. Volyn's towels were decorated with one horizontal strip of a geometrical ornament, executed by stringing, cross..." матиме вигляд:

$L^{S_1} = p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{12}, 0, h_1) \&$

$p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{13}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{11}, r_1, h_1) \&$

$p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{13}, 0, h_1),$

$L^{S_2} = [p_2(x_2, g_2, y_2, q_{21}, z_2, r_1, 0) \& p_2(x_2, g_2, y_2, q_{22}, z_2, r_1, 0)] \&$

$[p'_2(x'_2, 0, y'_{21}, 0, 0, 0, 0) \& p'_2(x'_2, 0, y'_{22}, 0, 0, 0, 0)],$

$L^{S_1} =$ were & traditional (towels, geometric, drawings, 0, figures, female, Podillia)&

were & traditional (towels, geometric, drawings, 0, birds, 0, Podillia)&

were & traditional (towels, geometric, drawings, 0, animals, 0, Podillia)&

were & traditional (towels, geometrized, drawings, 0, figures, female, Podillia)&

were & traditional (towels, geometrized, drawings, 0, birds, 0, Podillia)&

were & traditional (towels, geometrized, drawings, 0, animals, 0, Podillia),

$L^{S_2} =$ [decorated (towels, volyn's, strip, one, ornament, geometrical, 0) &

decorated (towels, volyn's, strip, horizontal, ornament, geometrical, 0)]&

*[executed (ornament, 0, stringing, 0, 0, 0, 0)] &
executed (ornament, 0, хрестиком, 0, 0, 0, 0)].*

У даному текстовому фрагменті для різних флективних мов (української, російської та англійської) здійснюється третій та перший принцип синтезу (твердження 1.1), внаслідок чого логіко-лінгвістичні моделі набувають вигляду:

$$L^{S_1} = p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{13}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{13}, 0, h_1),$$

$$L^{S_2} = [p_2(x_1, g_2, y_2, q_{21}, z_2, g_{11}, 0) \& p_2(x_1, g_2, y_2, q_{22}, z_2, g_{11}, 0)] \& [p'_2(z_2, 0, y'_{21}, 0, 0, 0, 0) \& p'_2(z_2, 0, y'_{22}, 0, 0, 0, 0)].$$

Принцип 4. Принцип базується на здійсненні заміни синонімів як з однієї, так і з різних частин мови. Це стає можливим завдяки синтаксичній або транспозиційній деривації: переході слова з однієї частини мови в іншу без заміни їх лексичних значень. Найбільш поширеними способами побудови такого роду синонімічних конструкцій є: «дієслово 1» – «дієслово 2 + спільнокореневий іменник до дієслова 1», «дієслово 1 – спільнокореневий іменник до дієслова 1» та «дієслово 1 – спільнокореневий прикметник до дієслова 1» [75]. За четвертим принципом суб'єкт $x_2^{(\mu)}$ другого речення S_2 замінюється предикатом $p_1^{(\lambda)}$ речення S_1 . Внаслідок приведення дієслівної форми предикатів до еквівалентних іменників можна здійснити синтез логіко-лінгвістичних моделей L^{S_1} та L^{S_2} , після чого:

$$L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}),$$

$$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(p_1^{(\mu)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_1^{(\lambda)}),$$

$$\alpha_{a_1}(l_1) = u_2(S_2) = p_2^{(\mu)},$$

$$\alpha_{a_2}(l_2) = u_1(S_1) = p_1^{(\lambda)}.$$

Для текстового фрагменту українською мовою "Книга дуже зацікавила вчених з технічних спеціальностей. Вона також буде цікавою спеціалістам з комп'ютерної лінгвістики" логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, z_1, r_1, h_1),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2, 0, y_2, 0, z_2, r_2, h_2),$$

L^{S_1} = зацікавила (книга, 0, вчених, 0, спеціальностей, технічних, дуже),

L^{S_2} = буде&цікавою (вона, 0, спеціалістам, 0, лінгвістики, комп'ютерної, також).

Логіко-лінгвістична модель цього ж текстового фрагменту російською мовою "Книга очень заинтересовала ученых по техническим специальностям. Она также будет интересна специалистам по компьютерной лингвистике" матиме вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, z_1, r_1, h_1),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2, 0, y_2, 0, z_2, r_2, h_2),$$

L^{S_1} = заинтересовала (книга, 0, ученых, 0, специальностей, технических, очень),

L^{S_2} = будет&интересна (она, 0, специалистам, 0, лингвистике, компьютерной, также).

Логіко-лінгвістична модель цього ж текстового фрагменту англійською мовою " Scientists in technical specialties were very interested with the book. It will also be interesting to specialists in computer linguistics" матиме вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, z_1, r_1, h_1),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2, 0, y_2, 0, z_2, r_2, h_2),$$

L^{S_1} = interested (book, 0, scientists, 0, specialties, technical, very),

L^{S_2} = be&interesting (it, 0, specialists, 0, linguistics, computer, also).

У даному текстовому фрагменті для різних флективних мов (української, російської та англійської) справджується четвертий принцип синтезу, внаслідок чого логіко-лінгвістичні моделі набувають вигляду:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, z_1, r_1, h_1),$$

$$L^{S_2} = p_{21} \& p_{22} (x_2, 0, y_2, 0, z_2, r_2, h_2).$$

Рекурсивне застосування принципів синтезу логіко-лінгвістичних моделей до всіх речень текстів дає можливість отримати кінцеву множину моделей t' (1.28), що відображають всі елементи змісту. Також сформульовані принципи синтезу логіко-лінгвістичних моделей використовуються для формування множини продукцій текстової бази.

2.2 Правила синтезу логіко-лінгвістичних моделей текстових фрагментів

Для того, щоб аналізувати зміст текстів, необхідно виявити зв'язки між лексико-граматичним та семантичним рівнями флективної мови, встановити, які саме когнітивні процеси до цього приводять. Це дасть можливість пояснити, як синтаксичні ланцюжки передають різний семантичний зміст. Саме синтаксичні ланцюжки є структурними компонентами логіко-лінгвістичних моделей текстової інформації.

Правила синтезу логіко-лінгвістичних моделей представляють собою умови, при виконанні яких в логіко-лінгвістичних моделях здійснюються заміни тотожних за змістом компонент, а також утворюються одномірні масиви, елементи яких є словами чи словосполученнями, що пов'язують речення природної мови одне з одним у частинах тексту [76].

Правила синтезу були сформульовані на базі основних принципів синтезу логіко-лінгвістичних моделей речень природної мови, тому правила перегукуються з принципами та являються своєрідною інструкцією для створення масивів характеристик кожного з речень природної мови при побудові логіко-лінгвістичної моделі тексту (1.24).

Правило 1. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких суб'єкти $x_1, \dots, x_i, \dots, x_k$ тотожні, виражені такими частинами мови, як іменники, займенники та чисельники або являються синонімами, то необхідно здійснити заміну суб'єктів: $x_1, \dots, x_i \equiv x_1, \dots, x_k \equiv x_1$ та сформувати одномірний масив зв'язків: $l_i = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k\}$.

Наприклад, нехай задано наступний фрагмент тексту: *"Комп'ютерне моделювання входить до переліку основних дисциплін. Воно вивчає основні методи та алгоритми моделювання комп'ютерними засобами. Комп'ютерне моделювання також передбачає створення імітаційних моделей"*.

Логіко-лінгвістичні моделі для речень цього фрагменту матимуть вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, 0, z_1, r_1, 0),$$

$L^{S_2} = p_2(x_1, g_1, y_{21}, q_2, z_1, g_1, 0) \& p_2(x_1, g_1, y_{22}, q_2, z_1, g_1, 0),$

$L^{S_3} = p_3(x_1, g_1, y_3, 0, x_1, r_3, h_3),$

$L^{S_1} =$ входить (моделювання, комп'ютерне, переліку, 0, дисциплін, основних, 0),

$L^{S_2} =$ вивчає (моделювання, комп'ютерне, методи, основні, моделювання, засобами, комп'ютерними, 0) & вивчає (моделювання, комп'ютерне, алгоритми, основні, моделювання, засобами, комп'ютерними, 0),

$L^{S_3} =$ передбачає (моделювання, комп'ютерне, створення, 0, моделей, імітаційних, також).

Таким чином, одномірний масив характеристик для кожного речення текстового фрагменту буде складатися з суб'єктів:

$l_1 = \{x_2, z_2, x_3, z_3\}, l_2 = \{x_1, g_1, x_3, g_3, z_3\}, l_3 = \{x_1, g_1, x_2, z_2\}.$

Правило 2. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких відношення $p_1, \dots, p_i, \dots, p_k$ тотожні, виражені такими частинами мови, як дієслово, іменник або являються синонімами, то необхідно здійснити заміну відношень: $p_1, \dots, p_i \equiv p_1, \dots, p_k \equiv p_1$ та сформуванати одномірний масив зв'язків: $l_i = p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_k$.

Наприклад, для текстового фрагменту, що складається із двох речень: "Хто говорить надто тихо, той справляє враження людини, яка не вірить у свої сили. Той же, хто говорить надто голосно, справляє враження агресивної людини" [67].

Логіко-лінгвістична модель фрагмента тексту українською мовою матиме вигляд:

$L^{S_1} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{12}) \&$

$p_{12}(x_1, 0, y_1, 0, z_1, 0, 0) \& p'_1(z_1, 0, y'_1, q'_1, 0, 0, 0),$

$L^{S_2} = p_{11}(x_1, 0, 0, 0, 0, 0, h_{11} \& h_{22}) \& p_{12}(x_1, 0, y_1, 0, z_1, r_2, 0),$

$L^{S_1} =$ говорить (хто, 0, 0, 0, 0, 0, надто&тихо) & справляє (той, 0, враження, 0, людини, 0, 0) & \neg вірить (людина, 0, сили, свої, 0, 0, 0),

$L^{S_2} =$ говорить (хто, 0, 0, 0, 0, 0, надто&голосно) & справляє (той, 0, враження, 0, людини, агресивної, 0).

Після застосування другого правила синтезу для кожного з речень буде сформовано такі одномірні масиви характеристик: $l_1 = \{p_{21}, p_{22}, y_2, z_2\}$, $l_2 = \{p_{11}, p_{12}, x_1, y_1, z_1\}$.

Правило 3. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких об'єкт або предмет відношення кожного попереднього речення тотожний, виражений займенником або являється синонімом по відношенню до суб'єкта наступного речення, то необхідно здійснити заміну: $y_1, \dots, x_i \equiv y_{i-1}, \dots, x_k \equiv y_{k-1}$ або $y_1, \dots, y_i \equiv z_{i-1}, \dots, y_k \equiv z_{k-1}$ та сформуванати одномірний масив зв'язків: $l_i = \{y_1, \dots, x_i, x_{i+1}, \dots, x_k\}$ або $l_i = \{y_1, \dots, z_i, z_{i+1}, \dots, z_k\}$, відповідно. Взаємозв'язки між об'єктами та предметами відношень попередніх речень з суб'єктами наступних можуть чергуватися.

Розглянемо текстовий фрагмент "*Вишитий рушник створював настрій, формував естетичні смаки, був узірцем людської працьовитості. Згадаймо про його утилітарне призначення...*" [77].

Логіко-лінгвістичні моделі для речень цього фрагменту матимуть вигляд:

$$L^{S_1} = p_{11}(x_1, g_1, y_{11}, 0, 0, 0, 0) \& p_{12}(x_1, g_1, y_{12}, q_1, 0, 0, 0) \&$$

$$p_{13}(x_1, g_1, y_{13}, 0, z_1, r_1, 0),$$

$$L^{S_2} = p_2(0, 0, y_2, q_2, x_1, g_1, 0),$$

$L^{S_1} =$ створював (рушник, вишитий, настрій, 0, 0, 0, 0) &
формував (рушник, вишитий, смаки, естетичні, 0, 0, 0) &
був (рушник, вишитий, узірцем, 0, працьовитості, людської, 0),

$L^{S_2} =$ згадаймо (0, 0, призначення, утилітарне, рушника, вишитого, 0).

У цьому випадку масив характеристик кожного з речень буде мати вигляд: $l_1 = \{z_2, r_2\}$ та $l_2 = \{x_1, g_2\}$.

Правило 4. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких суб'єкт або об'єкт першого речення поступово виступає характеристиками

суб'єктів, об'єктів та предметів відношень всіх наступних речень, то необхідно здійснити заміну: $x_1, \dots, g_i \equiv x_1, \dots, r_k \equiv x_1$ або $y_1, \dots, g_i \equiv y_1, \dots, r_k \equiv y_1$ та сформувати одномірний масив зв'язків: $l_i = \{x_1, \dots, g_i, q_{i+1}, \dots, r_k\}$ або $l_i = \{y_1, \dots, g_i, q_{i+1}, \dots, r_k\}$ відповідно. Характеристики суб'єктів, об'єктів та предметів відношень із наступних речень можуть чергуватися.

Нехай задано такий фрагмент: "*Комп'ютерне моделювання входить до переліку основних дисциплін. Воно передбачає використання різноманітних моделюючих пристроїв*".

Логіко-лінгвістичні моделі для речень цього фрагменту матимуть вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, 0, z_1, r_1, 0),$$

$$L^{S_2} = p_2(x_1, g_1, y_2, 0, z_2, r_{21}, 0) \& p_2(x_1, g_1, y_2, 0, z_2, x_1, 0),$$

L^{S_1} = входить (моделювання, комп'ютерне, переліку, 0, дисциплін, основних, 0),

L^{S_2} = передбачає (моделювання, комп'ютерне, використання, 0, пристроїв, різноманітних, 0) &

передбачає (моделювання, комп'ютерне, використання, 0, пристроїв, моделюючих, 0).

Масиви характеристик для речень заданого текстового фрагменту будуть мати вигляд: $l_i = \{x_2, g_2, r_{22}\}$, $l_2 = \{x_1, g_1\}$.

Правило 5. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких граматичні параметри відношень $p_1, \dots, p_i, \dots, p_k$ або суб'єктів $x_1, \dots, x_i, \dots, x_k$, або об'єктів $y_1, \dots, y_i, \dots, y_k$ однакові, виражені однаковими частинами мови або являються гіперонімами, то необхідно, не роблячи заміну в логіко-лінгвістичних моделях, сформувати для кожного речення текстового фрагменту одномірний масив зв'язків: $l_i = \{p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_k\}$, $l_i = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k\}$ або $l_i = \{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_k\}$ відповідно.

Наприклад, нехай задано наступний фрагмент тексту: "*Дядько Самійло не був ні професором, ні лікарем, ні інженером. Він не здатний був на високі посади*" [68].

Логіко-лінгвістична модель текстового фрагменту українською мовою матиме вигляд:

$$L^{S_1} = \neg p_1(x_1, g_1, \neg y_{11}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{12}, 0, 0, 0, 0) \& \neg p_1(x_1, g_1, \neg y_{13}, 0, 0, 0, 0),$$

$$L^{S_2} = \neg p_{21} \& p_1(x_1, g_1, y_2, q_2, 0, 0, 0),$$

$$L^{S_3} = \neg \text{був (дядько, Самійло, } \neg \text{професором, } 0, 0, 0, 0) \&$$

$$\neg \text{був (дядько, Самійло, } \neg \text{лікарем, } 0, 0, 0, 0) \&$$

$$\neg \text{був (дядько, Самійло, } \neg \text{інженером, } 0, 0, 0, 0),$$

$$L^{S_4} = \neg \text{здатний} \& \text{був (дядько, Самійло, } 0, \text{ посади, високі, } 0, 0, 0).$$

У цьому текстовому фрагменті виконано заміну суб'єктів за твердженням 2.1, проте об'єкти речень володіють відношенням частина-ціле, а відношення володіють однаковими граматичними параметрами, тому одномірний масив характеристик для речення S_1 заданого фрагменту буде мати вигляд: $l_1 = \{p_{22}, x_2, y_2\}$, а для речення S_2 : $l_2 = \{p_1, x_1, y_{11}, y_{12}, y_{13}\}$.

Правило 6. Якщо у тексті є послідовність речень природної мови $S_1, \dots, S_i, \dots, S_k$, кожному з яких поставлена у відповідність логіко-лінгвістична модель $L^{S_1}, \dots, L^{S_i}, \dots, L^{S_k}$, у яких суб'єкт, об'єкт або предмет відношення x_i, y_i, z_i , $i = \overline{1, k}$ відповідно, являються спільнокореневими словами з відношеннями інших речень p_i , то необхідно здійснити заміну: $p_i \equiv x_{i+1}$, $p_i \equiv y_{i+1}$ або $p_i \equiv z_{i+1}$ та сформувати одномірний масив зв'язків типу: $l_i = \{p_1, \dots, y_{i+1}, z_{i+2}, \dots, x_k\}$.

Наприклад, для текстового фрагменту "На Поділлі були традиційними геометричні та геометризовані рушники з малюнками жіночих фігурок, пташок, тварин. Волинські рушники оздоблювалися однією горизонтальною смугою геометричного орнаменту, виконаного нанизуванням, хрестиком..." [33] логіко-лінгвістична модель буде мати вигляд:

$$L^{S_1} = p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{11}, r_1, h_1) \& p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{12}, 0, h_1) \&$$

$$p_{11} \& p_{12}(x_1, g_{11}, y_1, 0, z_{13}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{11}, r_1, h_1) \&$$

$$p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{12}, 0, h_1) \& p_{11} \& p_{12}(x_1, g_{12}, y_1, 0, z_{13}, 0, h_1),$$

$L^S = [p_2(x_1, g_2, y_2, q_{21}, z_2, g_{11}, 0) \& p_2(x_1, g_2, y_2, q_{22}, z_2, g_{11}, 0)] \&$
 $[p'_2(z_2, 0, y'_{21}, 0, 0, 0, 0) \& p'_2(z_2, 0, y'_{22}, 0, 0, 0, 0)],$

$L^S =$ були&традиційними (рушники, геометричні, малюнками, 0, фігурок, жіночих, Поділлі)&

були&традиційними (рушники, геометричні, малюнками, 0, пташок, 0, Поділлі)&

були&традиційними (рушники, геометричні, малюнками, 0, тварин, 0, Поділлі)&

були&традиційними (рушники, геометризовані, малюнками, 0, фігурок, жіночих, Поділлі)&

були&традиційними (рушники, геометризовані, малюнками, 0, пташок, 0, Поділлі)&

були&традиційними (рушники, геометризовані, малюнками, 0, тварин, 0, Поділлі),

$L^S =$ [оздоблювалися (рушники, волинські, смугою, однією, орнаменту, геометричного, 0) &

оздоблювалися (рушники, волинські, смугою, горизонтальною, орнаменту, геометричного, 0)]&

[виконаний (орнамент, 0, нанизуванням, 0, 0, 0, 0)&

виконаний (орнамент, 0, хрестиком, 0, 0, 0, 0)].

Таким чином, одномірні масиви характеристик для заданого фрагменту будуть мати вигляд: $l_1 = \{p_2, x_2, r_2\}$ та $l_2 = \{p_{11}, x_1, g_1\}$.

Отже, з прикладів видно, що правила синтезу логіко-лінгвістичних моделей речень природної мови допомагають простежити спосіб формування логічних зв'язків у текстових фрагментах.

Правила синтезу логіко-лінгвістичних моделей речень природної мови можна застосовувати як для речень, розташованих у тексті послідовно, так і для речень з декількох абзаців. Розмірність одномірного масиву характеристик, утвореного внаслідок синтезу, відповідає кількості речень у тексті, пов'язаних між собою за змістом. Порівняння масивів характеристик кожного з речень одного з текстів з аналогічними масивами характеристик іншого тексту дає змогу знайти змістовний збіг. При порівнянні векторів характеристик важливим аспектом є правила, за якими їх було сформовано, тобто принципи утворення масивів. Таким чином, якщо елементами одномірних

масивів характеристик будуть синоніми або елементи інваріантних форм, але правила здійснення синтезу логіко-лінгвістичних моделей речень однакові, то відсоток збігу при цьому буде значно вищим, ніж якщо правила не будуть співпадати.

2.3 Абстрактні моделі логічних зв'язків між частинами текстових документів

Розрізняють два рівні побудови тексту. На першому рівні речення тексту зв'язуються за допомогою граматичних залежностей. Другий рівень необхідний для формування змістовної єдності тексту, що вбачається у змістовних зв'язках між фрагментами тексту, обумовлюючи його тип. Таким чином, кожен текст складається зі змістовних одиниць, які виконують у тексті певні функції [78].

Змістовний аналіз електронних текстових документів передбачає виявлення та встановлення логічних зв'язків між формальними моделями, якими описується зміст окремих частин – речень природної мови [79]. Виявлення певних текстових прототипів, тобто зразків текстової побудови, є однією з операцій в типологічних дослідженнях текстів [80].

Структурний принцип організації довільного тексту ґрунтується на вживанні розділових знаків і пов'язаний зі змістовним принципом організації тексту, оскільки структурно значущі частини є одночасно і логічно значимими [81]. У зв'язку з цим при змістовному аналізі тексту необхідно враховувати ряд проблем, що виникають при визначенні предикативності, перерозподілу функціональної та інформативної значимості по мірі розгортання думки.

Існування певного тексту передбачає певну форму його організації, де розгортання думки відбувається від відомого до нового шляхом використання різноманітних засобів членування текстів.

При цьому кожен текстовий фрагмент будується на основі певного типу тематичного зв'язку [82], комбінування яких забезпечує цілісність сприйняття тексту в цілому. За Ф. Данешом розрізняють п'ять типів абстрактних моделей, які лежать в основі

побудови текстів [83]. На їх основі можна вивести логічні моделі (шаблони) формування змістовних зв'язків між реченнями природної мови [84].

Нехай текстовий фрагмент t представлено у вигляді масиву логіко-лінгвістичних моделей речень $t = \{L^S, L^{S_2}, \dots, L^{S_\delta}, \dots, L^{S_{N(t)}}\}$, кожна з яких має вигляд [53]:

$$L^{S_\delta} = L_{p_\delta}^{S_\delta} (x_\delta^{(\lambda)}, g_\delta^{(\lambda)}, y_\delta^{(\lambda)}, q_\delta^{(\lambda)}, z_\delta^{(\lambda)}, r_\delta^{(\lambda)}, h_\delta^{(\lambda)}).$$

Абстрактна модель №1. Базується на послідовному розгортанні інформації, тобто коли нова інформація з попереднього речення стає темою для наступного речення, а об'єкти або предмети відношень кожного попереднього речення можуть виступати суб'єктами, об'єктами та предметами відношень кожного наступного. У такому тексті розгортання змісту відбувається від заданого до нового. Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції. Виявлення можливих комбінацій передбачає створення шаблонів масивів логіко-лінгвістичних моделей типу:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_\delta}, \dots, L^{S_{N(t)}}\} =$$

$$\{p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$$p_2(z_1, g_2, y_2, q_2, z_2, r_2, h_2) \& \dots \&$$

$$\dots \dots \dots$$

$$p_\delta(x_\delta, g_\delta, y_{\delta-1}, q_\delta, z_\delta, r_\delta, h_\delta) \& \dots \&$$

$$\dots \dots \dots$$

$$p_{N(t)}(z_{N(t)-1}, g_{N(t)}, y_{N(t)}, q_{N(t)}, z_{N(t)}, r_{N(t)}, h_{N(t)})\}.$$

Число створених шаблонів дорівнює кількості комбінацій із можливих компонентів логіко-лінгвістичних моделей, що підлягають заміні, у кожному наступному реченні, тобто $x_{\delta+1}, y_{\delta+1}$ або $z_{\delta+1}$ із кожного наступного на y_δ або z_δ попереднього. При цьому граматична характеристика часу для кожного із відношень логіко-лінгвістичних моделей однакова, тобто $h(p_1) = h(p_2) = \dots = h(p_\delta) = \dots = h(p_{N(t)})$.

Геометрична інтерпретація послідовного розгортання інформації у тексті відображена на рис. 2.1.

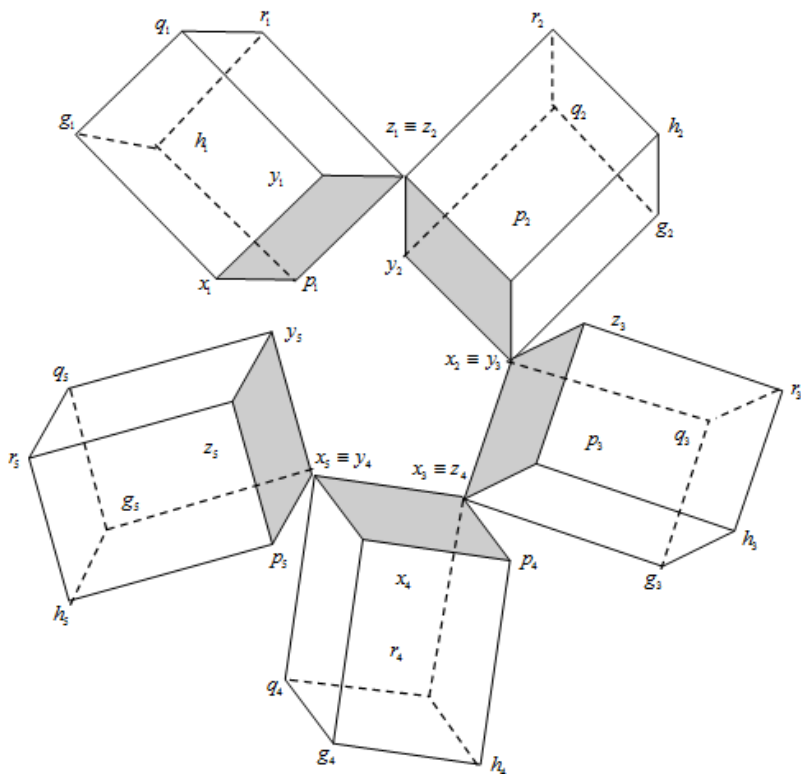


Рис. 2.1 Геометрична інтерпретація послідовного розгортання інформації у тексті

Абстрактна модель №2. Друга абстрактна модель характеризується тим, що у декількох реченнях природної мови вживається один і той самий суб'єкт, представлений у різних інтерпретаціях. Тобто одна і та сама тема пронизує увесь текст.

Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_6}, \dots, L^{S_{N(t)}}\} = \{p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$$p_2(x_1, g_2, y_2, q_2, z_2, r_2, h_2) \& \dots \&$$

$$p_\partial(x_1, g_\partial, y_\partial, q_\partial, z_\partial, r_\partial, h_\partial) \& \dots \&$$

$$p_{N(t)}(x_1, g_{N(t)}, y_{N(t)}, q_{N(t)}, z_{N(t)}, r_{N(t)}, h_{N(t)}) \& \dots \&$$

Компонентою, що буде присутня у логіко-лінгвістичних моделях всіх наступних речень фрагменту може бути x_1 , y_1 або z_1 із збереженням їх характеристик g_1 , q_1 або r_1 відповідно.

Геометричну інтерпретацію для одного із можливих варіантів вживання теми, яка повторюється у кожному реченні тексту, показано на рис. 2.2.

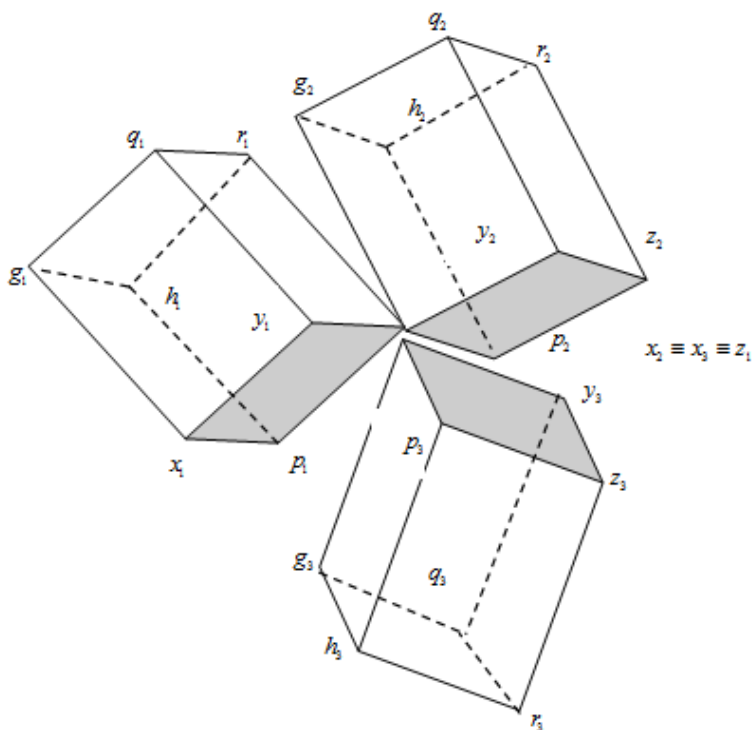


Рис. 2.2 Геометрична інтерпретація варіанту вживання теми, яка повторюється у кожному реченні тексту

На відміну від першого шаблону, граматична характеристика часу для першого та всіх наступних речень не однакова, тобто

$$h(p_2) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)}) = h(p_1) - 1 \quad \text{або}$$

$$h(p_2) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)}) = h(p_1) + 1.$$

Абстрактна модель №3. Для третьої моделі характерно, що суб'єкти логіко-лінгвістичних моделей всіх наступних речень природної мови є гіперонімами або конверсивами до суб'єкта, об'єкта, предмета відношень або характеристики відношення першого речення фрагменту, що розглядається. Таким чином, у кожній логіко-лінгвістичній моделі присутні елементи загальної тематичної направленості тексту. Внаслідок цього у текстовому фрагменті *тема розщеплена*, тобто в декількох наступних реченнях присутні початкові точки, від яких думка розвивається далі, а шаблон з масиву логіко-лінгвістичних моделей буде мати наступний тип:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} =$$

$$\{p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$$p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \& \dots \&$$

.....

$$p_{\delta}(y_{\delta-1}, g_{\delta}, y_{\delta}, q_{\delta}, z_{\delta}, r_{\delta}, h_{\delta}) \& \dots \&$$

.....

$$p_{N(t)}(y_{\delta-1}, g_{N(t)}, y_{N(t)}, q_{N(t)}, z_{N(t)}, r_{N(t)}, h_{N(t)})\}.$$

Компонентами логіко-лінгвістичних моделей, що можуть повторюватися в реченнях, від яких розвивається думка x_1 , y_1 або z_1 , при цьому суб'єктом кожного наступного речення може бути відповідно $y_{\delta-1}$ або $z_{\delta-1}$. При цьому граматична характеристика часу співпадає з граматичною характеристикою часу відношення, від якого розвивається думка, тобто $h(p_{\delta-1}) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)})$.

Геометричну інтерпретацію для одного із можливих варіантів вживання теми, що розщеплюється, показано на рис. 2.3.

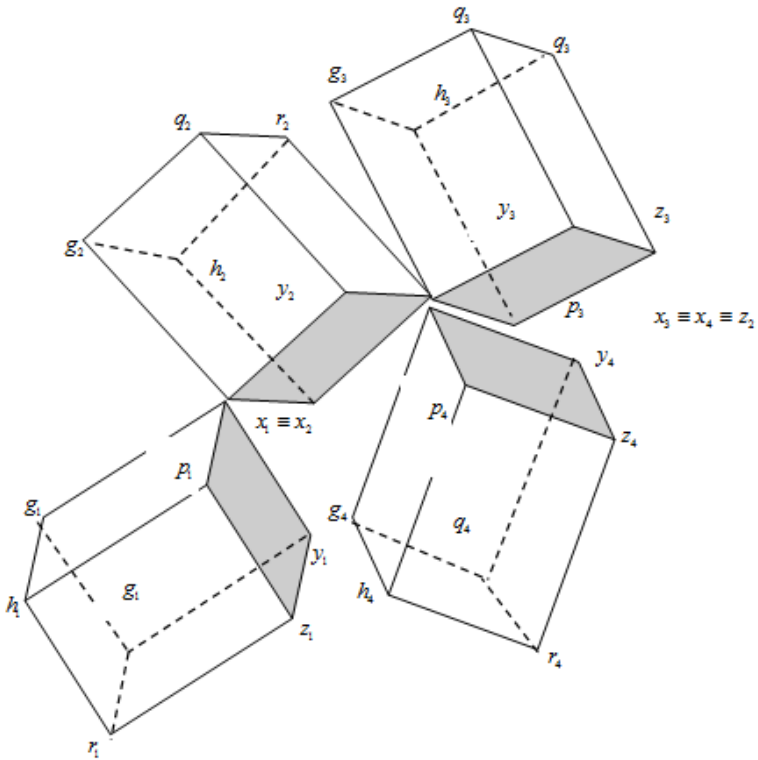


Рис. 2.3 Геометрична інтерпретація варіанту вживання розщепленої теми

Абстрактна модель №4. Для четвертої абстрактної моделі суб'єкт першого речення частково виступає суб'єктом для всіх наступних. Тобто всі речення фрагменту розкривають одну тему, основи якої закладені у першому реченні тексту, тобто має місце прогресія з похідними темами. У цьому випадку ряд окремих тем розкриває одну загальну гіпертему. Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \\ \{p_1(x_{11} \& x_{12} \& x_{13} \& \dots \& x_{1n}, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$p_2(x_{11}, g_2, y_2, q_2, z_2, r_2, h_2) \& \dots \&$

.....

$p_\theta(x_\theta, g_\theta, x_{13}, q_\theta, z_\theta, r_\theta, h_\theta) \& \dots \&$

.....

$p_{N(t)}(x_{N(t)}, g_{N(t)}, y_{N(t)}, q_{N(t)}, z_{1n}, r_{N(t)}, h_{N(t)})\}$.

Геометричну інтерпретацію для одного із можливих варіантів вживання похідних тем показано на рис. 2.4.

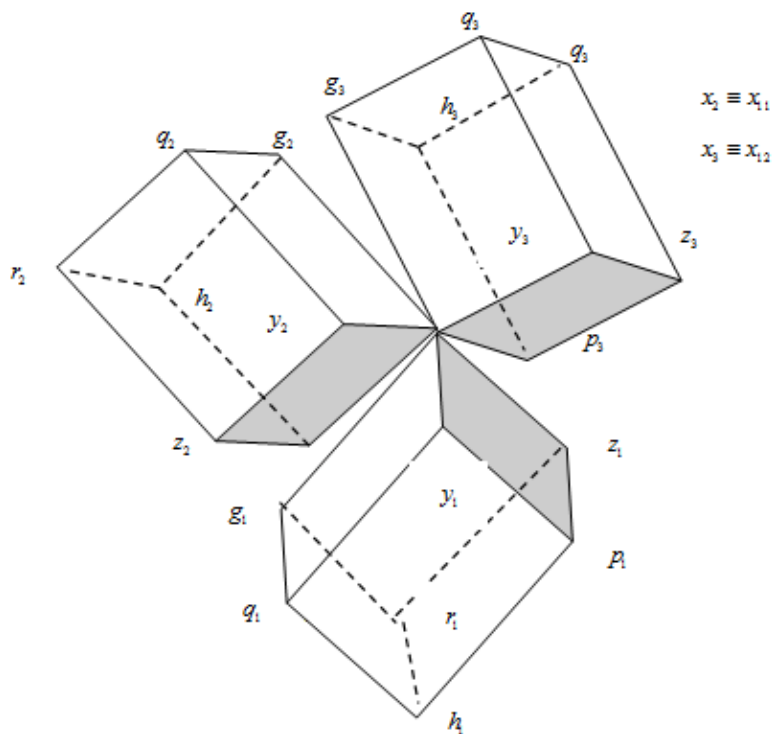


Рис. 2.4 Геометрична інтерпретація варіанту вживання похідних тем

Компонентами логіко-лінгвістичних моделей, що можуть повторюватися в реченнях, від яких розвивається думка $x_{11}, x_{12}, \dots, x_{1n}$, при цьому кожен із них може виступати суб'єктом,

об'єктом або предметом. При цьому граматична характеристика часу співпадає з граматичною характеристикою часу відношення, від якого розвивається думка, тобто $h(p_{\delta-1}) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)})$.

Абстрактна модель №5. У п'ятій моделі суб'єктами наступних речень можуть бути об'єкти, предмети та суб'єкти першого речення. У такому фрагменті передбачається відновлення відсутніх у тексті змістовних елементів шляхом їх заміни структурними компонентами логіко-лінгвістичних моделей першого речення тексту, що розглядається, тобто у тексті відбувається тематичний стрибок.

Логіко-лінгвістичні моделі речень природної мови, з'єднані за таким принципом, будуть пов'язані операцією кон'юнкції і матимуть вигляд:

$$t = \{L^{S_1}, L^{S_2}, \dots, L^{S_{\delta}}, \dots, L^{S_{N(t)}}\} =$$

$$\{p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \&$$

$$p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \& \dots \&$$

.....

$$p_{\delta}(y_{\delta-2}, g_{\delta}, y_{\delta}, q_{\delta}, z_{\delta}, r_{\delta}, h_{\delta}) \& \dots \&$$

.....

$$p_{N(t)}(x_{N(t)}, g_{N(t)}, y_{N(t)}, q_{N(t)}, z_1, r_{N(t)}, h_{N(t)})\}.$$

Граматична характеристика часу для першого та всіх наступних речень не однакова, тобто

$$h(p_2) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)}) = h(p_1) - 1 \text{ або}$$

$$h(p_2) = \dots = h(p_{\delta}) = \dots = h(p_{N(t)}) = h(p_1) + 1.$$

Геометричну інтерпретацію для одного із можливих варіантів вживання похідних тем показано на рис. 2.5.

Розгортання думки в текстах може відбуватися різними шляхами, тому запропоновані правила та шаблони можуть комбінуватися. Відповідно, шаблони можна надбудовувати.

Таким чином, послідовність розгортання змісту текстового фрагменту можна визначити за компонентами логіко-лінгвістичних моделей, зокрема, за зв'язками між їх суб'єктами, об'єктами та предметами відношень.

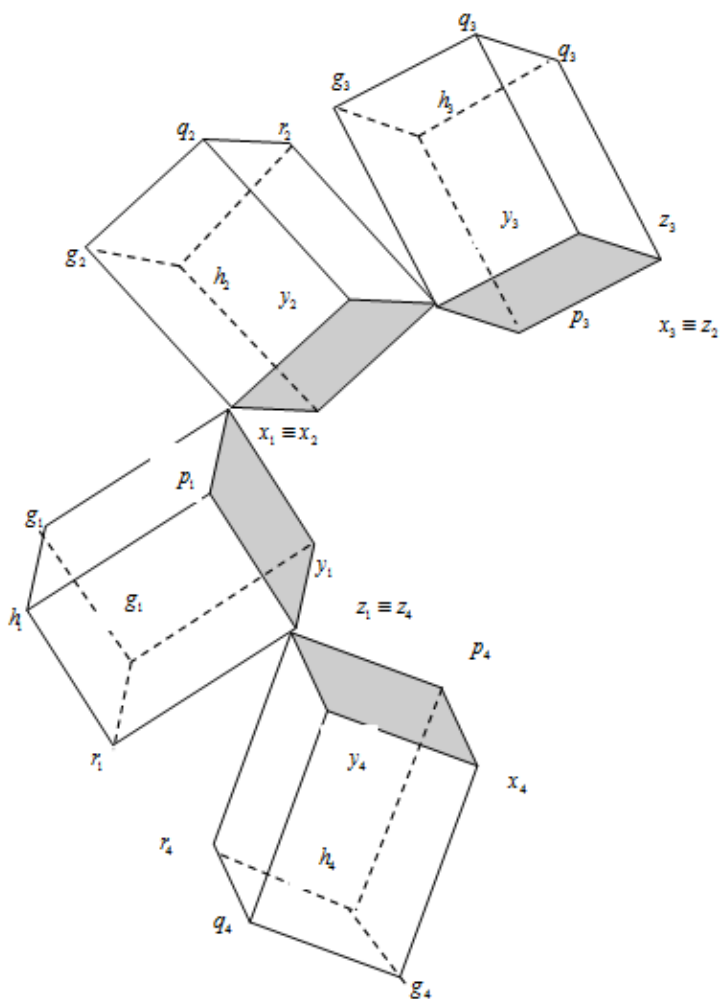


Рис. 2.5 Геометрична інтерпретація варіанту вживання тематично стрибка

Якщо побудований для текстового фрагменту масив логіко-лінгвістичних моделей буде співпадати з одним із типів наведених абстрактних моделей (шаблонів), то можна буде автоматично встановити характер розвитку думки у даному фрагменті. Сформульовані формальні умови та розроблені шаблони є базою

для встановлення змістовних зв'язків в електронних текстових документах в цілому.

Нехай необхідно визначити тип розгортання інформації для заданого текстового фрагменту: *"Цікава лекція заохочує лінивих і працюючих студентів та практикантів до подальшого навчання. Від її змісту, динаміки викладання та демонстраційних матеріалів залежить те, чи отримає знання студент. Тому викладачі повинні ретельно готуватися до лекцій, постійно удосконалювати свої знання та стежити за сучасними тенденціями у науці. Провідні висококваліфіковані викладачі регулярно включають нові сучасні дисципліни до навчальних планів"*.

По-перше, необхідно побудувати логіко-лінгвістичну модель текстового фрагменту, шляхом побудови логіко-лінгвістичних моделей речень природної мови. Логіко-лінгвістична модель (1.25) уже наводилася у першому розділі монографії:

$$\begin{aligned}
 t &= \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} = \\
 &\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \& \\
 &p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0), \\
 &L^{S_2} = [p_2(x_2, 0, y_{21}, q_2, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \& \\
 &p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(x'_2, 0, y'_2, 0, 0, 0, 0), \\
 &L^{S_3} = p_{31} \& p_{32}(x_3, 0, y_3, 0, 0, 0, h_{31}) \& \\
 &p_{31} \& p_{33}(x_3, 0, y_3, q_{31}, 0, 0, h_{32}) \& p_{31} \& p_{34}(x_3, 0, y_{33}, q_{32}, z_3, 0, 0), \\
 &L^{S_4} = p_4(x_4, g_{41}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_4, g_{41}, y_4, q_{42}, z_4, r_4, h_4) \& \\
 &p_4(x_4, g_{42}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_4, g_{42}, y_4, q_{42}, z_4, r_4, h_4)\}.
 \end{aligned}$$

Другим кроком є синтез логіко-лінгвістичних моделей речень, внаслідок чого елементи логіко-лінгвістичних моделей другого, третього і четвертого речень заданого текстового фрагменту будуть замінені на відповідні елементи логіко-лінгвістичної моделі першого речення $q_2 \equiv x_1$, $x'_2 \equiv y_{11}$, $y_3 \equiv x_1$, $y_3 \equiv y'_2$, $x_4 \equiv x_3$, $q_{42} \equiv q_{32}$, тобто:

$$\begin{aligned}
 t &= \{L^{S_1}, L^{S_2}, \dots, L^{S_8}, \dots, L^{S_{N(t)}}\} = \{L^{S_1}, L^{S_2}, L^{S_3}, L^{S_4}\} = \\
 &\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \& \\
 &p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0),
 \end{aligned}$$

$$\begin{aligned}
L^{S_2} &= [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \& \\
& p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0), \\
L^{S_3} &= p_{31} \& p_{32}(x_3, 0, x_1, 0, 0, 0, h_{31}) \& \\
& p_{31} \& p_{33}(x_3, 0, y'_2, q_{31}, 0, 0, h_{32}) \& p_{31} \& p_{34}(x_3, 0, y_{33}, q_{32}, z_3, 0, 0), \\
L^{S_4} &= p_4(x_3, g_{41}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{41}, y_4, q_{32}, z_4, r_4, h_4) \& \\
& p_4(x_3, g_{42}, y_4, q_{41}, z_4, r_4, h_4) \& p_4(x_3, g_{42}, y_4, q_{32}, z_4, r_4, h_4)\}.
\end{aligned}$$

Наступним етапом є визначення граматичних характеристик часу для відношень кожного з трьох речень:

$$h(p_1) = h(p_2) = h(p_{31} \& p_{32}) = h(p_{31} \& p_{33}) =$$

$$h(p_{31} \& p_{34}) = h(p_4) = \text{'теперішній'}.$$

Заключний етап – це зв'язка отриманого формального опису з запропонованими шаблонами. Рівність граматичних характеристик відношень, а також заміни, здійснені під час синтезу, дають змогу віднести заданий текстовий фрагмент до третього типу шаблону, що відповідає розщепленій темі.

III. АНАЛІЗ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ

*"Будь-яке знання починається з
сумніву і закінчується вірою."
Марія фон Ебнер-Ешенбах*

3.1 Критерії аналізу логіко-лінгвістичних моделей речень природної мови

Побудова логіко-лінгвістичних моделей речень природної мови [85] відбувається за наступними принципами:

- 1) визначення кількості простих атомарних предикатів v^S ;
- 2) визначення типу окремої форми логіко-лінгвістичної моделі речення природної мови $W(S)$;
- 3) фіксація значень елементів кортежу логічних операцій $O(S)$ речення S :

$$O(S) = [o_e(S) \mid e = \overline{1, h}],$$

де h – загальна кількість логічних операцій, наявних у реченні S ;

4) визначення множини відношень P^S між суб'єктами та об'єктами речення S , що фактично містить дії, стани або властивості суб'єкта;

5) фіксація множини предикатних констант H_p^S , що містять характеристики відношень між суб'єктами та об'єктами у реченні S ;

6) визначення множини суб'єктів $X_p^S(h)$, що входять до речення S та пов'язані з об'єктами речення p -им відношенням, яке володіє h -ю характеристикою;

7) фіксація множини предикатних констант $G_p^S(x, h)$, що містять характеристики (параметри) суб'єктів, які входять до речення S ;

8) визначення множини об'єктів $Y_p^S(x, g, h)$, що входять до речення S та пов'язані з суб'єктами речення $x \in X_p^S(h)$ з

характеристиками $g \in G_p^S(x, h)$ p -им відношенням, яке володіє h -ю характеристикою;

9) фіксація множини предикатних констант $Q_p^S(x, g, y, h)$, що містять характеристики (параметри) об'єктів з характеристиками $q \in Q_p^S(x, g, y, h)$, які входять до речення S ;

10) визначення множини предметів $Z_p^S(x, g, y, q, h)$ p -го відношення, яке володіє h -ю характеристикою, між суб'єктами $x \in X_p^S(h)$ з характеристиками $g \in G_p^S(x, h)$ та об'єктами $y \in Y_p^S(x, g, h)$ з характеристиками $q \in Q_p^S(x, g, y, h)$;

11) фіксація множини предикатних констант $R_p^S(x, g, y, q, z, h)$, що містить характеристики (параметри) предметів p -го відношення, яке володіє h -ю характеристикою, між суб'єктами $x \in X_p^S(h)$ з характеристиками $g \in G_p^S(x, h)$ та об'єктами $y \in Y_p^S(x, g, h)$;

12) формування логічного виразу L^S , шляхом підстановки у нього змінних, отриманих за попередніми принципами побудови логіко-лінгвістичних моделей відповідно до кількості елементів зафіксованих множин.

Логіко-лінгвістичні моделі, побудовані за описаними вище принципами, є інтерпретацією семантико-синтаксичної структури речень природної мови, так як будуються на основі наявних у реченні означальних (між означенням і підметом, означенням і додатком), обставинних (між присудком і обставинами), додаткових (між присудком і додатком), присудкових (між підметом і присудком) зв'язків [82].

Керуючись принципами побудови логіко-лінгвістичних моделей речень природної мови, можна сформулювати критерії для їх аналізу.

Критерій 1. Кількість простих предикатів в логіко-лінгвістичній моделі речення природної мови L^S .

Цей критерій дає змогу визначити інформацію про тип та склад речення природної мови.

Якщо кількість простих предикатів в логіко-лінгвістичній моделі рівна одному $v^S = 1$, тобто $L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1)$, то вона описує просте речення природної мови без однорідних членів.

Якщо кількість простих предикатів більше одного $v^S > 1$ і при цьому прості предикати відрізняються однією або декількома (одна з яких відповідає за характеристику) компонентами, то логіко-лінгвістична модель описує просте речення з однорідними членами. Наприклад, логіко-лінгвістична модель

$$L^S = p_{11}(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p_{12}(x_1, g_1, y_1, q_1, z_1, r_1, h_1)$$

описує речення з однорідними присудками, а модель

$$L^S = p_1(x_{11}, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_{11}, z_1, r_1, h_1) \& p_1(x_{11}, g_1, y_1, q_{12}, z_1, r_1, h_1) \& p_1(x_{12}, g_1, y_1, q_{12}, z_1, r_1, h_1)$$

описує просте поширене речення з однорідними підметами та однорідними означеннями, що характеризують об'єкт відношення.

Якщо кількість простих предикатів більше одного $v^S > 1$ і при цьому майже всі предикатні змінні та константи різні, то логіко-лінгвістична модель описує складне речення, яке в подальшому необхідно проаналізувати за ще декількома критеріями. Наприклад,

$$L^S = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \& p'_1(x'_1, g'_1, y'_1, q'_1, z'_1, r'_1, h'_1).$$

Критерій 2. Потужність множин предикатних змінних та констант логіко-лінгвістичної моделі.

Потужність множин предикатних змінних та констант впливає, в першу чергу, на кількість простих предикатів v^S , що обчислюється за формулами (1.5) – (1.8).

Також потужність множин визначає тип окремої форми логіко-лінгвістичної моделі.

Якщо просте речення природної мови містить лише одну частину, що має закінчений зміст, тобто $|P^S| = 1$; $|H^S_p| = 1$; $|X^S_p(h)| = 1$; $|G^S_p(x, h)| = 1$; $|Y^S_p(x, g, h)| = 1$; $|Q^S_p(x, g, y, h)| = 1$; $|Z^S_p(x, g, y, q, h)| = 1$; $|R^S_p(x, g, y, q, z, h)| = 1$.

Якщо множина H_p^S характеристик хоча б одного p -го відношення у реченні S між суб'єктом x та об'єктом y містить більше одного елемента і $|P^S|=1$; $|X_p^S(h)|=1$; $|G_p^S(x,h)|=1$; $|Y_p^S(x,g,h)|=1$; $|Q_p^S(x,g,y,h)|=1$; $|Z_p^S(x,g,y,q,h)|=1$; $|R_p^S(x,g,y,q,z,h)|=1$, то у реченні природної мови йдеться про різні обставини дії.

Якщо множина суб'єктів $X_p^S(h)$, що пов'язані з об'єктами речення p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів $|P^S|=1$; $|H_p^S|=1$; $|G_p^S(x,h)|=1$; $|Y_p^S(x,g,h)|=1$; $|Q_p^S(x,g,y,h)|=1$; $|Z_p^S(x,g,y,q,h)|=1$; $|R_p^S(x,g,y,q,z,h)|=1$, то речення просте з однорідними підметами.

Якщо множина об'єктів $Y_p^S(x,g,h)$, що пов'язані з суб'єктами речення p -им відношенням, яке володіє h -ю характеристикою, містить декілька елементів і $|P^S|=1$; $|H_p^S|=1$; $|X_p^S(h)|=1$; $|G_p^S(x,h)|=1$; $|Q_p^S(x,g,y,h)|=1$; $|Z_p^S(x,g,y,q,h)|=1$; $|R_p^S(x,g,y,q,z,h)|=1$, то це речення з однорідними додатками.

Якщо потужність множин $|P^S| \geq 2$ і $|X_p^S(h)| \geq 2$, то речення природної мови складне, а v^S завжди більше одиниці.

Тобто, фактично, визначення потужності множин компонентів логіко-лінгвістичної моделі речення природної мови визначає тип окремої форми логіко-лінгвістичної моделі.

Критерій 3. Логічні операції, що використані в логіко-лінгвістичній моделі речення природної мови L^S . Так:

– операція кон'юнкції «&» свідчить про наявність у реченні послідовності певних дій, перерахування, ствердження істинності простих предикатів, що з'єднані цією операцією;

– операція диз'юнкції « \vee » свідчить про наявність у реченні послідовності подій, що протиставляються, зіставляються, взаємно виключаються або чергуються;

– операція імплікації « \rightarrow » використовується між реченнями природної мови, у яких йдеться про певну подію та наслідок (результат) її виконання.

Таким чином, логічні операції дають змогу проаналізувати послідовність розмірковувань, про які йдеться у реченні природної мови.

Критерій 4. Наявність тотожних компонент у логіко-лінгвістичній моделі речення природної мови із різних множин предикатних змінних та констант.

Якщо тотожні елементи знайдені в різних множинах однієї і тієї ж логіко-лінгвістичної моделі, то це свідчить про наявність у реченні синонімів, займенників або вказівних слів, що дозволили зробити в моделі синтез [64]. Аналіз того, в яких саме множинах предикатних змінних або констант наявні такі елементи дасть можливість говорити про те, які прийоми змістовної заміни були зроблені у реченні природної мови.

Наприклад, якщо збігаються елементи множини суб'єктів та предметів відношень у двох різних простих предикатах, то розглядається складне означальне речення, у якому останнє слово першого речення виступає суб'єктом другого, залежного речення, виражене займенником.

Застосуємо сформульовані критерії аналізу для визначення та відновлення текстової інформації з речення природної мови, представленого такими логіко-лінгвістичними моделями:

$$L^S = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \&$$

$$p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0),$$

L^S = заохочує (лекція, цікава, студентів, лінивих, навчання, подальшого, 0) &

заохочує (лекція, цікава, практикантів, лінивих, навчання, подальшого, 0) &

заохочує (лекція, цікава, студентів, працьовитих, навчання, подальшого, 0) &

заохочує (лекція, цікава, практикантів, працьовитих, навчання, подальшого, 0).

Логіко-лінгвістична модель послідовно піддається аналізу за кожним із сформульованих критеріїв.

За першим критерієм визначається кількість простих предикатів у логіко-лінгвістичній моделі речення природної мови,

у даному випадку $v^{S_1} = 4$, відношення та суб'єкти відношень у всіх атомарних предикатах однакові, що говорить про те, що речення просте.

Тип речення також підтверджується згідно другого критерію аналізу, за яким встановлено потужності множин: $|P^S| = 1$;
 $|H_p^S| = 1$; $|X_p^S(h)| = 1$; $|G_p^S(x, h)| = 1$; $|Y_p^S(x, g, h)| = 2$;
 $|Q_p^S(x, g, y, h)| = 2$; $|Z_p^S(x, g, y, q, h)| = 1$; $|R_p^S(x, g, y, q, z, h)| = 1$. Тобто потужність множини об'єктів та їх відношень рівна двом, що в сукупності і дало загальну кількість простих предикатів $v^{S_1} = 4$. Такі характеристики підпадають під тип 12, що відображає типове сполучення членів речення природної мови та описує просте речення з однорідними означеннями і додатками (1.20), внаслідок чого кількість частин, що мають закінчений зміст, дорівнює сумі потужностей множин об'єктів відношень та їх характеристик:
 $v^S = |Y_p^S(x, g, h)| + |Q_p^S(x, g, y, h)|$.

Отже, ще раз підтверджено, що речення природної мови є простим.

За третім критерієм аналізу визначено, що у реченні йдеться про послідовність певних дій, перерахування, ствердження істинності простих предикатів, так як в логіко-лінгвістичній моделі використано лише операцію кон'юнкції.

Після застосування четвертого критерію аналізу видно, що у реченні відсутні тотожні компоненти у логіко-лінгвістичній моделі речення природної мови із різних множин предикатних змінних та констант. Це говорить про те, що для логіко-лінгвістичної моделі не застосовувалися правила синтезу.

Якщо зіставити результати, отримані в результаті застосування критеріїв аналізу, та оригінальний текст: *«Цікава лекція захоочує лінивих і працюючих студентів та практикантів до подальшого навчання»*, то видно, що критерії аналізу дали достатньо точні показники щодо структури речення природної мови.

Основні принципи формування логіко-лінгвістичних моделей сформульовані на основі граматичної організації речень природної мови [82]. Це дало можливість описати таких чотири критерії аналізу логіко-лінгвістичних моделей, за допомогою яких можна визначити інформацію про тип та склад речення природної

мови, а також проаналізувати послідовність розмірковувань, про які йдеться у реченні. Описані критерії аналізу логіко-лінгвістичних моделей необхідні для побудови формальних моделей електронних текстових документів і являються необхідними складовими процесу їх автоматичного лінгвістичного аналізу.

3.2 Умови тотожності логіко-лінгвістичних моделей простих речень природної мови

Для відновлення текстової інформації з логіко-лінгвістичних моделей шляхом їх аналізу необхідно володіти формальними правилами та умовами здійснення заміन. Такі заміни можуть робитися внаслідок синтезу логіко-лінгвістичних моделей, а також внаслідок вживання різноманітних засобів когезії [81 – 82].

Можна сформулювати ряд умов, за яких речення природної мови будуть формально вважатися однаковими за змістом. Для знаходження синонімічних конструкцій не має значення порядок розгляду речень, тобто умови можна застосовувати і у зворотню сторону, коли першим буде вважатися речення S_2 , а другим – S_1 .

Сформульовано умови тотожності, у яких виділено твердження, при виконанні яких прості речення природної мови будуть вважатися тотожними за змістом [83].

Введемо деякі позначення. Нехай H_k – слово речення S_1 , де $k = \overline{1, n}$, n – загальна кількість слів у реченні S_1 ; H_l – слово речення S_2 , де $l = \overline{1, m}$, m – загальна кількість слів у реченні S_2 .

Умова 1. Пов'язана з існуванням у флективних мовах такого явища, як синтаксична деривація, ще називається транспозиційною деривацією (перехід слова з однієї частини мови до іншої без зміни його лексичного значення) [85].

Твердження 1.1. Якщо у реченні S_1 існує послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$ та $Z_{k+2}(H_{k+2})$ відповідно, і такій послідовності можна поставити у відповідність послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2} \rightarrow H_{l+3}$ речення S_2 з відповідними характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$, $Z_{l+2}(H_{l+2})$ та $Z_{l+3}(H_{l+3})$, таку що $H_l \equiv H_k$, $H_{l+1} \neq H_{k+1}$,

$\widehat{H}_{l+2} \equiv \widehat{H}_{k+1}$, $H_{l+3} \equiv H_{k+2}$, то для атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$p_1 \equiv p_2, p_1 \in P^S, p_2 \in P^S$$

при $Z_k(H_k) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}$;

$Z_{k+1}(H_{k+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}$;

$Z_{k+2}(H_{k+2}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, t_{k+2}, h_{k+2}, l_{k+2}, ch_{k+2}\}$;

$Z_l(H_l) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}$;

$Z_{l+1}(H_{l+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}$;

$Z_{l+2}(H_{l+2}) = \{1, g_{l+2} \neq 1, n_{l+2}, k2_{l+2}, 0, 0, l_{l+2}, 2\}$;

$Z_{l+3}(H_{l+3}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, t_{k+2}, h_{k+2}, l_{k+2}, ch_{k+2}\}$;

$\widehat{H}_{l+2} \in \widehat{R}_q$ та $\widehat{H}_{k+1} \in \widehat{R}_q$ – слова, що входять до одного і того ж ряду $\widehat{R}_q \in R$ спільнокореневих слів із множини R .

Наприклад, розглянемо два речення природної мови «Традиція розпочинається у минулому столітті» та «Традиція бере початок у минулому столітті».

Логіко-лінгвістичні моделі речень мають вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_1} = \text{розпочинається (традиція, 0, столітті, минулому, 0, 0, 0)},$$

$$L^{S_2} = p_{21} \& p_{22}(x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \text{бере\&початок (традиція, 0, столітті, минулому, 0, 0, 0)}.$$

У цьому разі суб'єкти та об'єкти логіко-лінгвістичних моделей однакові, а предикати відповідають схемі: $p_2 \equiv p_{21} \& \hat{p}_1$, де \hat{p}_1 – слово, спільнокоренеve з p_1 .

Твердження 1.2. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$ та $Z_{k+2}(H_{k+2})$ відповідно, і такій послідовності можна поставити у відповідність послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2}$ речення S_2 з відповідними характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$ та $Z_{l+2}(H_{l+2})$, таку що $H_l \equiv H_k$, $\widehat{H}_{l+1} \equiv \widehat{H}_{k+1}$, $H_{l+2} \equiv H_{k+2}$, то для атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$p_1 \equiv p_2, p_1 \in P^S, p_2 \in P^S$$

при $Z_k(H_k) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}$;
 $Z_{k+1}(H_{k+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}$;
 $Z_{k+2}(H_{k+2}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $Z_l(H_l) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}$;
 $Z_{l+1}(H_{l+1}) = \{6, 1, n_k, k2_k, 0, 0, l_k, 2\}$;
 $Z_{l+2}(H_{l+2}) = \{cm_{k+2}, g_{r+2} \neq g_{k+2}, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $\widehat{H}_{l+1} \in \widehat{R}_q$ та $\widehat{H}_{k+1} \in \widehat{R}_q$ – слова, що входять до одного і того ж ряду $\widehat{R}_q \in R$ спільнокореневих слів із множини R .

Для двох речень природної мови «Робота дуже цікавить викладача» та «Робота дуже цікава викладачу» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, 0, 0, h_1),$$

$$L^{S_1} = \text{цікавить (робота, 0, викладача, 0, 0, 0, дуже)},$$

$$L^{S_2} = p_2(x_1, 0, y_1, 0, 0, 0, h_1),$$

$$L^{S_2} = \text{цікава (робота, 0, викладачу, 0, 0, 0, дуже)}.$$

У цьому разі суб'єкти та об'єкти логіко-лінгвістичних моделей однакові, а предикати відповідають, як і у твердженні 1.1, схемі: $p_2 \equiv \hat{p}_1$, де \hat{p}_1 – слово, спільнокореневе з p_1 .

Твердження 1.3. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$ та $Z_{k+2}(H_{k+2})$ відповідно, а у другому реченні S_2 знайдено послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2} \rightarrow H_{l+3} \rightarrow H_{l+4}$ характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$, $Z_{l+2}(H_{l+2})$, $Z_{l+3}(H_{l+3})$ та $Z_{l+4}(H_{l+4})$, таку що
 $Z_k(H_k) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}$;
 $Z_{k+1}(H_{k+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}$;
 $Z_{k+2}(H_{k+2}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $Z_l(H_l) = \{9, g_{r+1}, 0, 0, 0, 0, 0\}$;
 $Z_{l+1}(H_{l+1}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $Z_{l+2}(H_{l+2}) = \{cm_{l+2}, 1, n_{l+2}, k2_{l+2}, 0, 0, l_{l+2}, 2\}$;
 $Z_{l+3}(H_{l+3}) = \{9, g_{l+4}, 0, 0, 0, 0, 0\}$;

$$Z_{l+4}(H_{l+4}) = \{cm_k, g_{l+4} \neq 1, n_k, k2_k, 0, 0, l_k, 3\};$$

$$H_{l+1} \equiv H_{k+2}, \widehat{H}_{l+3} \equiv \widehat{H}_{k+2}, H_{l+4} \equiv H_k;$$

$$x_2 \in X_p^S(h) \text{ і } x_2 \equiv 0;$$

$\widehat{H}_{l+3} \in \widehat{R}_q$ та $\widehat{H}_{l+2} \in \widehat{R}_q$ – слова, що входять до одного і того ж ряду $\widehat{R}_q \in R$ спільнокорених слів із множини R , то в атомарних предикатах $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$p_1 \equiv p_2, p_1 \in P^S, p_2 \in P^S,$$

$$x_1 \equiv z_2, x_1 \in X_p^S(h), z_2 \in Z_p^S(x, g, y, q, h),$$

$$y_1 \equiv y_2, y_1 \in Y_p^S(x, g, h), y_2 \in Y_p^S(x, g, h).$$

Для двох речень природної мови «Робота дуже цікавить викладача» та «У викладача цікавість до роботи» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, 0, 0, h_1),$$

$$L^{S_1} = \text{цікавить (робота, 0, викладача, 0, 0, 0, дуже)},$$

$$L^{S_2} = p_2(0, 0, y_1, 0, x_1, 0, 0),$$

$$L^{S_2} = \text{цікавість (0, 0, викладача, 0, роботи, 0, 0)}.$$

У логіко-лінгвістичній моделі другого речення відсутній суб'єкт, предметом відношень другого речення виступає суб'єкт відношення першого речення, предикати відповідають, як і у твердженні 1.1 та твердженні 1.2, схемі: $p_2 \equiv \hat{p}_1$, де \hat{p}_1 – слово, спільнокореневе з p_1 .

Умова 2. Друга умова тотожності випливає з того, що синонімічними за змістом вважаються речення, у яких змінені синтаксичні позиції лексичних морфем, а денотативне значення залишається незмінним. Тобто предикати таких двох речень відображають ситуацію одного і того ж класу, у яких зміщено логічний акцент, проте зберігається зміст [86].

Твердження 2.1. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$ та $Z_{k+2}(H_{k+2})$ відповідно, і такій послідовності можна поставити у відповідність послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2}$ речення S_2 з відповідними характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$ та $Z_{l+2}(H_{l+2})$,

таку що $H_l \equiv H_{k+2}$, $H_{l+1} \equiv H_{k+1}$, $H_{l+2} \equiv H_k$, то для атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$\begin{aligned} p_1 &\equiv p_2, p_1 \in P^S, p_2 \in P^S, \\ x_1 &\equiv y_2, x_1 \in X_p^S(h), y_2 \in Y_p^S(x, g, h), \\ y_1 &\equiv x_2, y_1 \in Y_p^S(x, g, h), x_2 \in X_p^S(h) \end{aligned}$$

$$\begin{aligned} \text{при } Z_k(H_k) &= \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}; \\ Z_{k+1}(H_{k+1}) &= \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}; \\ H_{k+2}(S_{k+2}) &= \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}; \\ Z_l(H_l) &= \{cm_{k+2}, 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 1\}; \\ Z_{l+1}(H_{l+1}) &= \{5, 0, n_l, k2_l, t_{l+1}, h_{l+1}, l_l, 2\}; \\ Z_{l+2}(H_{l+2}) &= \{cm_k, g_{l+1} \neq 1, n_k, k2_k, 0, 0, l_k, 3\}. \end{aligned}$$

Розглянемо два речення природної мови «*Експерти використовують методи*» та «*Методи використовуються експертами*». Логіко-лінгвістичні моделі речень мають вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, 0, 0, 0),$$

$$L^{S_1} = \text{використовують (експерти, 0, методи, 0, 0, 0, 0)},$$

$$L^{S_2} = p_1(y_1, 0, x_1, 0, 0, 0, 0),$$

$$L^{S_2} = \text{використовуються (методи, 0, експертами, 0, 0, 0, 0)}.$$

У логіко-лінгвістичній моделі другого речення суб'єктом виступає об'єкт першого речення, а об'єктом – суб'єкт першого речення.

Твердження 2.2. Якщо у реченні S_1 існує послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$ та $Z_{k+2}(H_{k+2})$ відповідно, і такій послідовності можна поставити у відповідність послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2} \rightarrow H_{l+3}$ речення S_2 з відповідними характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$, $Z_{l+2}(H_{l+2})$ та $Z_{l+3}(H_{l+3})$, таку що $H_l \equiv H_{k+2}$, $H_{l+1} \equiv H_{k+1}$, $H_{l+3} \equiv H_k$, то для атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$\begin{aligned} p_1 &\equiv p_2, p_1 \in P^S, p_2 \in P^S, \\ x_1 &\equiv y_2, x_1 \in X_p^S(h), y_2 \in Y_p^S(x, g, h); \end{aligned}$$

$$y_1 \equiv x_2, y_1 \in Y_p^S(x, g, h), x_2 \in X_p^S(h)$$

$$\begin{aligned} \text{при } Z_k(H_k) &= \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\}; \\ Z_{k+1}(H_{k+1}) &= \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\}; \\ Z_{k+2}(H_{k+2}) &= \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}; \\ Z_l(H_l) &= \{cm_{k+2}, 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 1\}; \\ Z_{l+1}(H_{l+1}) &= \{5, 0, n_l, k2_l, t_{l+1}, h_{l+1}, l_l, 2\}; \\ Z_{l+2}(H_{l+2}) &= \{9, g_{l+3}, 0, 0, 0, 0, 0, 0\}; \\ Z_{l+3}(H_{l+3}) &= \{cm_k, g_{l+3} \neq 1, n_k, k2_k, 0, 0, l_k, 3\}. \end{aligned}$$

Для двох речень природної мови «*Магніт притягує залізо*» та «*Залізо притягується до магніту*» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, 0, 0, 0),$$

$$L^{S_1} = \text{притягує (магніт, 0, залізо, 0, 0, 0, 0)},$$

$$L^{S_2} = p_1(y_1, 0, x_1, 0, 0, 0, 0),$$

$$L^{S_2} = \text{притягується (залізо, 0, магніту, 0, 0, 0, 0)}.$$

Аналіз логіко-лінгвістичних моделей у двох описаних вище ситуаціях дає змогу зробити висновок про те, що якщо предикатна змінна (суб'єкт) одного речення дорівнює предикатній змінній (аргументу) другого речення, і навпаки, а також рівні предикати, то такі речення тотожні за змістом.

Твердження 2.3. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2} \rightarrow H_{k+3}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$, $Z_{k+2}(H_{k+2})$ та $Z_{k+3}(H_{k+3})$ відповідно, а у другому реченні S_2 знайдено послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2} \rightarrow H_{l+3} \rightarrow H_{l+4}$ характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$, $Z_{l+2}(H_{l+2})$, $Z_{l+3}(H_{l+3})$ та $Z_{l+4}(H_{l+4})$, таку що

$$Z_k(H_k) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\};$$

$$Z_{k+1}(H_{k+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\};$$

$$Z_{k+2}(H_{k+2}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\};$$

$$Z_{k+3}(H_{k+3}) = \{cm_{k+3}, g_{k+3} \neq 1, n_{k+3}, k2_{k+3}, 0, 0, l_{k+3}, 3\};$$

$$Z_l(H_l) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\};$$

$$Z_{l+1}(H_{l+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\};$$

$Z_{l+2}(H_{l+2}) = \{cm_{k+3}, (g_{l+2} \neq 1) \& (g_{l+2} \neq g_{k+3}), n_{k+3}, k2_{k+3}, 0, 0, l_{k+3}, 3\}$;
 $Z_{l+3}(H_{l+3}) = \{9, g_{l+4}, 0, 0, 0, 0, 0\}$;
 $Z_{l+4}(H_{l+4}) = \{cm_{k+2}, (g_{l+4} \neq 1) \& (g_{l+4} \neq g_{k+2}), n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $H_l \equiv H_k$, $H_{l+1} \equiv H_{k+1}$, $H_{l+2} \equiv H_{k+3}$, $H_{l+4} \equiv H_{k+2}$, то для аргументів атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$y_1 \equiv z_2, y_1 \in Y_p^S(x, g, h), z_2 \in Z_p^S(x, g, y, q, h);$$

$$z_1 \equiv y_2, z_1 \in Z_p^S(x, g, y, q, h), y_2 \in Y_p^S(x, g, h).$$

Для двох речень природної мови «Вчені постачають бібліотеки книжками» та «Вчені постачають книжки до бібліотек» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, 0, z_1, 0, 0),$$

$$L^{S_1} = \text{постачають (вчені, 0, бібліотеки, 0, книжками, 0, 0)},$$

$$L^{S_2} = p_1(x_1, 0, z_1, 0, y_1, 0, 0),$$

$$L^{S_2} = \text{постачають (вчені, 0, книжки, 0, бібліотек, 0, 0)}.$$

Умова 3. Третя умова тотожності пов'язана з існуванням у природній мові трансформацій, при яких змінюється прийменниково-відмінкова форма не більше, ніж у одного іменника, внаслідок чого такі синтаксичні конструкції вважаються синонімічними [15].

Твердження 3.1. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1}$ з характеристиками $Z_k(H_k)$ та $Z_{k+1}(H_{k+1})$ відповідно, і такій послідовності можна поставити у відповідність послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2}$ речення S_2 з відповідними характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$ та $Z_{l+2}(H_{l+2})$, таку що $\hat{H}_l \equiv \hat{H}_k$, $\hat{H}_{l+2} \equiv \hat{H}_{k+1}$, то для атомарних предикатів $L_p^{S_1}(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$p_1 \equiv p_2, p_1 \in P^S, p_2 \in P^S$$

при $Z_k(H_k) = \{5, 0, n_k, k2_k, t_k, h_k, l_k, 2\}$;

$$Z_{k+1}(H_{k+1}) = \{cm_{k+1}, g_{k+1} \neq 1, n_{k+1}, k2_{k+1}, 0, 0, l_{k+1}, ch_{k+1}\};$$

$$Z_l(H_l) = \{1, g_l \neq 1, n_l, k2_l, 0, 0, l_l, 2\};$$

$$Z_{l+1}(H_{l+1}) = \{9, g_{l+2}, 0, 0, 0, 0, 0\};$$

$$Z_{l+2}(H_{l+2}) = \{cm_{k+1}, g_{l+2} \neq 1, n_{k+1}, k2_{k+1}, 0, 0, l_{k+1}, ch_{l+2}\};$$

$\widehat{H}_l \in \widehat{R}_q$ та $\widehat{H}_k \in \widehat{R}_q$ – слова, що входять до одного і того ж ряду $\widehat{R}_q \in R$ спільнокорених слів із множини R , а $\widehat{H}_{l+2} \in \widehat{R}_{qm}$ та $\widehat{H}_{k+1} \in \widehat{R}_{qm}$ – слова, що входять до одного і того ж ряду $\widehat{R}_{qm} \in R$ спільнокорених слів із множини R .

Для двох речень природної мови «Цукор важить три кілограми» та «Цукор вагою в три кілограми» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^{S_1} = p_1(x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_1} = \text{важить (цукор, 0, кілограми, три, 0, 0, 0)},$$

$$L^{S_2} = p_2(x_1, 0, y_1, q_1, 0, 0, 0),$$

$$L^{S_2} = \text{вагою (цукор, 0, кілограми, три, 0, 0, 0)}.$$

Суб'єкт та об'єкт відношення у першому та другому реченні однакові, а предикати відповідають схемі: $p_2 \hat{=} \hat{p}_1$, де \hat{p}_1 – слово, спільнокоренеve з p_1 .

Умова 4. Четверта умова тотожності базується на використанні у реченнях природної мови конверсивів, які описують одну й ту саму ситуацію з різних точок зору.

Твердження 4.1. Якщо у реченні S_1 знайдено послідовність слів $H_k \rightarrow H_{k+1} \rightarrow H_{k+2} \rightarrow H_{k+3}$ з характеристиками $Z_k(H_k)$, $Z_{k+1}(H_{k+1})$, $Z_{k+2}(H_{k+2})$ та $Z_{k+3}(H_{k+3})$ відповідно, а у другому реченні S_2 знайдено послідовність слів $H_l \rightarrow H_{l+1} \rightarrow H_{l+2} \rightarrow H_{l+3} \rightarrow H_{l+4}$ характеристиками $Z_l(H_l)$, $Z_{l+1}(H_{l+1})$, $Z_{l+2}(H_{l+2})$, $Z_{l+3}(H_{l+3})$ та $Z_{l+4}(H_{l+4})$, таку що

$$Z_k(H_k) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\};$$

$$Z_{k+1}(H_{k+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\};$$

$$Z_{k+2}(H_{k+2}) = \{cm_{k+2}, g_{k+2} \neq 1, n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\};$$

$$Z_{k+3}(H_{k+3}) = \{cm_{k+3}, g_{k+3} \neq 1, n_{k+3}, k2_{k+3}, 0, 0, l_{k+3}, 3\};$$

$$Z_l(H_l) = \{cm_k, 1, n_k, k2_k, 0, 0, l_k, 1\};$$

$$Z_{l+1}(H_{l+1}) = \{5, 0, n_k, k2_k, t_{k+1}, h_{k+1}, l_k, 2\};$$

$$Z_{l+2}(H_{l+2}) = \{cm_{k+3}, (g_{l+2} \neq 1) \& (g_{l+2} \neq g_{k+3}), n_{k+3}, k2_{k+3}, 0, 0, l_{k+3}, 3\};$$

$$Z_{l+3}(H_{l+3}) = \{9, g_{l+4}, 0, 0, 0, 0, 0\};$$

$Z_{l+4}(H_{l+4}) = \{cm_{k+2}, (g_{l+4} \neq 1) \& (g_{l+4} \neq g_{k+2}), n_{k+2}, k2_{k+2}, 0, 0, l_{k+2}, 3\}$;
 $H_l \equiv H_{k+3}$, $\vec{H}_{l+1} \equiv \vec{H}_{k+1}$, $H_{l+3} \equiv H_{k+2}$, $H_l \equiv H_{k+3}$, а $\vec{H}_{l+1} \in \vec{K}_q$ та $\vec{H}_{k+1} \in \vec{K}_q$ – слова, що належать до одного і того ж ряду $\vec{K}_q \in K$ конверсивів із множини K , то для атомарних предикатів $L_p^S(h)$ та $L_p^{S_2}(h)$ можна здійснити таку заміну:

$$\begin{aligned}
 p_1 &\equiv p_2, \quad p_1 \in P^S, \quad p_2 \in P^S, \\
 x_1 &\equiv y_2, \quad x_1 \in X_p^S(h), \quad y_2 \in Y_p^S(x, g, h), \\
 y_1 &\equiv x_2, \quad y_1 \in Y_p^S(x, g, h), \quad x_2 \in X_p^S(h).
 \end{aligned}$$

Для двох речень природної мови «Сергій подарував батькові машину» та «Батько отримав від Сергія машину» логіко-лінгвістичні моделі будуть мати вигляд:

$$L^S_1 = p_1(x_1, 0, y_1, 0, z_1, 0, 0),$$

$$L^S_1 = \text{подарував (Сергій, 0, батькові, 0, машину, 0, 0)},$$

$$L^S_2 = p_2(y_1, 0, x_1, 0, z_1, 0, 0),$$

$$L^S_2 = \text{отримав (батько, 0, Сергія, 0, машину, 0, 0)}.$$

У логіко-лінгвістичній моделі другого речення суб'єктом відношення виступає об'єкт першого речення, а об'єктом – суб'єкт першого речення, при цьому предмети відношень однакові, а предикати являються конверсивами і відповідають схемі: $p_2 \equiv \vec{p}_1$, де \vec{p}_1 – конверсив до p_1 .

3.3 Алгоритм аналізу логіко-лінгвістичних моделей текстових документів

Для вилучення знань із текстової інформації необхідно застосувати до вхідного тексту метод, який дозволив би розглядати кожне введене речення як систему взаємозв'язаних елементів з метою дослідити зв'язки між ними. Таким методом є аналіз [85].

У логіці цей термін понятійний, тобто робить можливим утворення понять, розклад єдиного на множини, цілого – на частини, складного – на його компоненти, події – на її окремі ступені, змісту – на його елементи, поняття – на його ознаки [86].

Аналіз логіко-лінгвістичних моделей речень природної мови – це процес дослідження окремих компонент логіко-лінгвістичних моделей та їх граматичних характеристик з метою виявлення та встановлення логічних зв'язків між словами, які інтерпретуються цими компонентами.

Усі методи здійснення інформаційного пошуку поділяються на статистичні, методи пошуку за семантичними мережами та комбіновані методи [87].

Основною ідеєю статистичних методів є визначення ваги кожного слова у документі. Для них притаманна якісна математична модель, що дозволяє отримати правильні оцінки релевантності для документів. Недоліком статистичних методів є те, що вони не враховують змістовне навантаження текстів та тексту запиту. Статистичні методи лежать в основі роботи пошукових машин Google, Yandex, Yahoo та ін. [88].

Методи пошуку за семантичними мережами використовують дані, які представлені у вигляді онтологій, а пошук відбувається шляхом задання властивостей шуканого об'єкта. Такі методи враховують змістовне навантаження, проте застосовувати їх можна тільки для таких електронних документів, які містять семантичний опис контенту.

Комбіновані методи, окрім статистичних, використовують методи семантичного аналізу текстів. Саме до цієї групи методів пошуку інформації відносяться подальші дослідження.

Аналіз логіко-лінгвістичних моделей текстових документів представляє собою складний процес отримання інформації про структуру та зміст тексту, що розглядається, на основі виявлення закономірностей і тенденцій синтаксичної, семантичної та лексичної побудови тексту. Результатом проведення аналізу логіко-лінгвістичної моделі текстового документу є відновлений текст [89].

Нехай у відповідність деякому тексту t_1 поставлена його логіко-лінгвістична модель. Лінгвістична та семантико-синтаксична складова логіко-лінгвістичної моделі заданого тексту матимуть вигляд:

$$t_1 = \langle CQ, F_1, B_1, A_1 \rangle,$$

$$t'_1 = \begin{cases} L_{p^{(\lambda)}}^{S_1}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)}) \\ L_{p^{(\lambda)}}^{S_2}(x_2^{(\lambda)}, g_2^{(\lambda)}, y_2^{(\lambda)}, q_2^{(\lambda)}, z_2^{(\lambda)}, r_2^{(\lambda)}, h_2^{(\lambda)}) \\ \dots \\ L_{p^{(\lambda)}}^{S_{\delta}}(x_{\delta}^{(\lambda)}, g_{\delta}^{(\lambda)}, y_{\delta}^{(\lambda)}, q_{\delta}^{(\lambda)}, z_{\delta}^{(\lambda)}, r_{\delta}^{(\lambda)}, h_{\delta}^{(\lambda)}), \\ \dots \\ L_{p^{(\lambda)}}^{S_{N(t_1)}}(x_{N(t_1)}^{(\lambda)}, g_{N(t_1)}^{(\lambda)}, y_{N(t_1)}^{(\lambda)}, q_{N(t_1)}^{(\lambda)}, z_{N(t_1)}^{(\lambda)}, r_{N(t_1)}^{(\lambda)}, h_{N(t_1)}^{(\lambda)}). \end{cases}$$

Тут $\delta = \overline{1, N(t_1)}$ – номер речення у тексті, $N(t_1)$ – загальна кількість речень у тексті t_1 .

У процесі аналізу логіко-лінгвістичних моделей використовується база правил формування зв'язків між складними частинами тексту, а також між моделями $L^{S_{\delta_1}}$, $\delta_1 = \overline{1, N(t_1)}$ безпосередньо. Аналіз логіко-лінгвістичних моделей текстових документів потрібно здійснювати за чітко визначеним алгоритмом (рис. 3.1).

1. Аналізується перший параметр лінгвістичної складової моделі CQ_1 . Так як тип тексту визначає його структуру, а також стилістичні, семантичні та синтаксичні особливості, то в залежності від значення CQ_1 для тексту будуть характерні певні граматичні особливості, на які буде звертатися увага при подальшому аналізі. Тоді можна сказати, що існує такий оператор $Q_1(w_1)$, який ставить у відповідність конкретному значенню змінної CQ_1 із множини можливих значень $CQ \in C$ масив граматичних параметрів w_1 :

$$Q_1(w_1): C \rightarrow CQ_1.$$

2. Фіксується кількість складних частин електронного текстового документу $f_j \in F_1$, $j = \overline{1, m}$, m – кількість складних синтаксичних частин.

3. На відміну від двох попередніх параметрів, текстова база B_1 є однією із значущих змінних для формування змістовного портрету документу.

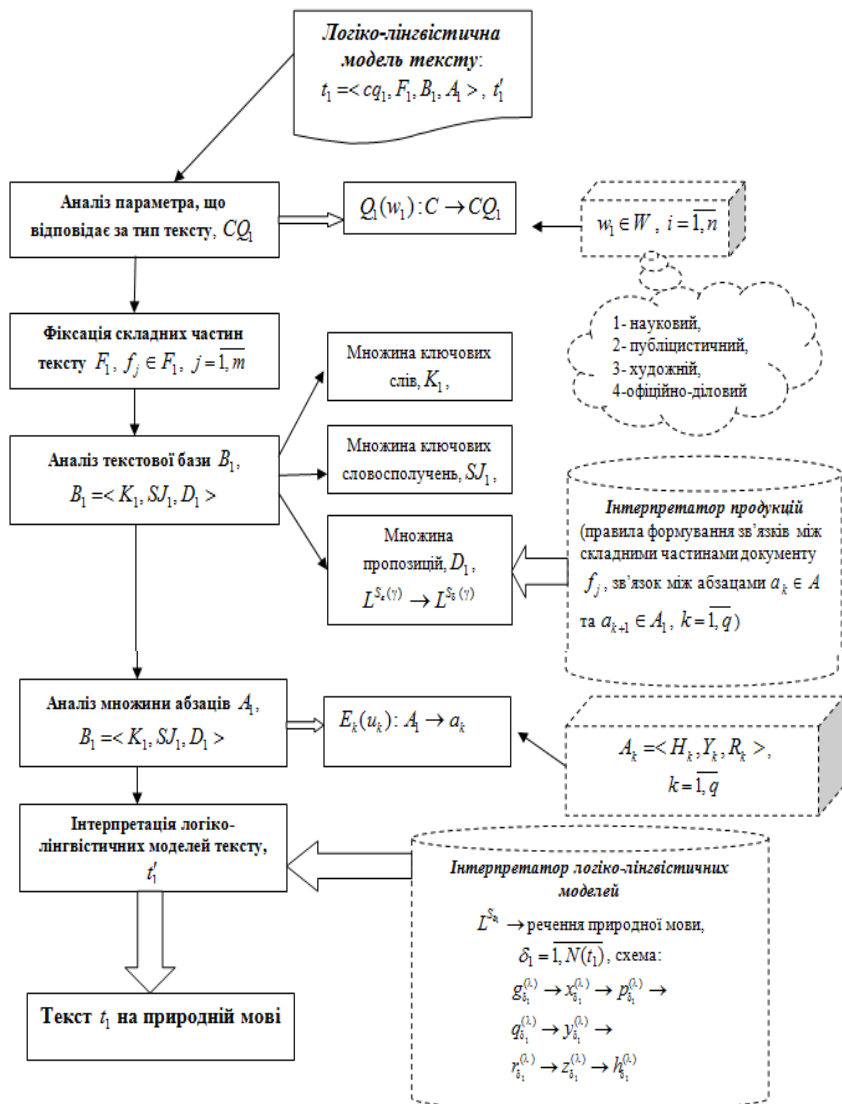


Рис. 3.1 – Алгоритм аналізу логіко-лінгвістичних моделей текстових документів

Множина ключових слів, множина ключових словосполучень, а також множина пропозицій формують основу для вилучення знань з електронного текстового документу.

Для аналізу множини пропозицій застосовується інтерпретатор продукцій, який працює циклічно. У кожному циклі він переглядає правила формування зв'язків між складними частинами документу f_j із множини заданих в моделі F_1 , щоб з'ясувати, які посилання із заданої множини D_1 збігаються з відомими на даний момент фактами з робочої пам'яті.

Після вибору правило спрацьовує, його висновок заноситься в робочу пам'ять, а цикл повторюється спочатку.

Тобто в результаті роботи інтерпретатора продукцій шукаються такі моделі $L^{S_e(\gamma)}$, що належать до абзацу $a_k \in A$, які за змістом передують або з яких випливають моделі $L^{S_\delta(\gamma)}$ з абзацу $a_{k+1} \in A$, $k = \overline{1, q}$: $L^{S_e(\gamma)} \rightarrow L^{S_\delta(\gamma)}$.

4. Аналізується множина абзаців A_1 , у кожному з них ($a_k \in A_1$) відмічається тип зв'язку між реченнями H_k , тип тематичної прогресії Y_k , тип рематичної домінанти R_k . Тобто, буде існувати такий оператор $E_k(u_k)$, який ставить у відповідність кожному абзацу a_k із множини A_1 масив параметрів u_k , що формують зміст відповідного абзацу:

$$E_k(u_k): A_1 \rightarrow a_k \text{ або } E_k(u_k): A_1 \rightarrow \langle H_k, Y_k, R_k \rangle.$$

5. Застосовується інтерпретатор логіко-лінгвістичних моделей речень природної мови.

В основу змістовного аналізу логіко-лінгвістичних моделей покладено функціональні залежності між головними та другорядними членами речень природної мови. При цьому використовуються такі основні принципи перетворення.

1) Визначити кількість атомарних предикатів v^S , що описують частини речення S , які відображають закінчений зміст.

2) Зафіксувати множини P^S , H_p^S , $X_p^S(h)$, $G_p^S(x, h)$, $Y_p^S(x, g, h)$, $Q_p^S(x, g, y, h)$, $Z_p^S(x, g, y, q, h)$, $R_p^S(x, g, y, q, z, h)$.

3) Зафіксувати значення елементів кортежу логічних операцій $O(S)$ речення S .

4) Визначити потужності множин P^S , H_p^S , $X_p^S(h)$, $G_p^S(x, h)$, $Y_p^S(x, g, h)$, $Q_p^S(x, g, y, h)$, $Z_p^S(x, g, y, q, h)$, $R_p^S(x, g, y, q, z, h)$.

5) Якщо хоча б одна із потужностей множин $|P^S|$ або $|X_p^S(h)|$ дорівнює одиниці, то речення природної мови просте.

6) Інакше, якщо хоча б одна із потужностей множин $|P^S| \geq 2$ або $|X_p^S(h)| \geq 2$, то речення природної мови складне.

7) Визначити тип окремої форми логіко-лінгвістичної моделі (1.1) – (1.4).

8) Для кожного атомарного предикату $L_p^S(x, g, y, q, z, r, h)$ сформувати словосполучення між такими елементами логіко-лінгвістичних моделей, як:

– відношенням $p \in P^S$ та характеристикою $h \in H_p^S$ цього відношення;

– відношенням $p \in P^S$ та об'єктом $y \in Y_p^S(x, g, h)$;

– об'єктом $y \in Y_p^S(x, g, h)$ та предметом відношення $z \in Z_p^S(x, g, y, q, h)$;

– суб'єктом $x \in X_p^S(h)$ та його характеристикою $g \in G_p^S(x, h)$;

– об'єктом $y \in Y_p^S(x, g, h)$ та його характеристикою $q \in Q_p^S(x, g, y, h)$;

– предметом відношення $z \in Z_p^S(x, g, y, q, h)$ та його характеристикою $r \in R_p^S(x, g, y, q, z, h)$.

9) Відновити прямий порядок слів у реченні природної мови шляхом об'єднання отриманих в п.8) словосполучень, а також враховуючи логічні операції між атомарними предикатами.

Отже, інтерпретатор перетворює модель $L^{S_{n_1}}, \delta_1 = \overline{1, N(t_1)}$ у речення природної мови шляхом синтезу простих предикатів (з урахуванням масиву граматичних параметрів w_1 , набору

пропозицій $L^{S_e(\gamma)} \rightarrow L^{S_\delta(\gamma)}$ та вектора параметрів u_k) за такою схемою:

$$g_{\delta_1}^{(\lambda)} \rightarrow x_{\delta_1}^{(\lambda)} \rightarrow p_{\delta_1}^{(\lambda)} \rightarrow q_{\delta_1}^{(\lambda)} \rightarrow y_{\delta_1}^{(\lambda)} \rightarrow r_{\delta_1}^{(\lambda)} \rightarrow z_{\delta_1}^{(\lambda)} \rightarrow h_{\delta_1}^{(\lambda)}.$$

Запропонований алгоритм дозволяє автоматично аналізувати логіко-лінгвістичні моделі текстів різних типів та довільної складності. Створена база правил формування зв'язків між складними частинами тексту, а також всередині абзаців є неодмінною складовою функціонування алгоритму, через те що саме вона дає можливість вилучати зміст з текстової інформації.

Логіко-лінгвістична модель (1.24) – (1.25) є засобом, що дозволяє формалізувати тексти природної мови за єдиним принципом. У свою чергу, аналіз таких моделей дає змогу зробити зворотну операцію – відновити текст. Формальний апарат трансформації тексту у логіко-лінгвістичну модель і навпаки виступає єдиним засобом автоматизації процесу обробки текстової інформації. Створення автоматичної системи лінгвістичного аналізу електронних документів спростовує такі гіпотези щодо складності текстів, як:

- чим більша кількість термінів у тексті, тим складніший він для перекладу;
- чим складніше дерево предикатної структури та синтаксис тексту, тим складніше парсинг тексту;
- складність тексту прямо пропорційна середній довжині слова та середній довжині речення.

Усі вище наведені гіпотези сформульовані, виходячи з статистичного аналізу природно мовних текстів. Змістовна ж компонента аналізу електронних документів містить у собі поєднання бази правил, що створена на основі досліджень лінгвістів, з методами обробки масивів текстової інформації.

Наприклад, нехай деякий текстовий фрагмент задано такими логіко-лінгвістичними моделями, з яких потрібно відновити текст:

$$\begin{aligned} \{L^{S^1} = & p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \& \\ & p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0), \\ L^{S^2} = & [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \& \\ & p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0)\}, \end{aligned}$$

$\{L^{S_1} = \text{заохочує (лекція, цікава, студентів, лінивих, навчання, подальшого, 0)} \&$

$\text{заохочує (лекція, цікава, практикантів, лінивих, навчання, подальшого, 0)} \&$

$\text{заохочує (лекція, цікава, студентів, працьовитих, навчання, подальшого, 0)} \&$

$\text{заохочує (лекція, цікава, практикантів, працьовитих, навчання, подальшого, 0),}$

$L^{S_2} = [\text{залежить (те, 0, змісту, лекції, 0,0,0)} \&$

$\text{залежить (те, 0, динаміки, 0, викладання,0,0)} \&$

$\text{залежить (те, 0, матеріалів, демонстраційних, 0,0,0)}] \rightarrow$

$\text{отримає (студент, 0, знання,0,0,0,0)}\}.$

1) Кількість атомарних предикатів, що описують частини речення S , які відображають закінчений зміст, дорівнює $\nu^{S_1} = 4$, $\nu^{S_2} = 4$.

2) Множини, що входять до логіко-лінгвістичної моделі містять такі елементи:

$P^{S_1} = \{\text{заохочує}\};$

$X_p^{S_1}(h) = \{\text{лекція}\};$

$G_p^{S_1}(x, h) = \{\text{цікава}\};$

$Y_p^{S_1}(x, g, h) = \{\text{студентів, практикантів}\};$

$Q_p^{S_1}(x, g, y, h) = \{\text{лінивих, працьовитих}\};$

$Z_p^{S_1}(x, g, y, q, h) = \{\text{навчання}\};$

$R_p^{S_1}(x, g, y, q, z, h) = \{\text{подальшого}\};$

$P^{S_2} = \{\text{залежить, отримає}\};$

$X_p^{S_2}(h) = \{\text{те, студент}\};$

$Y_p^{S_2}(x, g, h) = \{\text{змісту, динаміки, матеріалів, знання}\};$

$Q_p^{S_2}(x, g, y, h) = \{\text{лекцій, демонстраційних}\};$

$Z_p^{S_2}(x, g, y, q, h) = \{\text{викладання}\}.$

3) Кортеж логічних операцій речення S_1 містить такі елементи: $O(S_1) = [\&, \&, \&]$, для речення S_2 – $O(S_2) = [\&, \&, \rightarrow]$.

4) Потужності множин $|P^{S_1}|=1$; $|H_p^{S_1}|=\emptyset$; $|X_p^{S_1}(h)|=1$;
 $|G_p^{S_1}(x,h)|=1$; $|Y_p^{S_1}(x,g,h)|=2$; $|Q_p^{S_1}(x,g,y,h)|=2$;
 $|Z_p^{S_1}(x,g,y,q,h)|=1$; $|R_p^{S_1}(x,g,y,q,z,h)|=1$, $|P^{S_2}|=2$; $|H_p^{S_2}|=\emptyset$;
 $|X_p^{S_2}(h)|=2$; $|G_p^{S_2}(x,h)|=\emptyset$; $|Y_p^{S_2}(x,g,h)|=4$; $|Q_p^{S_2}(x,g,y,h)|=2$;
 $|Z_p^{S_2}(x,g,y,q,h)|=1$; $|R_p^{S_2}(x,g,y,q,z,h)|=\emptyset$,

5) Потужності множин відношень та суб'єктів першого речення $|P^{S_1}|=1$ і $|X_p^{S_1}(h)|=1$, тому речення природної мови просте з однорідними додатками та означеннями.

6) Потужності множин відношень та суб'єктів другого речення $|P^{S_2}|=2$ і $|X_p^{S_2}(h)|=2$, тому речення природної мови складне.

7) Так як елементи множини відношень та суб'єктів другого речення природної мови містять як різні елементи, так і такі, що дублюються, то це свідчить про наявність однорідних членів в простих реченнях, які входять до складного. Операція імплікації між простими предикатами у логіко-лінгвістичній моделі вказує на те, що речення складнопідрядне.

8) Логіко-лінгвістична модель першого речення природної мови S_1 складається з чотирьох атомарних предикатів:

$$\{L^{S_1} = p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0) \& p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0)\}.$$

Словосполучення, відновлені з атомарного предикату $p_1(x_1, g_1, y_{11}, q_{11}, z_1, r_1, 0)$: *«цікава лекція, заохочує студентів, лінивих студентів, студентів до навчання, подальшого навчання».*

Словосполучення, відновлені з атомарного предикату $p_1(x_1, g_1, y_{12}, q_{11}, z_1, r_1, 0)$: *«цікава лекція, заохочує практикантів, лінивих практикантів, практикантів до навчання, подальшого навчання».*

Словосполучення, відновлені з атомарного предикату $p_1(x_1, g_1, y_{11}, q_{12}, z_1, r_1, 0)$: *«цікава лекція, заохочує студентів, працюючих студентів, студентів до навчання, подальшого навчання».*

Словосполучення, відновлені з атомарного предикату $p_1(x_1, g_1, y_{12}, q_{12}, z_1, r_1, 0)$: *«цікава лекція, заохочує практикантів, працюючих практикантів, практикантів до навчання, подальшого навчання».*

Логіко-лінгвістична модель другого речення природної мови S_2 також складається з чотирьох атомарних предикатів:

$$L^{S_2} = [p_2(x_2, 0, y_{21}, x_1, 0, 0, 0) \& p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0) \& p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)] \rightarrow p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0)\}.$$

Словосполучення, відновлені з атомарного предикату $p_2(x_2, 0, y_{21}, x_1, 0, 0, 0)$: *«залежить від змісту, змісту лекції».*

Словосполучення, відновлені з атомарного предикату $p_2(x_2, 0, y_{22}, 0, z_{21}, 0, 0)$: *«залежить від динаміки, динаміки викладання».*

Словосполучення, відновлені з атомарного предикату $p_2(x_2, 0, y_{23}, q_{22}, 0, 0, 0)$: *«залежить від матеріалів, демонстраційних матеріалів».*

Словосполучення, відновлені з атомарного предикату $p'_2(y_{11}, 0, y'_2, 0, 0, 0, 0)$: *«отримає знання».*

9) Відновлений порядок слів у реченні природної мови S_1 за схемою $g_{\delta_1}^{(\lambda)} \rightarrow x_{\delta_1}^{(\lambda)} \rightarrow p_{\delta_1}^{(\lambda)} \rightarrow q_{\delta_1}^{(\lambda)} \rightarrow y_{\delta_1}^{(\lambda)} \rightarrow r_{\delta_1}^{(\lambda)} \rightarrow z_{\delta_1}^{(\lambda)} \rightarrow h_{\delta_1}^{(\lambda)}$: *«цікава лекція заохочує лінивих і працюючих студентів і практикантів до подальшого навчання».*

10) Відновлений порядок слів у реченні природної мови S_2 за схемою $g_{\delta_1}^{(\lambda)} \rightarrow x_{\delta_1}^{(\lambda)} \rightarrow p_{\delta_1}^{(\lambda)} \rightarrow q_{\delta_1}^{(\lambda)} \rightarrow y_{\delta_1}^{(\lambda)} \rightarrow r_{\delta_1}^{(\lambda)} \rightarrow z_{\delta_1}^{(\lambda)} \rightarrow h_{\delta_1}^{(\lambda)}$: *«те залежить від змісту лекції, від динаміки викладання, від демонстраційних матеріалів, внаслідок чого студент отримає знання».*

Оригінальний текстовий фрагмент, для якого були побудовані задані логіко-лінгвістичні моделі має вигляд: *"Цікава лекція заохочує лінивих і працюючих студентів та практикантів до подальшого навчання. Від її змісту, динаміки викладання та демонстраційних матеріалів залежить те, чи отримає знання студент".*

IV. ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕЛЕКТРОННИХ ТЕКСТОВИХ ДОКУМЕНТІВ НА ОСНОВІ ВИКОРИСТАННЯ ЛОГІКО-ЛІНГВІСТИЧНИХ МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ

"Використання порівнянь – один із кращих засобів обдурити себе та інших. Людина, що застосовує порівняння, не відчуває у собі здатності впевнити інших за допомогою фактів та логіки"
Альфред Адлер

4.1. Основні способи виявлення рерайту за допомогою логіко-лінгвістичного моделювання

Для порівняльного аналізу формальних моделей, що відображають зміст певних речень природної мови, необхідно зафіксувати можливі способи утворення тотожних формальних моделей (інваріантних форм логіко-лінгвістичних моделей).

Запропоновані способи виявлення тотожних за змістом речень природної мови [90] виведені на базі основних способів рерайтингу, які застосовуються сьогодні [91]:

- зміна конструкції речення природної мови;
- заміна прямої мови;
- використання синонімів, гіперонімів та конверсивів;
- зміна порядку речень та абзаців у тексті;
- видалення деяких слів та словосполучень з тексту [92].

Спосіб 1. Якщо одне речення природної мови написано у стверджувальній формі, а друге – у вигляді умови, то у першому випадку логіко-лінгвістична модель представляє собою атомарний предикат, а у другому – два атомарних предиката, з'єднаних логічною операцією імплікації, і при цьому зберігається еквівалентність наступних компонент логіко-лінгвістичних моделей: $y_1 \equiv y_2$, $z_1 \equiv z_2$, $h_1 \equiv \tilde{h}_2$ (антоніми), $x_1 \equiv x'_2$, $g_1 \equiv g'_2$,

$p_1 \equiv \bar{p}_2$ (конверсиви), $y_1 \equiv \bar{p}'_2$ (спільнокореневі), $x_2 \equiv 0$, то речення однакові за змістом.

Наприклад, є два речення: "Флективні мови погано піддаються опису формалізмами" та "Якщо формалізмами не вдається правильно зробити опис, то вважають, що описуються флективні мови".

Перевіримо за першою умовою, чи являється друге і третє речення рерайтом першого.

Для цього необхідно побудувати логіко-лінгвістичні моделі цих двох речень:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \rightarrow p'_2(x'_2, g'_2, y'_2, q'_2, z'_2, r'_2, h'_2) =$$

$$-p_{21} \& p_{22}(0, 0, y_2, 0, z_2, 0, h_2) \rightarrow p'_2(x'_2, g'_2, 0, 0, 0, 0, h'_2),$$

L^{S_2} = -вдається&зробити(0, 0, опис, 0, формалізмами, 0, правильно) \rightarrow описуються(мови, флективні, 0, 0, 0, 0, 0).

Спосіб 2. Якщо для перефразування речення вжитий спосіб перенесення дії у пасивний стан, а в логіко-лінгвістичних моделях двох речень, що порівнюються, простежується еквівалентність таких компонент $x_2 \equiv 0$, $x_1 \equiv y_2$, $g_1 \equiv q_2$, $z_1 \equiv z_2$, $h_1 \equiv \hat{h}_2$ (синоніми), то одне з речень природної мови являється рерайтом другого.

Нехай є два речення: "Флективні мови погано піддаються опису формалізмами" та "Для флективних мов дуже важко підібрати формалізми".

Логіко-лінгвістичні моделі речень будуть мати вигляд, відповідно:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) = p_2(0, 0, y_2, q_2, z_2, 0, h_{21} \& h_{22}),$$

L^{S_2} = підібрати(0, 0, мов, флективних, формалізми, 0, дуже&важко).

Спосіб 3. Якщо у двох реченнях природної мови було здійснено пряму заміну слів синонімами, то логіко-лінгвістичні моделі таких речень будуть еквівалентними, тобто $p_1 \equiv \widehat{p}_2$ (спільнокореневі), $x_1 \equiv \widehat{x}_2$, $z_1 \equiv \widehat{z}_2$, $h_1 \equiv \widehat{h}_2$ (синоніми), при цьому характеристики об'єктів, суб'єктів та предметів відношень можуть бути упущені або додані, тому їх еквівалентність не являється обов'язковою, то речення однакові за змістом, і одне з них являється рерайтом іншого.

Наприклад, нехай є два речення: "Флективні мови погано піддаються опису формалізмами" та "Природні мови описуються математичними формулами не дуже добре".

Логіко-лінгвістичні моделі таких речень будуть мати вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) = p_2(x_2, g_2, y_2, q_2, 0, 0, -h_{21} \& h_{22}),$$

L^{S_2} = описуються(мови, природні, формулами, математичними, 0, 0, \neg дуже&добре).

Спосіб 4. Якщо дія, про яку йдеться у першому реченні, замінена на дієприслівниковий зворот у другому, а логіко-лінгвістичні моделі таких речень матимуть однакову структуру атомарних предикатів, проте у логіко-лінгвістичній моделі другого речення буде наявний ще один предикат, при цьому $x_1 \equiv y'_2$, $g_1 \equiv q'_2$, $z_1 \equiv z'_2$, $x'_2 \equiv 0$, $y_1 \equiv \widehat{p}'_2$ (спільнокореневі), то одне з речень природної мови являється рерайтом другого.

Наприклад, два речення "Флективні мови погано піддаються опису формалізмами" та "Погано піддаючись опису, флективні мови все-таки намагаються описати формалізмами".

Логіко-лінгвістичні моделі речень мають вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \& p'_2(x'_2, g'_2, y'_2, q'_2, z'_2, r'_2, h'_2) =$$

$$p_2(x_2, g_2, y_2, 0, 0, 0, h_2) \& p'_{21} \& p'_{22}(0, 0, y'_2, q'_2, z'_2, 0, h'_2),$$

L^{S_2} = піддаються (мови, флективні, опису, 0, 0, 0, погано)&

намагаються&описати(0,0,мови,флексивні,формалізмами,0,все-таки).

Спосіб 5. Якщо у першому реченні, що порівнюється, присутня пряма мова, а у другому відсутня, і при цьому логіко-лінгвістичні моделі речень мають такий вигляд

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \rightarrow p'_1(x'_1, g'_1, y'_1, q'_1, z'_1, r'_1, h'_1),$$

$$L^{S_2} = p_2(\hat{x}'_1, \hat{g}_2, \hat{p}'_1, q_2, z_2, r_2, h_2) \rightarrow p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1),$$

то зміст речень однаковий, а одне з них являється рерайтом другого.

Наприклад, нехай є два речення "*Флексивні мови погано піддаються опису формалізмами*", – *констатують вчені*" та "*Вченими була дана констатація факту про те, що флексивні мови погано піддаються опису формалізмами*".

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) \rightarrow p'_1(x'_1, g'_1, y'_1, q'_1, z'_1, r'_1, h'_1) =$$

$$p_1(x_1, g_1, y_1, 0, z_1, 0, h_1) \rightarrow p'_1(x'_1, 0, 0, 0, 0, 0, 0),$$

$$L^{S_1} = \text{піддаються (мови, флексивні, опису, 0, формалізмами, 0, погано)} \rightarrow \text{констатують(вчені, 0, 0, 0, 0, 0, 0)}.$$

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \rightarrow p'_2(x'_2, g'_2, y'_2, q'_2, z'_2, r'_2, h'_2) =$$

$$p_{21} \& p_{22}(x_2, 0, y_2, 0, z_2, 0, 0) \& p'_2(x'_2, g'_2, y'_2, 0, z'_2, 0, h'_2),$$

$$L^{S_2} = \text{була\&дана(вченими, 0, констатація, 0, факту, 0, 0)} \rightarrow$$

$$\text{піддаються(мови, флексивні, опису, 0, формалізмами, 0, погано)}.$$

Спосіб 6. Якщо дія в першому реченні заміняється однаковим за змістом іменником, при цьому логіко-лінгвістична модель другого речення представляє собою атомарні предикати, з'єднані логічною операцією імплікації з однаковими відношеннями, а $x_1 \equiv y_{21}$, $g_1 \equiv q_{21}$, $y_1 \equiv z_{21}$, $z_1 \equiv r_{21}$, $h_1 \equiv g_{22}$, то одне з речень природної мови являється рерайтом другого.

Нехай є два речення: "*Флексивні мови погано піддаються опису формалізмами*" та "*Піддавання флексивних мов опису формалізмами – дуже складна справа*".

Логіко-лінгвістичні моделі речень будуть мати вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_{21}, g_{21}, y_{21}, q_{21}, z_{21}, r_{21}, h_{21}) \& p_2(x_{22}, g_{22}, y_{22}, q_{22}, z_{22}, r_{22}, h_{22}) = p_2(0, 0, y_{21}, q_{21}, z_{21}, r_{21}, 0) \& p_2(x_{22}, g_{22}, 0, 0, 0, 0, h_{22}),$$

L^{S_2} = піддавання (0, 0, мов, флективних, опису, формалізмами, 0) → піддавання (справа, складна, 0, 0, 0, 0, дуже).

Спосіб 7. Якщо в реченні була здійснена зміна структури за рахунок об'єднання або роз'єднання змістовних складових, то в логіко-лінгвістичних моделях це відобразиться на кількості атомарних предикатів з однаковими компонентами, що фактично, дублюють одна одну, при цьому $x_1 \equiv y_2$, $g_1 \equiv q_2$, $z_1 \equiv z'_2$, $x_2 \equiv 0$, $g_2 \equiv 0$, $y_1 \equiv \hat{p}_2$ (спільнокореневі), $h_1 \equiv \hat{h}_2 \equiv \hat{h}'_2$ (синоніми), то над реченнями, що порівнюються, здійснено рерайтинг.

Наприклад, нехай є речення "Флективні мови погано піддаються опису формалізмами" і "Описувати флективні мови дуже складно. Ще складніше описувати флективні мови формалізмами".

Логіко-лінгвістичні моделі речень будуть мати вигляд:

$$L^{S_1} = p_1(x_1, g_1, y_1, q_1, z_1, r_1, h_1) = p_1(x_1, g_1, y_1, 0, z_1, 0, h_1),$$

L^{S_1} = піддаються (мови, флективні, опису, 0, формалізмами, 0, погано).

$$L^{S_2} = p_2(x_2, g_2, y_2, q_2, z_2, r_2, h_2) \& p'_2(x'_2, g'_2, y'_2, q'_2, z'_2, r'_2, h'_2) = p_2(0, 0, y_2, q_2, 0, 0, h_2) \& p_2(0, 0, y_2, q_2, z'_2, 0, h'_2),$$

L^{S_2} = описувати (0, 0, мови, флективні, 0, 0, дуже & складно) & описувати (0, 0, мови, флективні, формалізмами, 0, складніше).

Для успішного проведення пошуку рерайту необхідно здійснювати перевірку всіх запропонованих способів паралельно, що дасть можливість виявляти прийоми комбінованого рерайтингу. Дослідження різних методів рерайтингу дало можливість виявити закономірності та можливості заміни елементів логіко-лінгвістичних моделей для формального встановлення їх тотожності.

4.2. Метод порівняльного аналізу електронних текстових документів за змістом

Метод порівняльного аналізу текстових електронних документів базується на формуванні змістовних моделей текстових документів на основі використання математичного апарату логіки предикатів і складається з таких етапів.

1. Побудова логіко-лінгвістичних моделей. Для довільного електронного документу можна побудувати його логіко-лінгвістичну модель (1.24) – (1.25). Для побудови таких моделей використовується метод автоматичного формування логіко-лінгвістичних моделей речень природної мови [93], а також алгоритм побудови логіко-лінгвістичної моделі тексту [94]. В результаті виконання цього етапу для двох електронних текстових документів t_1 і t_2 формуються такі логіко-лінгвістичні моделі:

$$t_1 \ll CQ_1, F_1, B_1, A_1 \gg, \quad (4.1)$$

$$t'_1 = \bigwedge_{\delta_1=1}^{N(t_1)} L^{S_{\delta_1}}, \quad (4.2)$$

$$t_2 \ll CQ_2, F_2, B_2, A_2 \gg, \quad (4.3)$$

$$t'_2 = \bigwedge_{\delta_2=1}^{N(t_2)} L^{S_{\delta_2}}. \quad (4.4)$$

2. Порівняння семантико-синтаксичних складових. На цьому етапі до лінгвістичних складових (4.2) та (4.4) двох електронних документів застосовується метод пошуку текстових збігів у реченнях природної мови довільної складності. За методом почергово порівнюються логіко-лінгвістичні моделі $L^{S_{\delta_1}}$ та $L^{S_{\delta_2}}$, де $\delta_1 = \overline{1, N(t_1)}$, $\delta_2 = \overline{1, N(t_2)}$. Це дає можливість визначити ступінь схожості речень, що входять до текстів. Основою порівняльного аналізу є формування інваріантних логіко-лінгвістичних моделей $Q^{S_{\delta_1}}$ та $Q^{S_{\delta_2}}$, які базуються на основних способах виявлення рерайту за допомогою логіко-лінгвістичного моделювання. Інваріантні форми логіко-лінгвістичних моделей безпосередньо впливають на розрахунок відсотку збігів у реченнях текстів. Після виконання даного етапу буде знайдено відсоток збігів $\eta(S)$ для електронних текстових документів t_1 і t_2 .

3. Порівняння лінгвістичних складових. Аналізуються лінгвістичні складові логіко-лінгвістичних моделей (4.1) та (4.3).

3.1. Порівнюються перші параметри лінгвістичних складових CQ_1 та CQ_2 . На основі порівняння визначається відсоток збігу $\eta(CQ)$:

$$\eta(CQ) = \frac{CQ_1}{CQ_2} \cdot 100\%.$$

3.2. Порівнюються другі параметри лінгвістичних складових F_1 та F_2 . На основі порівняння визначається відсоток збігу $\eta(F)$:

$$\eta(F) = \frac{f_1 + \dots + f_{j_1} + \dots + f_{m_1}}{f_1 + \dots + f_{j_2} + \dots + f_{m_2}} \cdot 100\%,$$

де $j_1 = \overline{1, m_1}$, m_1 - це загальна кількість складних синтаксичних частин першого електронного текстового документу;

$j_2 = \overline{1, m_2}$, m_2 - це загальна кількість складних синтаксичних частин другого електронного текстового документу.

3.3. При порівнянні третього параметру лінгвістичної складової оцінюється схожість ключових слів та словосполучень, а також множина продукцій, отримана на основі синтезу логіко-лінгвістичних моделей. В результаті порівняння продукцій розраховується відсоток збігу $\eta(B)$:

$$\eta(B) = \frac{\{L^{S_{q_1}(y)} \rightarrow L^{S_{q_1}(y)}\}}{\{L^{S_{q_2}(y)} \rightarrow L^{S_{q_2}(y)}\}} \cdot 100\%.$$

3.4. Порівнюються четверті параметри лінгвістичних складових, на основі чого визначається відсоток збігу $\eta(A)$:

$$\eta(A) = \frac{\max(a_{k_1})}{\max(a_{k_2})} \cdot 100\%,$$

де $k_1 = \overline{1, q_1}$, q_1 - кількість абзаців у тексті t_1 ,

$k_2 = \overline{1, q_2}$, де q_2 - кількість абзаців у тексті t_2 .

Визначення відсотку збігів за різними компонентами логіко-лінгвістичної моделі електронного текстового документу дає можливість проаналізувати характер збігів, а також визначити засоби утворення текстових дублікатів.

У свою чергу порівняння семантико-синтаксичних складових логіко-лінгвістичних моделей електронних текстових документів (тобто другий етап) здійснюється за методом порівняння логіко-лінгвістичних моделей речень природної мови довільної складності, що передбачає виконання таких етапів.

Нехай є два речення природної мови довільної складності S_1 та S_2 , кожному з яких ставиться у відповідність логіко-лінгвістична модель, побудована згідно з формулами (1.1) – (1.4):

$$L^{S_1} = \bigwedge_{\lambda=1}^{v^{S_1}} L_{p^{(\lambda)}}^{S_1}(h), \quad (4.5)$$

$$L^{S_2} = \bigwedge_{\mu=1}^{v^{S_2}} L_{p^{(\mu)}}^{S_1}(h), \quad (4.6)$$

де $L_{p^{(\lambda)}}^{S_1}(h) = p_1^{(\lambda)}(x_1^{(\lambda)}, g_1^{(\lambda)}, y_1^{(\lambda)}, q_1^{(\lambda)}, z_1^{(\lambda)}, r_1^{(\lambda)}, h_1^{(\lambda)})$ – простий предикат, що описує частину речення S_1 , яка відображає закінчений зміст;

$L_{p^{(\mu)}}^{S_2}(h) = p_2^{(\mu)}(x_2^{(\mu)}, g_2^{(\mu)}, y_2^{(\mu)}, q_2^{(\mu)}, z_2^{(\mu)}, r_2^{(\mu)}, h_2^{(\mu)})$ – простий предикат, що описує частину речення S_2 , яка відображає закінчений зміст;

$\lambda = 1, \overline{v^{S_1}}, v^{S_1}$ – кількість частин речення S_1 , які мають закінчений зміст, і обчислюється за формулами (1.5) – (1.8);

$\mu = 1, \overline{v^{S_2}}, v^{S_2}$ – кількість частин речення S_2 , які мають закінчений зміст, і обчислюється за формулами (1.5) – (1.8).

1. Побудова вектора характеристик речення. На цьому етапі для кожного складного речення природної мови S_1 та S_2 ідентифікуються параметри заданих моделей (4.5) та (4.6). Для кожної логіко-лінгвістичної моделі встановлюються значення кількості атомарних предикатів v^{S_1} та v^{S_2} , фіксуються логічні операції, використані у моделях (4.5) та (4.6). Внаслідок цього формується вектор характеристик $U(S)$ складного речення:

$$U(S) = \{v^S, O(S), W(S)\},$$

де v^S – кількість атомарних предикатів, що описують частини речення S , які відображає закінчений зміст;

$O(S)$ – кортеж логічних операцій речення S ;

$$O(S) = [o_e(S) \mid e = \overline{1, \bar{h}}],$$

де \bar{h} – загальна кількість логічних операцій, наявних у реченні S ;

$W(S)$ – вид окремої форми уніфікованої логіко-лінгвістичної моделі, $W(S) = \{1, \dots, w_1, \dots, 15\}$ – множина типів окремих форм, елементи якої набувають значення $w_i = \overline{1, 15}$.

Після виконання першого кроку алгоритму порівняльного аналізу речень природної мови довільної складності логіко-лінгвістичні моделі (4.5) та (4.6) будуть мати такі вектори характеристик відповідно:

$$U(S_1) = \{v^{S_1}, O(S_1), W(S_1)\} \text{ та}$$

$$U(S_2) = \{v^{S_2}, O(S_2), W(S_2)\}.$$

2. Ідентифікація елементів логіко-лінгвістичних моделей.

Для кожного атомарного предиката $L_{p^{(\lambda)}}^{S_1}(h)$ та $L_{p^{(\mu)}}^{S_2}(h)$ фіксуються предикатні змінні та константи, внаслідок чого кожній компоненті логіко-лінгвістичної моделі ставиться у відповідність конкретне значення концепту з моделі (4.5) та (4.6) відповідно. Результатом виконання даного етапу є дві таблиці відношень (табл. 4.1, табл. 4.2).

Таблиця 1 – Відношення складного речення S_1

№	p	x	g	y	q	z	r	h
1	$p_1^{(1)}$	$x_1^{(1)}$	$g_1^{(1)}$	$y_1^{(1)}$	$q_1^{(1)}$	$z_1^{(1)}$	$r_1^{(1)}$	$h_1^{(1)}$
...
λ	$p_1^{(\lambda)}$	$x_1^{(\lambda)}$	$g_1^{(\lambda)}$	$y_1^{(\lambda)}$	$q_1^{(\lambda)}$	$z_1^{(\lambda)}$	$r_1^{(\lambda)}$	$h_1^{(\lambda)}$
...
v^{S_1}	$p_1^{(v^{S_1})}$	$x_1^{(v^{S_1})}$	$g_1^{(v^{S_1})}$	$y_1^{(v^{S_1})}$	$q_1^{(v^{S_1})}$	$z_1^{(v^{S_1})}$	$r_1^{(v^{S_1})}$	$h_1^{(v^{S_1})}$

Таблиця 2 – Відношення складного речення S_2

№	p	x	g	y	q	z	r	h
1	$p_2^{(1)}$	$x_2^{(1)}$	$g_2^{(1)}$	$y_2^{(1)}$	$q_2^{(1)}$	$z_2^{(1)}$	$r_2^{(1)}$	$h_2^{(1)}$
...
μ	$p_2^{(\mu)}$	$x_2^{(\mu)}$	$g_2^{(\mu)}$	$y_2^{(\mu)}$	$q_2^{(\mu)}$	$z_2^{(\mu)}$	$r_2^{(\mu)}$	$h_2^{(\mu)}$
...

№	p	x	g	y	q	z	r	h
v^{S_2}	$p_2^{(v^{S_2})}$	$x_2^{(v^{S_2})}$	$g_2^{(v^{S_2})}$	$y_2^{(v^{S_2})}$	$q_2^{(v^{S_2})}$	$z_2^{(v^{S_2})}$	$r_2^{(v^{S_2})}$	$h_2^{(v^{S_2})}$

3. Формування інваріантних логіко-лінгвістичних моделей.

Кожному реченню притаманний предикат (присудок), його визначення – логічний суб'єкт, що задає граматико-змістовну перспективу речення. Інваріантна форма логіко-лінгвістичної моделі формується за шаблоном (1.1) – (1.4) для різних типів складнопідрядних речень природної мови.

Залежно від типу логіко-лінгвістичної моделі, до якого належать задані формули (4.5) та (4.6), формуються інваріантні форми, що відповідають тотожним за змістом реченням природної мови:

$$(L^{S_1} \rightarrow Q^{S_1}) \& (Q^{S_1} \rightarrow L^{S_1}) \text{ та} \\ (L^{S_2} \rightarrow Q^{S_2}) \& (Q^{S_2} \rightarrow L^{S_2}),$$

де $Q^{S_1} \subseteq Q1$, де $Q1$ – множина можливих інваріантних форм логіко-лінгвістичної моделі L^{S_1} ;

$Q^{S_2} \subseteq Q2$, де $Q2$ – множина можливих інваріантних форм логіко-лінгвістичної моделі L^{S_2} .

Після виконання етапу формування інваріантних логіко-лінгвістичних моделей для кожного речення природної мови табл. 4.1 і табл. 4.2 доповнюються та будуть містити всі можливі варіанти відношень, що відповідають синонімічно тотожним конструкціям (табл. 4.3, табл. 4.4).

Таблиця 3 – Відношення складного речення S_1 з урахуванням інваріантних логіко-лінгвістичних моделей

№	p	x	g	y	q	z	r	h
1	$p_1^{(1)}$	$x_1^{(1)}$	$g_1^{(1)}$	$y_1^{(1)}$	$q_1^{(1)}$	$z_1^{(1)}$	$r_1^{(1)}$	$h_1^{(1)}$
...
λ	$p_1^{(\lambda)}$	$x_1^{(\lambda)}$	$g_1^{(\lambda)}$	$y_1^{(\lambda)}$	$q_1^{(\lambda)}$	$z_1^{(\lambda)}$	$r_1^{(\lambda)}$	$h_1^{(\lambda)}$
...
v^{S_1}	$p_1^{(v^{S_1})}$	$x_1^{(v^{S_1})}$	$g_1^{(v^{S_1})}$	$y_1^{(v^{S_1})}$	$q_1^{(v^{S_1})}$	$z_1^{(v^{S_1})}$	$r_1^{(v^{S_1})}$	$h_1^{(v^{S_1})}$

Продовження таблиці 3

№	p	x	g	y	q	z	r	h
Φ_1	$p_1^{(\Phi_1)}$	$x_1^{(\Phi_1)}$	$g_1^{(\Phi_1)}$	$y_1^{(\Phi_1)}$	$q_1^{(\Phi_1)}$	$z_1^{(\Phi_1)}$	$r_1^{(\Phi_1)}$	$h_1^{(\Phi_1)}$
...
\mathfrak{S}^{S_1}	$p_1^{(S_1)}$	$x_1^{(S_1)}$	$g_1^{(S_1)}$	$y_1^{(S_1)}$	$q_1^{(S_1)}$	$z_1^{(S_1)}$	$r_1^{(S_1)}$	$h_1^{(S_1)}$

Таблиця 4 – Відношення складного речення S_2 з урахуванням інваріантних логіко-лінгвістичних моделей

№	p	x	g	y	q	z	r	h
1	$p_2^{(1)}$	$x_2^{(1)}$	$g_2^{(1)}$	$y_2^{(1)}$	$q_2^{(1)}$	$z_2^{(1)}$	$r_2^{(1)}$	$h_2^{(1)}$
...
μ	$p_2^{(\mu)}$	$x_2^{(\mu)}$	$g_2^{(\mu)}$	$y_2^{(\mu)}$	$q_2^{(\mu)}$	$z_2^{(\mu)}$	$r_2^{(\mu)}$	$h_2^{(\mu)}$
...
ν^{S_2}	$p_2^{(\nu^{S_2})}$	$x_2^{(\nu^{S_2})}$	$g_2^{(\nu^{S_2})}$	$y_2^{(\nu^{S_2})}$	$q_2^{(\nu^{S_2})}$	$z_2^{(\nu^{S_2})}$	$r_2^{(\nu^{S_2})}$	$h_2^{(\nu^{S_2})}$
Φ_2	$p_2^{(\Phi_2)}$	$x_2^{(\Phi_2)}$	$g_2^{(\Phi_2)}$	$y_2^{(\Phi_2)}$	$q_2^{(\Phi_2)}$	$z_2^{(\Phi_2)}$	$r_2^{(\Phi_2)}$	$h_2^{(\Phi_2)}$
...
\mathfrak{S}^{S_2}	$p_2^{(S_2)}$	$x_2^{(S_2)}$	$g_2^{(S_2)}$	$y_2^{(S_2)}$	$q_2^{(S_2)}$	$z_2^{(S_2)}$	$r_2^{(S_2)}$	$h_2^{(S_2)}$

4. Порівняння векторів характеристик. На даному етапі відбувається порівняння векторів характеристик двох заданих речень S_1 та S_2 .

1) Фіксація значення $W(S_1)$ та $W(S_2)$ видів окремої форми уніфікованої логіко-лінгвістичної моделі (1.1) – (1.4).

2) Якщо кількість атомарних предикатів, що описують частини речення, які відображають закінчений зміст, для обох речень однакова $\nu^{S_1} = \nu^{S_2}$, то перейти до п. 3), інакше перейти до п. 5.

3) Порівняти кортежі логічних операцій, що використані у реченнях S_1 та S_2 :

$$O(S_1) = [o_{e_1}(S_1) \mid e_1 = \overline{1, \bar{h}_1}],$$

$$O(S_2) = [o_{e_2}(S_2) \mid e_2 = \overline{1, \bar{h}_2}].$$

5. Порівняння логіко-лінгвістичних моделей. Цей етап передбачає перевірку ряду умов, зокрема, умов протиріччя та

синонімії двох речень природної мови, можливість побудови імплікативних та пресуппозиційних зв'язків між атомарними предикатами, а також умови організації граматичного зв'язку між предикативними частинами речень. У результаті перевірки встановлюється істинність або хибність факту виявлення тотожних за змістом речень природної мови довільної складності.

1) Здійснити порівняння простого предиката $L_{p^{(\lambda)}}^{S_1}(h)$ логіко-лінгвістичної моделі (4.5) та його інваріантної форми $Q_{p^{(\varphi_1)}}^{S_1}(h)$ з простим предикатом $L_{p^{(\mu)}}^{S_2}(h)$ логіко-лінгвістичної моделі (4.6) та його інваріантною формою $Q_{p^{(\varphi_2)}}^{S_2}(h)$ методом порівняння простих речень за змістом.

2) Нехай змінна $\theta(S)$ – загальна кількість збігів простих предикатів речення природної мови S_1 з простими предикатами речення природної мови S_2 .

3) Нехай змінна $\tau(S)$ – загальна кількість збігів простих інваріантних предикатів речень природної мови S_1 та S_2 .

4) Перевірити істинність умов та зафіксувати загальну кількість збігів:

– якщо істинною є умова

$$(L_{p^{(\lambda)}}^{S_1}(h) \equiv L_{p^{(\mu)}}^{S_2}(h)),$$

то $\theta(S)$ збільшується на одиницю, тобто $\theta(S) = \theta(S) + 1$;

– якщо істинною є умова

$$(Q_{p^{(\varphi_1)}}^{S_1}(h) \equiv L_{p^{(\mu)}}^{S_2}(h)) \vee (L_{p^{(\lambda)}}^{S_1}(h) \equiv Q_{p^{(\varphi_2)}}^{S_2}(h)) \vee (Q_{p^{(\varphi_1)}}^{S_1}(h) \equiv Q_{p^{(\varphi_2)}}^{S_2}(h)),$$

то $\tau(S)$ збільшується на одиницю, тобто $\tau(S) = \tau(S) + 1$.

5) Загальна кількість порівнянь $\omega(S)$ обчислюється як $\omega(S) = (v^{S_1} + \vartheta^{S_1})(v^{S_2} + \vartheta^{S_2})$, тобто дорівнює добутку суми кількості простих та простих інваріантних предикатів речення S_1 і суми кількості простих та простих інваріантних предикатів речення S_2 .

6) Обчислити відсоток збігів $\eta(S)$ у реченнях природної мови S_1 та S_2 :

– якщо $W(S_1) = W(S_2)$ і $\Delta h = 0$, то

$$\eta(S) = \frac{\theta(S) + \tau(S)}{v^{S_1}} \cdot 100\% = \frac{\theta(S) + \tau(S)}{v^{S_2}} \cdot 100\% ;$$

- якщо $W(S_1) = W(S_2)$ і $\Delta h = h_1 - h_2 > 0$, то

$$\eta(S) = \frac{\theta(S) + \tau(S)}{v^{S_2}} \cdot 100\% ;$$

- якщо $W(S_1) = W(S_2)$ і $\Delta h = h_1 - h_2 < 0$, то

$$\eta(S) = \frac{\theta(S) + \tau(S)}{v^{S_1}} \cdot 100\% ;$$

- якщо $W(S_1) \neq W(S_2)$, то

$$\eta(S) = \frac{\tau(S)}{v^{S_1}} \cdot 100\% , \text{ при } h_1 < h_2 ;$$

$$\eta(S) = \frac{\tau(S)}{v^{S_2}} \cdot 100\% , \text{ при } h_2 < h_1 .$$

Таким чином, метод порівняльного аналізу речень природної мови довільної складності дає можливість визначити ступінь схожості речень за змістом. Це стає можливим завдяки формуванню інваріантних логіко-лінгвістичних моделей, що будуються на основі перевірки способів виявлення рерайту за допомогою логіко-лінгвістичного моделювання.

4.3. Інформаційна технологія порівняльного аналізу електронних текстових документів за змістом

Інформаційна технологія порівняльного аналізу електронних текстових документів за змістом представляє собою сукупність методів, процесів та програмно-технічних засобів, об'єднаних у технологічний ланцюжок, який забезпечує введення, обробку, зберігання, розподілення та відображення [95] текстової інформації, що представлена у вигляді природно мовних електронних документів, для отримання нового показника якості у вигляді відсотку збігів. Технічними засобами для створення інформаційної технології є апаратне, програмне та математичне забезпечення, що дає можливість здійснювати переробку первинної інформації, представленої у вигляді природно мовних електронних документів, на інформацію нової якості.

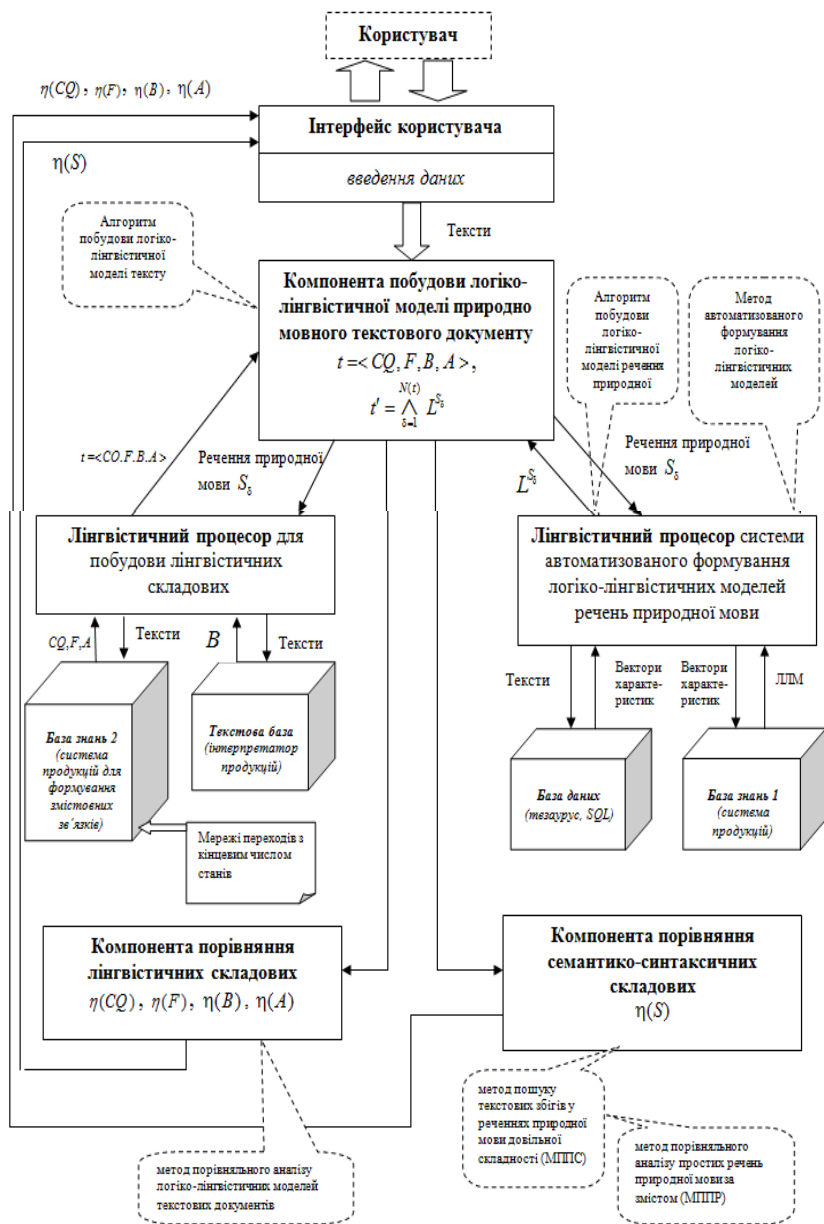


Рис. 4.1 – Функціональна схема інформаційної технології порівняльного аналізу текстових документів

Функціональна структура інформаційної технології порівняльного аналізу двох текстових документів представлена на рис. 4.1, де відображено її основні структурні одиниці [96, 97].

Інтерфейс користувача представляє собою програмний модуль, який відповідає етапу введення текстової інформації.

У полях вводу обираються два тексти, порівняльний аналіз яких необхідно здійснити. Для текстів, що порівнюються, існує ряд обмежень:

- 1) система обробляє речення української мови;
- 2) речення повинно бути грамотно написано;
- 3) слова у реченні розділені пробілами;
- 4) кінцем речення вважаються такі знаки пунктуації, як «.», «!», «?»;
- 5) знаки пунктуації необхідно вказувати безпосередньо після слова, далі пробіл;
- 6) тире і дефіс позначаються у реченні одним і тим самим символом «-», після тире ставиться пробіл, після дефіса – одразу записується слово.

Інтерфейс користувача інформаційної технології порівняльного аналізу двох електронних текстових документів представляє собою два поля для введення заданих текстів та поля для отримання інформації щодо характеристик слів, що входять до складу введених текстів, та логічних зв'язків між ними (рис. 4.2).

Компонента побудови логіко-лінгвістичної моделі природно мовного текстового документу відповідає першому етапу методу порівняльного аналізу логіко-лінгвістичних моделей текстових документів – побудові таких моделей. Ця компонента входить до блоку обробки інформації і включає в себе роботу двох лінгвістичних процесорів: один з яких працює на основі методу автоматизованого формування логіко-лінгвістичних моделей текстової інформації, а другий – на основі системи продукцій для формування змістовних зв'язків у тексті.

Комп'ютерна обробка знань представляє собою обробку їх змісту правилами перетворення тих форм, якими описуються знання в машині. Тому при обробці знань фундаментальною і важливою проблемою є опис змістового вмісту проблем широкого діапазону, а також наявність опису такої форми знань, яка

гарантуватиме, що їх обробка формальними правилами перетворення буде здійснюватися правильно [98].

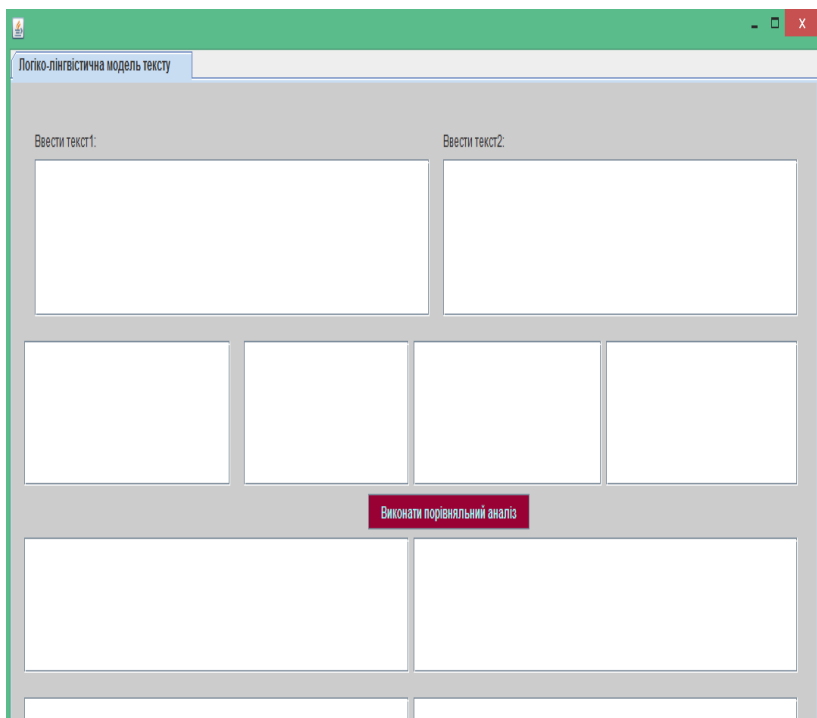


Рис. 4.2 – Інтерфейс користувача інформаційної технології порівняльного аналізу двох текстових документів

Однорідне представлення приводить до спрощення механізму управління логічним виводом та спрощення управління знаннями [99]. Показником інтелектуальності системи, з точки зору представлення знань, вважається здатність системи використовувати в потрібний момент необхідні (релевантні) знання. У проблемі доступу до знань можна виділити три аспекти:

1) зв'язність (агрегація) знань є основним способом, що забезпечує прискорення пошуку релевантних знань; знання організуються навколо найбільш важливих об'єктів (сутностей) предметної області;

2) механізм доступу до знань – деякий опис сутності, призначений для відшукування у базі знань об'єктів, що відповідають цьому опису; очевидно, що впорядкування та структурування знань може значно пришвидшити процес пошуку;

3) способи співставлення: синтаксичні, параметричні, семантичні та примусові.

У випадку синтаксичного співставлення співвідносяться форми (зразки), а не зміст об'єктів. Успішним вважається співставлення, за результатами якого зразки ідентичні. Результат синтаксичного співставлення є бінарним. У параметричному співставленні вводиться параметр, що визначає ступінь співставлення. У випадку семантичного співставлення співвідносяться не зразки об'єктів, а їх функції [100].

Лінгвістичний процесор для автоматичного формування логіко-лінгвістичних моделей (рис. 4.3) представляє собою транслятор, що встановлює відповідність між текстом та його змістом і здатен працювати з природною мовою у всьому її об'ємі [101].

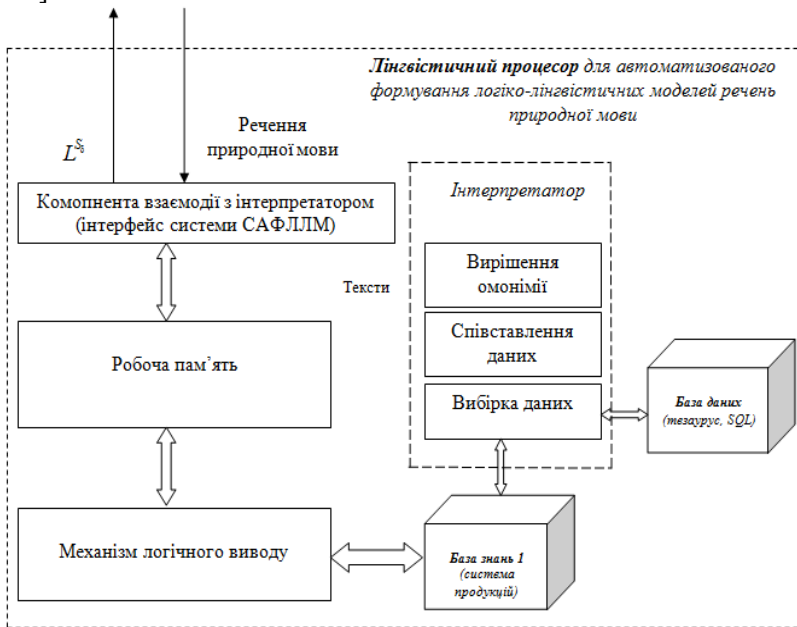


Рис. 4.3 – Структура лінгвістичного процесора для автоматичного формування логіко-лінгвістичних моделей

Дві основні функції лінгвістичного процесору полягають у вилученні змісту із заданої текстової інформації та вираженні отриманого змісту мовою логіки предикатів для можливості подальшого порівняння сформованих моделей [102]. Тобто лінгвістичний процесор для автоматичного формування логіко-лінгвістичних моделей здійснює аналіз введеної текстової інформації згідно з алгоритмом методу автоматизованого формування логіко-лінгвістичних моделей тексту, в результаті чого на виході автоматично створюються логіко-лінгвістичної моделі (1.1) – (1.4) для кожного з речень тексту.

База даних, що використовується в інформаційній технології порівняльного аналізу двох електронних текстових документів представляє собою тезаурус, словник слів природної мови, розміщених у вигляді набору реляційних таблиць в SQL.

Отже, на вхід лінгвістичного процесора для автоматичного формування логіко-лінгвістичних моделей надходять речення природної мови першого та другого електронного текстового документу, а на виході отримуємо двовимірні масиви логіко-лінгвістичних моделей заданих текстів, що представляють собою

семантико-синтаксичні складові $t'_1 = \bigwedge_{\delta_1=1}^{N(t_1)} L^{S_{\delta_1}}$ і $t'_2 = \bigwedge_{\delta_2=1}^{N(t_2)} L^{S_{\delta_2}}$. Перший

текст природної мови вводиться в поле під назвою «Ввести текст 1:», друге речення природної мови вводиться в поле під назвою «Ввести текст 2:» (рис. 4.4).

Текстова інформація повинна містити правильно написані слова української мови. Якщо слово буде написано не коректно, програма не знайде його у системі і видасть про це інформацію у вигляді характеристик. Внаслідок цього у логіко-лінгвістичній моделі будуть відсутні не ідентифіковані компоненти. Після введення речень природної мови на екрані з'являються логіко-лінгвістичні моделі введених речень у формульному вигляді та у вигляді слів (рис. 4.5).

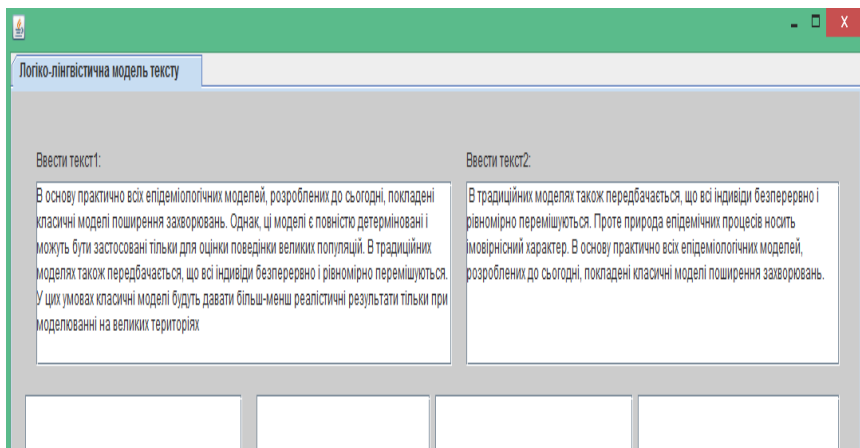


Рис. 4.4 – Приклад введення даних для опрацювання лінгвістичним процесором автоматичного формування логіко-лінгвістичних моделей

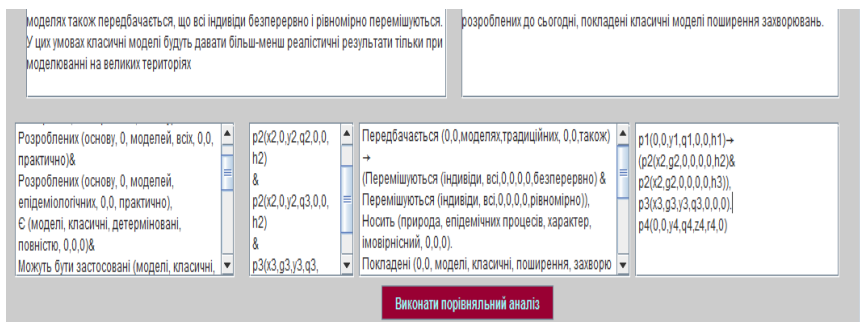


Рис. 4.5 – Результат побудови логіко-лінгвістичних моделей речень природної мови двох заданих текстів

Лінгвістичний процесор для побудови лінгвістичної складової логіко-лінгвістичної моделі кожного із заданих електронних текстових документів $t_1 = \langle CQ, F_1, B_1, A_1 \rangle$ та $t_2 = \langle CQ_2, F_2, B_2, A_2 \rangle$ представляє собою транслятор, який перетворює речення природної мови на впорядковану четвірку, складовими якої є тип тексту, множина складних синтаксичних

конструкції тексту, текстова база та множина абзаців тексту (рис. 4.6).

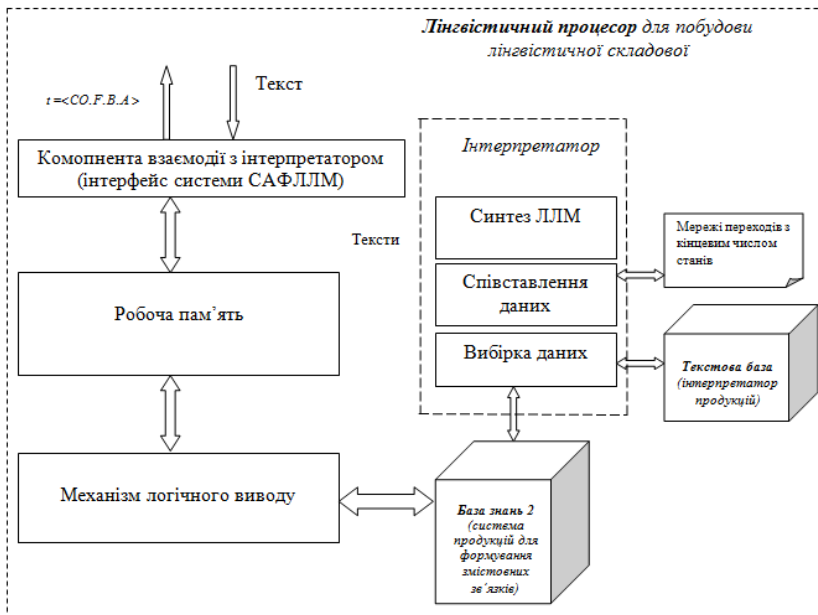


Рис. 4.6 – Структура лінгвістичного процесора для побудови лінгвістичної складової логіко-лінгвістичної моделі текстового документу

База знань, що використовується лінгвістичним процесором для побудови лінгвістичної складової логіко-лінгвістичної моделі текстового документу, представляє собою множину продукцій, що містить: правила здійснення синтезу логіко-лінгвістичних моделей; правила формування інваріантних логіко-лінгвістичних моделей, абстрактні моделі логічних зв'язків між частинами текстових документів.

Компонента порівняння семантико-синтаксичних складових та компонента порівняння лінгвістичних складових логіко-лінгвістичних моделей текстових документів відповідає другому та третьому етапам методу їх порівняння відповідно, на основі чого можна описати математичне

забезпечення інформаційної технології порівняльного аналізу двох електронних текстових документів (рис. 4.7).



Рис. 4.7 – Математичне забезпечення інформаційної технології порівняльного аналізу двох електронних текстових документів за змістом

Після натискання кнопки «Виконати порівняльний аналіз» виводиться результат із зазначенням відсотків збігу за різними компонентами логіко-лінгвістичних моделей двох заданих текстів (рис. 4.8).

Інформаційна технологія порівняльного аналізу текстових документів спроектована на основі використання таких алгоритмів та методів: алгоритму побудови логіко-лінгвістичної моделі речення природної мови, методу автоматизованого формування логіко-лінгвістичних моделей текстової інформації, алгоритму побудови логіко-лінгвістичної моделі текстового документу, методу порівняльного аналізу простих речень природної мови, методу порівняння речень природної мови довільної складності, методу порівняльного аналізу логіко-лінгвістичних моделей текстових документів.

В основу практично всіх епідеміологічних моделей, розроблених до сьогодні, покладені класичні моделі поширення захворювань. Однак, ці моделі є повністю детерміновані і можуть бути застосовані тільки для оцінки поведінки великих популяцій. В традиційних моделях також передбачається, що всі індивіди безперервно і рівномірно перемішуються. У цих умовах класичні моделі будуть давати більш-менш реалістичні результати тільки при моделюванні на великих територіях

В традиційних моделях також передбачається, що всі індивіди безперервно і рівномірно перемішуються. Проте природа епідемічних процесів носить імовірнісний характер. В основу практично всіх епідеміологічних моделей, розроблених до сьогодні, покладені класичні моделі поширення захворювань.

Розроблених (основу, 0, моделей, всіх, 0,0, практично)&	$p_2 \times 2,0, y_2, q_2, 0,0, h_2$	Передбачається (0,0, моделях, традиційних, 0,0, також) →	$p_1(0,0, y_1, q_1, 0,0, h_1) \rightarrow$
Розроблених (основу, 0, моделей, епідеміологічних, 0,0, практично), Є (моделі, класичні, детерміновані, повністю, 0,0,0)&	$p_2 \times 2,0, y_2, q_3, 0,0, h_2$	→ (Перемішуються (індивіди, всі, 0,0,0,0, безперервно) & Перемішуються (індивіди, всі, 0,0,0,0, рівномірно)). Носить (природа, епідемічних процесів, характер, імовірнісний, 0,0,0).	$p_2 \times 2, g_2, 0,0,0,0, h_2) \&$
Можуть бути застосовані (моделі, класичні).	$p_3 \times 3, g_3, y_3, q_3,$	Покладені (0,0, моделі, класичні, поширення, захворю	$p_3 \times 3, g_3, y_3, q_3, 0,0, 0) \&$ $p_4(0,0, y_4, q_4, z_4, r_4, 0)$

Виконати порівняльний аналіз

В основу практично всіх епідеміологічних моделей, розроблених до сьогодні, покладені класичні моделі поширення захворювань. В традиційних моделях також передбачається, що всі індивіди безперервно і рівномірно перемішуються.	В традиційних моделях також передбачається, що всі індивіди безперервно і рівномірно перемішуються. В основу практично всіх епідеміологічних моделей, розроблених до сьогодні, покладені класичні моделі поширення захворювань.
---	---

J(S)=40%	J(CQ)=95%
J(F)=99%, J(A)=40%	J(B)=76%

Рис. 4.8 – Результат порівняння логіко-лінгвістичних моделей двох текстів

Розробка інформаційної технології порівняльного аналізу електронних текстових документів за змістом базується на теоретичних основах синтезу та аналізу речень природної мови, що висвітлені у даній монографії.

СПИСОК ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1. Кормен Т. Алгоритмы: построение и анализ 3-е изд. / Т.Кормен, И. Чарльз, Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн.Спб., Вильямс, 2013. – 1328 с.
2. Джарратано Д. Экспертные системы: принципы разработки и программирование / Д. Джарратано, Г. Райли; пер. с англ. К.А. Птицына. – 4-е издание. – М.: ООО «Вильямс», 2007. – 1152 с.
3. Про затвердження переліку пріоритетних тематичних напрямів наукових досліджень і науково-технічних розробок на період до 2020 року [Електронний ресурс]. – Режим доступу: <http://zakon3.rada.gov.ua/laws/show/942-2011-%D0%BF>.
4. Лінгвістичі та технологічні основи тлумачної лексикографії / [В. А. Широков, В. М. Білоноженко, О. В. Бугаков та ін.]. – К.: Довіра, 2010. – 295 с.
5. Корпусна лінгвістика / [Широков В. А., Бугаков О. В., Грязнухіна Т. О. та ін.]. – К.: Довіра, 2005. – 471 с.
6. Shyrovkov V. System semantics of explanatory dictionaries / Shyrovkov V. // *Cognitive Studies| Études cognitives*. – 2015. – 12. – P. 95-106.
7. Evans V. *Cognitive Linguistics* / V. Evans, M. Green. – Edinburg: Edinburg university press Publ., 2006. – 830 p.
8. Поспелов Д. А. Логико-лингвистические модели в системах управления / Поспелов Д. А. – М.: Энергоиздат., 1981. – 232 с.
9. Осуга С. Приобретение знаний / С. Осуга, Ю. Саэки. – М.: Мир, 1990. – 304 с.
10. Вагин В. Н. Дедукция и обобщение в системе принятия решений / Вагин В. Н. – М.: Наука, 1988. – 384 с.
11. Chang C. Practical linguistic steganography using contextual synonym substitution and a novel vertex coding method/ C. Chang, S. Clark // *Computational linguistics*. – 2014. – Vol.40, №2. – P. 403–448.
12. Zhang Y. Discriminative syntax-based word ordering for text generation / Y. Zhang // *Computational linguistics*. – 2015. – Vol.41, №3. – P. 503–538.
13. Гальперин И. Р. Текст как объект лингвистического исследования / И. Р. Гальперин. – [5-е изд.]. – М.: КомКнига, 2007. – 144 с.

14. Лайонз Дж. Лингвистическая семантика: монография / Лайонз Дж. – М.: Языки славянской культуры, 2003. – 400 с.
15. Никитин М. В. Курс лингвистической семантики / Никитин М. В. – СПб.: Изд-во РГПУ им. Герцена, 2007. – 819 с.
16. Апресян Ю. Д. Лексическая семантика: в 2-х т. / Апресян Ю. Д. – М.: «Восточная литература», 1995. – Т. 1. – 422 с.
17. Чернявская З. Е. Лингвистика текста: поликодовость, интертекстуальность, интердискурсивность / Чернявская З. Е. – М.: Книжный дом «Либроком», 2009. – 248 с.
18. Рубашкин В.Ш. Представление и анализ смысла в интеллектуальных информационных системах / Рубашкин В.Ш. – М.: Наука. Гл. ред. физ.-мат. лит., 1989. – 192 с.
19. Гладкий А. В. Синтаксические структуры естественного языка в автоматизированных системах общения / Гладкий А. В. – М.: Наука. Гл. ред. физ.-мат. лит., 1985. – 144 с.
20. Jurafsky D. Natural Language Processing / D. Jurafsky, C. Manning. – Stanford University. – Available at: <https://www.coursera.org/course/nlp/> – 15.05.2017.
21. Ландэ Д. В. Интернетика: Навигация в сложных сетях: модели и алгоритмы / Ландэ Д. В., Снарский А. А., Безсуднов И. В. – М.: Либроком (Editorial URSS), 2009. – 264 с.
22. Ланде Д. В. Аналіз інформаційних потоків у глобальних комп'ютерних мережах / Ланде Д. В. – Вісник НАН України, 2017. – № 3. – С. 46–54. <https://doi.org/10.15407/visn2017.03.045>
23. Глибовець М. М. Системи штучного інтелекту: навч. посібник / М.М. Глибовець, О.В. Олецький. – К.: вид. "Км "Академія", 2002. – 366 с.
24. Глибовець А.М. Один підхід до побудови інтелектуальної пошукової системи / А.М. Глибовець, А.С. Шабінський // Наукові записки НаУКМА. – (Серія "Комп'ютерні науки"). – 2010. – Т. 112. – С. 26-29.
25. Глибовець М. М. Інтелектуальні мережі : навчальний посібник / М. М. Глибовець, А. М. Глибовець, М. В. Поляков ; Нац. ун-т "Києво-Моги́л. акад.", Корпорація "Noosphere Ventures USA Inc." – Дніпропетровськ : [Нова ідеологія], 2014. – 462 с.
26. Алгоритм извлечения информации в АБВУУ Comreno [Электронный ресурс]. – Режим доступа: <https://habrahabr.ru/company/abbyy/blog/269191/> – 15.05.2017.

27. Проверка уникальности текста Advego Plagiatus [Электронный ресурс]. – Режим доступа: <http://advego.ru/> – 15.05.2017.
28. Text.ru [Электронный ресурс]. – Режим доступа: <https://text.ru/about>. – 15.05.2017.
29. StrikePlagiarism [Электронный ресурс]. – Режим доступа:<http://strikeplagiarism.com>. – 15.05.2017.
30. Content watch [Электронный ресурс]. – Режим доступа: <https://content-watch.ru>. – 15.05.2017.
31. Плющ М.Я. Українська мова. Довідник / М.Я. Плющ, Н.Я. Грипас. – К.: Радянська школа, 1990. – 255 с.
32. Валгина Н.С. Теория текста. – М.: Логос, 2003.
33. Мацько Л.І. Українська мова: посібник / Л.І. Мацько, О.М. Сидоренко. – 2-ге вид. – К.: Либідь, 1996. – 432 с.
34. Вавіленкова А. І. Вплив функціонально-істинних операцій на зміст речень природної мови / А. І. Вавіленкова // Наукоємні технології. – 2016. – № 1(29). – С. 31 – 35.
35. Мускатов Д. В. Теорема Геделя о полноте исчисления предикатов [Электронный ресурс]. – Режим доступа: <http://lectoriy.mipt.ru/lecture/Maths-MathemLogic-L12-Musatov-141119.03>. – 15.05.2017.
36. Вавіленкова А. І. Логіко-лінгвістичні моделі речень як засіб порівняння текстових документів за змістом / А. І. Вавіленкова // Математичні машини і системи. – 2012. – № 1. – С. 166 – 173.
37. Никконен А. Ю. Устранение избыточности и дублирования сюжетов новостных сообщений / А. Ю. Никконен // Интернет-Математика. Сборник работ участников конкурса. – Екатеринбург: Изд-во Урал. ун-та, 2007. – С. 157 – 167.
38. Мельчук И.А. Опыт теории лингвистических моделей СМЫСЛ-ТЕКСТ. – М.: Языки русской культуры, 1999. – 346с.
39. Вавіленкова А. І. Извлечение смысла из предложений естественного языка / А. І. Вавіленкова // Программные продукты и системы. – 2012. – № 4 (100). – С. 87 – 90.
40. Кузнецов О.П. Дискретная математика для инженера / О.П. Кузнецов. – 6-е изд. – Спб.: Издательство "Лань", 2009. – 400 с.
41. Складне речення: види складних речень [Електронний ресурс]. – Режим доступу: <http://l-ponomar.com/skladne-rechennya-vydy-skladnyh-rechen/>. – 15.05.2017.

42. Литвиненко О. Є. Окремі форми уніфікованої логіко-лінгвістичної моделі / О. Є. Литвиненко, А. І. Вавіленкова // Міжнар. наук.-техн. конф. «Інтелектуальні технології лінгвістичного аналізу»: тези доповідей. – К.: НАУ, 2014. – С. 11.
43. Уэно Х. Представление и использование знаний / Х. Уэно, М. Исидзука. – М.: Мир, 1989. – 220 с.
44. Ланде Д.В. Поиск знаний в Интернет. Профессиональная работа / Ланде Д.В. – СПб.: Издательский дом "Диалект Вильямс", 2005. – 272 с.
45. Снитюк В.Є. Прогнозування. Моделі. Методи. Алгоритми: навч. посібник / Снитюк В.Є. – К.: "Маклаут", 2008. – 364 с.
46. Оксіюк О. Г. Модель оптимізації семантичної мережі/ Оксіюк О. Г. // Системи управління, навігації та зв'язку. – 2015. – №1. – С. 24-31.
47. Снарский А.А., Ландэ Д.В. Моделирование сложных сетей: учебное пособие. – К.: Инжиниринг, 2015. – 212 с.
48. Вавіленкова А. І. Проект комп'ютерної технології лінгвістичного аналізу електронних документів / А. І. Вавіленкова // International Scientific Journal Acta Universitatis Pontica Euxinus. Special number for X international conference «Strategy of quality in industry and education». – Varna (Bulgaria), 2014. – P. 388 – 394.
49. Глибовец Н.Н. Применение онтологий и методов текстового анализа при создании интеллектуальных поисковых систем // Н.Н. Глибовец, А. Н. Глибовец, А.С. Шабинский. – Проблемы управления и информатики. – 2011. – 6. – С. 95– 103.
50. Вавіленкова А. І. Алгоритм побудови логіко-лінгвістичної моделі текстової інформації / А. І. Вавіленкова // Електротехнічні та комп'ютерні системи. – 2014. – № 16 (92). – С. 85 – 91.
51. Філіппов К.А. Лингвистика текста. Курс лекцій. – СПб.: Издательство С.-Петербургского университета, 2008. – 336с.
52. Вавіленкова А. І. Автоматизація процесу вилучення знань з електронних документів / А. І. Вавіленкова // International Scientific Journal Acta Universitatis Pontica Euxinus. Special number for IX international conference «Strategy of quality in industry and education». – Varna (Bulgaria), 2013. – Vol. II. – P. 396 – 399.
53. Вавіленкова А. І. Алгоритм побудови математичної моделі тексту / А. І. Вавіленкова // XXI міжнар. наук.-практ. конф. «Інформаційні технології: наука, техніка, технологія, освіта,

здоров'я (MicroCAD-2013)»: тези доповідей. – Харків: НТУ "ХПІ", 2013. – Ч. IV. – С. 6.

54. Звегинцев В. А. Новое в зарубежной лингвистике. Логический анализ естественного языка: вып. 18 / Звегинцев В. А. – М.: Изд-во «Прогресс», 1986. – 392 с.

55. Вавіленкова А. І. Теоретичні основи аналізу електронних текстів: монографія / Вавіленкова А. І. – К.: ТОВ «СІК ГРУП УКРАЇНА», 2016. – 192 с.

56. Поспелов Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов / Поспелов Д.А. – М.: Радио и связь, 1989. – 114 с.

57. Гальперин И.Р. К проблеме зависимости предложения от контента / Гальперин И.Р. - 1977

58. Шемакин Ю.И. Компьютерная семантика / Ю.И. Шемакин, А.А. Романов – М.: НОЦ «Школа Китайгородской», 1995. – 344 с.

59. Греймас А.-Ж. Структурная семантика. Поиск метода / Греймас А.-Ж. – М.: Академический проект, 2004. – 368 с.

60. Кобозева И.М. Лингвистическая семантика / Кобозева И.М. – М.: Эдитореал УРСС, 2000. – 352 с.

61. Лукашевич, 2011, Когезия как структурная связность текста <http://like-money.ru/stati/271-kogeziya-kak-strukturnaya-svyaznost-teksta>

62. Головкина С. Х. Лингвистический анализ текста / С. Х. Головкина, С. Н. Смольников. – Вологда: Издательский центр ВИРО, 2006. – 124 с.

63. Маслова В. А. Когнитивная лингвистика: учеб. пособ. / Маслова В. А. – Минск: ТетраСистемс, 2008. – 272 с.

64. Вавіленкова А. И. Основные принципы синтеза логико-лингвистических моделей / А. И. Вавіленкова // Кибернетика и системный анализ. – 2015. – Т. 51, № 5. – С. 176 – 185.

65. Вавіленкова А. І. Аналіз методів обробки текстової інформації / А. І. Вавіленкова // Вісник Національного технічного університету «Харківський політехнічний університет»: зб. наук. праць. – (Серія «Інформатика та моделювання»). – 2013. – № 39 (1012). – С. 35 – 40.

66. Вавіленкова А. І. Проект комп'ютерної технології лінгвістичного аналізу електронних документів / А. І. Вавіленкова // International Scientific Journal Acta Universitatis Pontica Euxinus.

Special number for X international conference «Strategy of quality in industry and education». – Varna (Bulgaria), 2014. – P. 388 – 394.

67. Томан Іржі Мистецтво говорити / Томан Іржі; пер з чеської В.І. Романця. – К.: Полтвидав України. – 1986. – 221 с.

68. Довженко О. П. Зачарована Десна [Електронна копія] / О. Довженко. – Х.: Фоліо, 2012. – (Шкільна бібліотека української та світової літератури).

69. Нечуй-Левицький І.С. Микола Джеря: повість / Нечуй-Левицький І.С. – К.: Радянська Україна, 1990. – 93 с.

70. Бухараев Р. Г. Семантический анализ в вопросно-ответных системах / Р. Г. Бухараев, Д. Ш. Сулейманов. – Казань: Издательство Казанского университета, 1990. – 124 с.

71. Васильев Л. М. Современная лингвистическая семантика / Васильев Л. М. – М.: Высш. шк., 1990. – 176 с.

72. Вавіленкова А. І. Синтез логіко-лінгвістичних моделей речень природної мови як засіб побудови змістовної моделі тексту / А. І. Вавіленкова // Системи підтримки прийняття рішень. Теорія і практика: зб. доп. наук.-практ. конф. з міжнар. участю. – Київ: ПММС НАНУ, 2013. – С. 49 – 51.

73. Гончар О. Перекоп / Гончар О. – К., 2011. – ТЗ. – С. 289 – 646.

74. Вавіленкова А. І. Принципи побудови семантичних графів / А. І. Вавіленкова // Міжнар. наук.-практ. конф. «Інтелектуальні технології лінгвістичного аналізу»: тези доповідей. – К.: НАУ, 2012. – С. 8.

75. Вавіленкова А. І. Формування семантичних зв'язків у тексті / А. І. Вавіленкова // Тези всеукр. наук.-метод. конф. студентів та молодих науковців «Прикладна лінгвістика 2014: проблеми і рішення»: тези доповідей. – Миколаїв: НУК, 2014. – С. 18 – 20.

76. Вавіленкова А. І. Правила синтезу логіко-лінгвістичних моделей речень природної мови / А. І. Вавіленкова // Наукоємні технології. – 2017. – № 1(33). – С. 3 – 7.

77. Стельмах М. Гуси-лебеді летять / Стельмах М. [Електронний ресурс]. – Режим доступу: <http://www.ukrlib.com.ua/books/printit.php?tid=323>.

78. Geeraerts, D. Cognitive linguistics: basic readings research / Geeraerts D., Dirven R., John R. – Berlin-New York: Publ. mouton de cruyter, 2006. – 486 p.

79. Jiawei, H. Data Mining: Concepts and Techniques / H. Jiawei, K. Micheline // Morgan Kaufmann Publishers. – 2000.

80. Gries, S. Th. Corpus-based methods and cognitive semantics: the many meanings of to run / S. Th. Gries // Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis. – 2006. – P. 57–99.

81. Geeraerts, D. Cognitive linguistics: basic readings research / Geeraerts D., Dirven R., John R. – Berlin-New York: Publ. mouton de cruyter, 2006. – 486 p.

82. Вавіленкова А. І. Критерії аналізу логіко-лінгвістичних моделей речень природної мови / А. І. Вавіленкова // Вісник Національного технічного університету «Харківський політехнічний університет»: зб. наук. праць. – (Серія «Нові рішення в сучасних технологіях»). – 2017. – № 7 (1229). – С. 118 – 122.

83. Вавіленкова А. І. Аналіз методів пошуку синонімів в електронних документах / А. І. Вавіленкова // Вісник Чернігівського державного технологічного університету: зб. наук. праць. – (Серія «Технічні науки»). – 2014. – № 2 (73). – С. 119 – 128.

84. Хант Э. Искусственный интеллект / Хант Э. – М.: Мир, 1978. – 560 с.

85. Золотова Г.А., Онипенко Н.К., Сидорова М.Ю. Коммуникативная грамматика русского языка. М., 1988

86. Вавіленкова А. І. Умови тотожності логіко-лінгвістичних моделей простих речень природної мови / А. І. Вавіленкова // Вісник Національного технічного університету «Харківський політехнічний університет»: зб. наук. праць. – (Серія «Інформатика та моделювання»). – 2015. – № 32 (1141). – С. 27 – 35.

87. Беляков И. А. Применение интеллектуальных технологий в процессе сертификации программного обеспечения / И. А. Беляков, М. А. Еремеев // Молодой ученый. – 2011. – Т.1, №11. – С. 23 –31.

88. Методы и модели анализа данных: OLAP и Data Mining / [Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И.]. – Спб.: БХВ-Петербург, 2007. – 384 с.

89. Вавіленкова А. І. Пошук змісту висловлювань через побудову логіко-лінгвістичних моделей / А. І. Вавіленкова //

Дев'ята міжнар. наук.-практ. конф. «Математичне та імітаційне моделювання систем. МОДС 2014»: тези доповідей. – Чернігів: ЧДІЕУ, 2014. – С. 263 – 266.

90. Вавіленкова А. І. Основи автоматичної обробки електронних текстів при боротьбі з рерайтом / А. І. Вавіленкова // Наукоємні технології. – 2017. – № 2(34). – С. 97 – 101.

91. Рерайтинг [Електронний ресурс]. – Режим доступу: http://uniofweb.ru/wiki/rewriting/#full_version.

92. Величковский Б.М. Значение в глобальных когнитивных моделях Б.М. Величковский. Современная когнитивная психология. М., МГУ, 1982. С. 206-221.

93. Vavilenkova A. I. The formal model of knowledge retrieval and processing / A. I. Vavilenkova // Black Sea Scientific journal of academic research. – (Series «Technical and applied sciences»). – 2014. – Vol.16, N 09. – P. 115 – 119.

94. Вавіленкова А. І. Побудова змістовної моделі тексту на основі використання логіко-лінгвістичних моделей / А. І. Вавіленкова // Вісник Національного університету «Львівська політехніка». – (Серія «Комп'ютерні науки та інформаційні технології»). – 2015. – № 826. – С. 169 – 175.

95. Коцюба И.Ю. Основы проектирования информационных систем: учеб. пособ./ И.Ю. Коцюба, А.В. Чунаев, А.Н. Шиков. – Спб.: Университет ИТМО, 2015. – 206 с.

96. Вавіленкова А.І. Виявлення компонентів змісту в реченнях природної мови /А.І. Вавіленкова // Сьома наук.-практ. конф. з міжнар. участю «Математичне та імітаційне моделювання систем. МОДС 2012»: тези доповідей. – Чернігів, 2012. – С. 240–242.

97. Вавіленкова А. І. Структура інформаційної технології порівняльного аналізу текстових документів / А. І. Вавіленкова // Технічні науки та технології. –2016. – № 1 (3). – С. 103 – 109.

98. Вавіленкова А. І. Логіко-лінгвістичні моделі речень як засіб порівняння текстових документів за змістом / А. І. Вавіленкова // Математичні машини і системи. – 2012. – № 1. – С. 166 – 173.

99. Вавіленкова А. І. Програмне забезпечення для виявлення текстових документів, ідентичних за змістом / А. І. Вавіленкова // Вісник Чернігівського державного технологічного університету: наук. зб. – (Серія «Технічні науки»). – 2013. – № 2 (65). – С. 125 – 132.

100. Вавіленкова А. І. Система порівняльного аналізу текстових документів / А. І. Вавіленкова // Вісник Житомирського державного технологічного університету: зб. наук. праць. – (Серія «Технічні науки»). – 2015. – № 4 (75). – С. 94 – 100.

101. Вавіленкова А. І. Структура лінгвістичного процесору системи порівняльного аналізу текстів за змістом / А. І. Вавіленкова // Системи підтримки прийняття рішень. Теорія і практика: зб. доп. наук.-практ. конф. з міжнар. участю. – Київ: ІПММС НАНУ, 2011. – С. 153 – 156.

102. Вавіленкова А. І. Інформаційна технологія обробки текстової інформації на основі побудови логіко-лінгвістичних моделей / А. І. Вавіленкова // International Scientific Journal Acta Universitatis Pontica Euxinus. Special number for XI international conference «Strategy of quality in industry and education». – Varna (Bulgaria), 2015. – Vol. II. – P. 377 – 380.

Наукове видання

БАВІЛЕНКОВА Анастасія Ігорівна

**АНАЛІЗ І СИНТЕЗ ЛОГІКО-ЛІНГВІСТИЧНИХ
МОДЕЛЕЙ РЕЧЕНЬ ПРИРОДНОЇ МОВИ**

Монографія
(Українською мовою)

В авторській редакції
Комп'ютерна верстка Бумажний В.П.

Підп. до друку 12.06.17. Формат 60×84/16.
Друк офс. Ум. друк. арк.. 10,3. Обл.-вид. арк. 9.
Тираж 50 пр. Замовлення № 275.

Віддруковано у ТОВ «Наш формат»,
02105, м. Київ, пр. Миру, 7

Свідоцтво про внесення суб'єкта видавничої справи до Державного
реєстру видавців, виготівників і розповсюджувачів видавничої продукції
ДК № 4540 від 07.05.2013 р.